Geoscientific
Model Development

Competition between plant functional types in the Canadian Terrestrial Ecosystem Model (CTEM) v. 2.0

J. R. Melton[1] and V. K. Arora[2]

[1]Climate Research Division, Environment Canada, Victoria, BC, V8W 2Y2, Canada
[2]Canadian Centre for Climate Modelling and Analysis, Climate Research Division, Environment Canada, Victoria, BC, V8W 2Y2, Canada

Correspondence to: J. R. Melton (joe.melton@canada.ca)

Abstract. The Canadian Terrestrial Ecosystem Model (CTEM) is the interactive vegetation component in the Earth system model of the Canadian Centre for Climate Modelling and Analysis. CTEM models land–atmosphere exchange of CO_2 through the response of carbon in living vegetation, and dead litter and soil pools, to changes in weather and climate at timescales of days to centuries. Version 1.0 of CTEM uses prescribed fractional coverage of plant functional types (PFTs) although, in reality, vegetation cover continually adapts to changes in climate, atmospheric composition and anthropogenic forcing. Changes in the spatial distribution of vegetation occur on timescales of years to centuries as vegetation distributions inherently have inertia. Here, we present version 2.0 of CTEM, which includes a representation of competition between PFTs based on a modified version of the Lotka–Volterra (L–V) predator–prey equations. Our approach is used to dynamically simulate the fractional coverage of CTEM's seven natural, non-crop PFTs, which are then compared with available observation-based estimates. Results from CTEM v. 2.0 show the model is able to represent the broad spatial distributions of its seven PFTs at the global scale. However, differences remain between modelled and observation-based fractional coverage of PFTs since representing the multitude of plant species globally, with just seven non-crop PFTs, only captures the large-scale climatic controls on PFT distributions. As expected, PFTs that exist in climate niches are difficult to represent either due to the coarse spatial resolution of the model, and the corresponding driving climate, or the limited number of PFTs used. We also simulate the fractional coverage of PFTs using unmodified L–V equations to illustrate its limitations.

The geographic and zonal distributions of primary terrestrial carbon pools and fluxes from the versions of CTEM that use prescribed and dynamically simulated fractional coverage of PFTs compare reasonably well with each other and observation-based estimates. The parametrization of competition between PFTs in CTEM v. 2.0 based on the modified L–V equations behaves in a reasonably realistic manner and yields a tool with which to investigate the changes in spatial distribution of vegetation in response to future changes in climate.

1 Introduction

Dynamic global vegetation models (DGVMs) are now considered an integral component of Earth system models (ESMs). DGVMs model vegetation as a dynamic component of the Earth system, allowing simulation of the atmosphere–land flux of CO_2 and other trace gases. A corresponding ocean carbon cycle component in ESMs simulates the atmosphere–ocean flux of CO_2. Together they allow for modelling of the atmospheric CO_2 concentration as a prognostic variable in ESMs and quantification of the biogeochemical feedbacks that affect the carbon cycle and physical climate system (Arora et al., 2013; Arora and Boer, 2014). Modelling vegetation as a dynamic component of the climate system also permits simulation of biophysical feedbacks. Land surface vegetation in ESMs responds to variations in weather and climate and, in turn, influences the climate, on hourly to centennial timescales. Over short periods (hourly to daily), the simulated latent heat flux is affected by changes in veg-

etation's stomatal conductance, which controls the rate of transpiration. On seasonal timescales, changes in leaf area index (LAI) influence the land surface albedo and the transpiration and evaporation of intercepted precipitation from canopy leaves. On timescales from inter-annual to centennial, changes in structural attributes of vegetation such as LAI, vegetation height, rooting depth and, in particular, the spatial distribution of vegetation, affect the land surface albedo and various components of the land surface energy and water balance. These biophysical interactions between vegetation and climate are well documented (Wang et al., 2011; Kim and Wang, 2007; Gobron et al., 2010).

DGVMs typically represent vegetation in terms of a modest number of plant functional types (PFTs). This simplification is justified by the ability to categorize plant species on the basis of their form and interaction with the environment (Box, 1996). For example, fir (*Abies*), spruce (*Picea*) and pine (*Pinus*) can all be classified as needleleaf evergreen trees. From a purely practical standpoint, PFTs are used because it is impossible, at present, to obtain model parameters for the thousands of plant species based on the sparse available data.

In its simplest form, the dynamic behaviour of vegetation can be described by two aspects: changes in plant structure and changes in areal coverage. First, as vegetation responds to changes and variability in climate, its structure changes affecting its height, LAI and rooting depth, typically on seasonal to decadal timescales. Second, vegetation also adapts by changing its areal extent. Areal changes are slower than the structural changes, typically occurring on decadal to centennial timescales for woody vegetation and sub-decadal to decadal timescales for herbaceous vegetation. Together, these changes in structure and areal extent capture vegetation's dynamic behaviour in response to changes and variability in climate at different timescales.

DGVMs may incorporate only the structural, or both the structural and areal, aspects of vegetation dynamics. The Canadian Terrestrial Ecosystem Model (CTEM) (Arora, 2003; Arora and Boer, 2005a; Arora et al., 2009; Melton and Arora, 2014), for example, has so far incorporated only structural vegetation dynamics, i.e. it uses prescribed fractional coverage of its nine PFTs (seven natural and two crop PFTs; see Table 1). The fractional coverage of PFTs in CTEM is based on a reconstruction of historical land cover from 1850 to present as described in Wang et al. (2006) and Arora and Boer (2010). Other DGVMs with prescribed fractional coverage of their PFTs include the Community Land Model version 4 (CLM4) (Lawrence et al., 2011) and the Terrestrial Ecosystem Model (TEM) (Tian et al., 1999). Obviously, this approach does not account for changes in plant distribution due to changing climate or atmospheric CO_2 concentration. Examples of DGVMs that account for both structural and areal vegetation dynamics include Top-down Representation of Interactive Foliage and Flora Including Dynamics (TRIFFID), which is currently implemented in the Hadley Centre's Earth System Model (HadGEM2-ES) (Jones et al., 2011), and the Lund–Potsdam–Jena (LPJ) DGVM (Sitch et al., 2003).

Removing the constraint of specified fractional coverage of PFTs and simulating the areal vegetation dynamics adds another degree of freedom to DGVMs. Simulating both structural and areal vegetation dynamics in a realistic manner is a more stringent test of a DGVM's abilities. An incorrect simulation of the geographical distribution of PFTs will lead to a similarly flawed distribution of primary terrestrial carbon pools and fluxes, regardless if the model correctly simulates the structural vegetation dynamics.

Modelling areal dynamics of vegetation requires simulating competition for available space and resources between PFTs. In the real world, plants compete for space to acquire both aboveground (light) and belowground (water and nutrients) resources. Attempts to capture competition between plants to allow these interactions have generally been accomplished using three different kind of models: (i) theoretical models, (ii) *gap* models, and (iii) DGVM-based models.

Theoretical models (e.g. Tilman, 1982 and Pacala and Tilman, 1994) generally relate mortality and colonization to bulk parameters that are not easily related to physiological traits or mechanisms (although Pacala and Tilman, 1994, allow some mechanistic approaches). Gap models, which attempt to explicitly represent plant growth that occurs following the creation of a forest gap, provide a more mechanistic approach that simulates resource competition between individual trees at the spatial scale of a few metres (e.g. Bugmann, 2001). There have been attempts to integrate gap model dynamics into DGVMs (e.g. the Spatially Explicit Individual-Based DGVM (SEIB-DGVM), Sato et al., 2007; the LPJ General Ecosystem Simulator (LPJ-GUESS), Smith et al., 2001; and the Ecosystem Demography (ED) model approach, Moorcroft et al., 2001). At the large spatial scales at which DGVMs are typically implemented (0.5 to 4° resolution), gap model dynamics impose a high computational load and thus have generally been applied only at local and regional scales (e.g. eastern North America in Fisher et al., 2015). An exception is that SEIB-DGVM has been globally applied in the Kyousei2 model albeit with difficulties representing the global vegetation distribution (for e.g. see Fig. 13 in Sato et al., 2007). DGVM-based competition schemes are either based on (i) the Lotka–Volterra (L–V) equations (e.g. TRIFFID, Cox, 2001; University of Sheffield Conifer Model (USCM), Brentnall et al., 2005; and CTEM, Arora and Boer, 2006a, b), (ii) simpler approaches that assume PFTs occupy area in proportion to their net primary productivity (NPP) (e.g. LPJ, Sitch et al., 2003; the Organizing Carbon and Hydrology in Dynamic Ecosystems model (ORCHIDEE), Krinner et al., 2005; Zhu et al., 2015) or (iii) account for competitiveness via both NPP and mortality (e.g. Jena Scheme for Biosphere–Atmosphere Coupling in Hamburg (JSBACH); Brovkin et al., 2009). These latter DGVM-based approaches generally suffer from an amplified expression of dominance,

Table 1. Parameters used in the competition module of CTEM v. 2.0. S_{sap} is the multiplying factor for converting biomass density to sapling density (Eq. 15 in Sect. 2.1.2). LAI_{min} is the minimum LAI below which a PFT does not expand and LAI_{max} is the maximum LAI above which a PFT expands at a maximum rate (Eqs. 17 and 18; Sect. 2.1.2). A_{max} is the PFT maximum age, which is used in the calculation of intrinsic mortality (Eq. 20; Sect. 2.1.3). $m_{ge,max}$ is the maximum growth-related mortality used in Eq. (21) for calculating growth-related mortality. Some parameter values differ between the model versions when PFTs' fractional coverage is specified and when they are dynamically modelled using competition between PFTs. If modified, the parameter values for the dynamically modelled version are shown in parentheses.

Plant functional type (PFT)	PFT short name	S_{sap} –	LAI_{min} $m^2 m^{-2}$	LAI_{max} $m^2 m^{-2}$	A_{max} yr	$m_{ge,max}$ yr^{-1}
Needleleaf evergreen trees	NDL-EVG	0.32	1.00	4.00	800	0.005
Needleleaf deciduous trees	NDL-DCD	0.20	1.00	3.00	500	0.005
Broadleaf evergreen trees	BDL-EVG	0.08	1.50	6.00	700	0.005
Broadleaf cold deciduous trees	BDL-DCD-COLD	0.14	1.00	5.00	450	0.005
Broadleaf drought/dry deciduous trees	BDL-DCD-DRY	0.13	1.00	5.00	500	0.005
C_3 crop*	C3-CROP	–	–	–	–	–
C_4 crop*	C4-CROP	–	–	–	–	–
C_3 grass	C4-GRASS	0.20	0.01	4.00	–	0.00 (0.05)
C_4 grass	C4-GRASS	0.20	0.01	4.00	–	0.00 (0.10)

* Prescribed fractional cover.

i.e. the dominant PFT ends up occupying a disproportionately large fraction of a grid cell. As a result, forests are simulated in regions, which, in reality, are largely savannas (e.g. see Fig. 2 in Cramer et al., 2001, for Hybrid, TRIFFID, LPJ and the Integrated Biosphere Simulator (IBIS) models).

Here, we evaluate CTEM v. 2.0, which explicitly models the competition for space between PFTs using a modified version of the L–V equations (Lotka, 1925; Volterra, 1926) to interactively determine the fractional coverage of seven non-crop PFTs as a function of climate and soils. The competition scheme used here was presented in Arora and Boer (2006a, b), but this study is the first evaluation of the scheme on a global scale. Section 2 summarizes CTEM v. 2.0 and includes a description of its competition parametrization based on a modified version of the L–V equations. Section 3 describes the experimental set-up and observation-based data sets used for model evaluation. Results from simulations with prescribed and dynamically simulated fractional coverage of CTEM's seven non-crop PFTs, along with discussion, are presented in Sect. 4. We also present results from simulations that use an unmodified version of the L–V equations for comparison. Simulations with unmodified L–V equations highlight the tendency of the unmodified L–V relations to simulate excessive coverage of dominant PFTs. Finally, conclusions are presented in Sect. 5. A detailed model description of CTEM v. 2.0 is provided in the Appendix.

2 CTEM structure and processes

Version 1 of the CTEM is the terrestrial carbon cycle component of the second generation Canadian Earth System Model (CanESM2) (Arora et al., 2011) where it is coupled to version 2.7 of the Canadian Land Surface Scheme (CLASS).

CTEM v. 2.0, used here, is currently coupled to the CLASS v. 3.6 (Verseghy, 2012). The coupled CLASS–CTEM model is capable of being run online in the CanESM model or offline, as is the case in this study, driven by observation-based meteorological forcings. CTEM models terrestrial ecosystem processes for nine PFTs, two of which are crop PFTs (see Table 1), by tracking the flow of carbon through three living vegetation components (leaves, stem and roots) and two dead carbon pools (litter and soil). The carbon budget equations for the model's five pools are summarized in Sect. A7 of the Appendix. The amount of carbon in these five carbon pools is simulated prognostically. CLASS models the land surface energy and water balance and calculates liquid and frozen soil moisture, and soil temperature for three soil layers (with thicknesses 0.1, 0.25 and 3.75 m). In the CLASS–CTEM framework, CLASS uses structural vegetation attributes (including LAI, vegetation height, canopy mass and rooting depth) simulated by CTEM, and CTEM uses soil moisture, soil temperature and net radiation calculated by CLASS. Combined, CLASS and CTEM simulate the atmosphere–land fluxes of energy, water and CO_2.

Version 1.0 of CTEM is described in a collection of papers detailing parametrization of photosynthesis, autotrophic and heterotrophic respiration (Arora, 2003); phenology, carbon allocation, biomass turnover and conversion of biomass to structural attributes (Arora and Boer, 2005a); dynamic root distribution (Arora and Boer, 2003); and disturbance (fire) (Arora and Boer, 2005b). These processes are modelled over prescribed fractional coverage of nine PFTs (Wang et al., 2006) and determine the structural vegetation dynamics including vegetation biomass, LAI, vegetation height, fraction of roots in each of the three soil layers, leaf onset and offset times and primary CO_2 fluxes of gross primary productivity

(GPP) and NPP. A full description of CTEM and changes from v. 1.0 to v. 2.0 are included in the Appendix.

A parametrization for competition between PFTs in an earlier version of CTEM is described by Arora and Boer (2006a, b) where it was evaluated at select locations. Here we present CTEM v. 2.0, which builds upon the model framework of CTEM v. 1.0 and can be run in two different modes, either (i) using specified fractional coverage of its nine PFTs, or (ii) allowing the fractional coverage of its seven non-crop PFTs to be dynamically determined based on competition between PFTs. The parametrization for simulating competition between PFTs is summarized in Sect. 2.1. The fire parametrization has also been refined in the new model version as described in Appendix A9. The CLASS–CTEM modelling framework has the capability of representing the subgrid scale variability of PFTs using either a composite or a mosaic configuration (Li and Arora, 2012; Melton and Arora, 2014). In the composite (or single tile) configuration, the vegetation attributes for all PFTs present in a grid cell are averaged and used in energy and water balance calculations that determine the physical land surface conditions including soil moisture, soil temperature and thickness and fractional coverage of snow (if present). In the mosaic (or multi-tile) configuration each PFT is allocated its own tile for which separate energy and water balance calculations are performed. As a result, the simulated carbon balance evolves somewhat differently in the two configurations despite being driven with identical climate forcing (see Melton and Arora, 2014). The results presented in this paper are obtained using the composite configuration.

2.1 Representation of competition between PFTs in CTEM

Competition between PFTs in CTEM is based upon modified L–V equations (Arora and Boer, 2006a, b). The L–V equations (Lotka, 1925; Volterra, 1926) have been adapted from their initial application for simulating predator–prey interactions in ecosystem models as described below.

2.1.1 Competition parametrization

The change in fractional coverage (f) of a PFT α through time, $\frac{df_\alpha}{dt}$, is expressed as the result of mortality, and competition and colonization (CC) interactions with the other PFTs present in a grid cell and bare ground, collectively represented as B where $\alpha \notin B$:

$$\frac{df_\alpha}{dt} = g(f_\alpha, f_B) - m_\alpha f_\alpha. \tag{1}$$

The CC interactions are represented symbolically by the $g(f_\alpha, f_B)$ function. Mortality is assumed to be proportional to the number density of plants and represented by the mortality term, $m_\alpha f_\alpha$. The PFT-dependent mortality rate (m_α; day^{-1}) (described further in Sect. 2.1.3) produces bare

ground via a number of processes, and that bare ground is subsequently available for colonization. We consider the fractional coverage for N PFTs plus bare ground ($f_{N+1} = f_{\text{bare}}$) where $\sum_{j=1}^{N+1} f_j = 1$. For competition between unequal competitors, the PFTs are ranked in terms of their dominance. If PFT α is the most dominant, it will invade the area of other PFTs and the bare ground ($f_B, \alpha \notin B$). Woody PFTs are all more dominant than grass PFTs since trees can successfully invade grasses by overshading them (Siemann and Rogers, 2003) and thus are ranked higher. Within tree or grass PFTs the dominance rank of a PFT is calculated based upon its colonization rate (c_α; day^{-1}) with higher colonization rates giving a higher dominance ranking. For the general case of PFT α with a dominance rank of i, we describe the ranking from most dominant to least as 1, 2, ..., $i-1$, i, $i+1$, ..., N. Equation (1) can then be reformulated following a phenomenological approach as

$$\begin{aligned}
\frac{df_\alpha}{dt} &= f_\alpha^b(c_{\alpha,i+1} f_{i+1} + c_{\alpha,i+2} f_{i+2} + \ldots + c_{\alpha,N} f_N) \\
&\quad - f_\alpha(c_{1,\alpha} f_1^b + c_{2,\alpha} f_2^b + \ldots + c_{(i-1),\alpha} f_{i-1}^b) \\
&\quad - m_\alpha f_\alpha,
\end{aligned} \tag{2}$$

where the exponent b is an empirical parameter, which controls the behaviour of the L–V equations. In the original L–V formulation, b is 1, but we modify the L–V relations by using $b = 0$ following Arora and Boer (2006a, b) (implications of this choice are expanded upon below). The fractional cover of PFT α then changes depending on the gains it makes into the area of less dominant PFTs and the losses it suffers due to mortality and encroachment by more dominant PFTs. The rate of change of the bare fraction, f_{bare}, is expressed as

$$\frac{df_{\text{bare}}}{dt} = \sum_{\beta=1}^{N} (m_\beta f_\beta - c_{\beta,\text{bare}} f_\beta^b f_{\text{bare}}). \tag{3}$$

The rate at which PFT α invades another PFT β is given by

$$c_{\alpha,\beta} f_\alpha^b f_\beta = c_\alpha \left(\frac{c_{\alpha,\beta}}{c_\alpha} \right) f_\alpha^b f_\beta = c_\alpha \delta_{\alpha,\beta} f_\alpha^b f_\beta. \tag{4}$$

A PFT invading bare ground has an unimpeded *invasion* rate, c_α. The ratio of the invasion rate by PFT α into area covered by another PFT β and its unimpeded invasion rate ($\frac{c_{\alpha,\beta}}{c_\alpha}$) gives the relative efficiency of colonization, termed $\delta_{\alpha,\beta}$, which is a scalar between 0 and 1. δ is 1 for invasion of any PFT into bare ground and 1 for tree PFT invasion into grass PFTs. If a PFT β has a lower dominance ranking than another PFT α then $\delta_{\beta,\alpha} = 0$ implying that sub-dominant PFTs do not invade dominant PFTs, but get invaded by them, i.e. $\delta_{\alpha,\beta} = 1$. Equation (2) can then be written more succinctly for each PFT as

$$\frac{df_\alpha}{dt} = \sum_{\beta=1}^{N+1} (c_\alpha \delta_{\alpha,\beta} f_\alpha^b f_\beta - c_\beta \delta_{\beta,\alpha} f_\alpha f_\beta^b) - m_\alpha f_\alpha. \tag{5}$$

The value of parameter b is related to the manner in which two PFTs interact, represented by $f_\alpha^b f_\beta$, in Eqs. (2)–(4). As a result, the value of b affects the equilibrium solution for fractional coverage of PFTs as well as how f_i evolves over time.

For the usual form of the L–V equations with $b = \delta = 1$, and for the case of a grid cell with two PFTs, the competition–colonization equations are

$$\frac{df_1}{dt} = c_1 f_1 (f_2 + f_{\text{bare}}) - m_1 f_1$$
$$= c_1 f_1 (1 - f_1) - m_1 f_1, \tag{6}$$

$$\frac{df_2}{dt} = c_2 f_2 f_{\text{bare}} - c_1 f_1 f_2 - m_2 f_2$$
$$= c_2 f_2 (1 - f_1 - f_2) - c_1 f_1 f_2 - m_2 f_2, \tag{7}$$

where the dominant PFT 1 invades PFT 2 and the bare fraction, and PFT 2 invades only the bare fraction. The equilibrium solutions for f_1 and f_2 in this case are

$$f_1 = \max\left[\frac{c_1 - m_1}{c_1}, 0\right], \tag{8}$$

$$f_2 = \max\left[\frac{c_2 - c_2 f_1 - c_1 f_1 - m_2}{c_2}, 0\right]$$
$$= \max\left[\frac{(c_2 - m_2) - (1 + \frac{c_2}{c_1})(c_1 - m_1)}{c_2}, 0\right], \tag{9}$$

In Eq. (9), as long as $(c_1 - m_1) > (c_2 - m_2)$ the equilibrium solution for f_2 will always be zero and coexistence is not possible.

For $b = 0$ and $\delta = 1$, the competition–colonization equations are

$$\frac{df_1}{dt} = c_1 (f_2 + f_{\text{bare}}) - m_1 f_1$$
$$= c_1 (1 - f_1) - m_1 f_1, \tag{10}$$

$$\frac{df_2}{dt} = c_2 f_{\text{bare}} - c_1 f_2 - m_2 f_2$$
$$= c_2 (1 - f_1 - f_2) - c_1 f_2 - m_2 f_2, \tag{11}$$

and the corresponding equilibrium fractions are

$$f_1 = \frac{c_1}{c_1 + m_1}, \tag{12}$$

$$f_2 = \frac{c_2 (1 - f_1)}{(c_1 + c_2 + m_2)}. \tag{13}$$

In Eqs. (12) and (13), as long as $m_1 > 0$ and $c_2 > 0$, then PFT 2 will always exist and equilibrium coexistence is possible. Values of parameter b between 1 and 0 yield equilibrium values of f_2 that vary between 0 (Eq. 9) and those obtained using Eq. (13). $b = 0$ yields a maximum value of equilibrium f_2 allowing PFT 2 to coexist maximally.

In the standard L–V equations for predator–prey interactions coexistence is possible because the predator depends on prey for its food and so the predator population suffers as the prey population declines. This is in contrast to the application of the equations for competition between PFTs where the dominant PFT does not depend on sub-dominant PFTs for its existence and is thus able to exclude them completely. The PFTs interact with each other through the invasion term $(-c_\beta \delta_{\beta,\alpha} f_\alpha f_\beta^b)$ in Eq. (5), where $\delta_{\alpha,\beta} = 1$ or 0 depending on whether PFT α can or cannot invade PFT β, respectively, as mentioned earlier. This interaction through invasion is represented by $-c_1 f_1 f_2$ in Eq. (7) (for $b = 1$) and by $-c_1 f_2$ in Eq. (11) (for $b = 0$). The magnitude of this interaction thus depends on the value of parameter b. When $b = 1$ the interaction is proportional to the product of the fractional coverage of the two PFTs ($f_1 f_2$). When $b = 0$, the interaction is proportional to the fractional coverage of the PFT being invaded (f_2). The use of $b = 0$ thus reduces the product term $f_\alpha^b f_\beta$ to f_β and implies that the invasion of sub-dominant PFT β does not depend on the current fractional coverage of the dominant PFT α. This case may be thought of as corresponding to the general availability of the seeds of the dominant PFT α that may germinate and invade the coverage of the sub-dominant PFT β provided the climate is favourable, even if PFT α does not exist in the grid cell, i.e. $f_\alpha = 0$ (in the case where $f_\alpha = 0$, the PFT is always assumed to have a dormant seed bank in the grid cell given the long lifetimes of seeds and their wide dispersion). In contrast, in the standard version of the L–V equations, as implemented for predator–prey interactions, b always equals 1 since the amount of predation, and hence the reduction in the number of prey, depends on the product of the number of predators and the number of prey. Using $b = 0$ is thus consistent with invasion of the sub-dominant PFT β being unaffected by the fractional coverage of the dominant PFT α.

2.1.2 Colonization rate

The PFT-dependent colonization rate (c_α; day^{-1}) is calculated based on the fraction (Λ_α) of positive NPP (kg C m^{-2} day^{-1}) that is used for spatial expansion

$$c_\alpha = \Lambda_\alpha \, \text{NPP}_\alpha \, \xi_\alpha, \tag{14}$$

where ξ_α ((kg C)$^{-1}$ m^2) is the inverse sapling density calculated as the reciprocal of vegetation biomass ($C_{\text{veg},\alpha}$; kg C m^{-2}) multiplied by a PFT-dependent constant ($S_{\text{sap},\alpha}$; unitless; see Table 1)

$$\xi_\alpha = \frac{1}{S_{\text{sap},\alpha} \max[0.25, \min(5.0, C_{\text{veg},\alpha})]}. \tag{15}$$

The fraction of NPP used for spatial expansion, Λ_α, is calculated using the leaf area index (LAI$_\alpha$; m^2 leaf (m^2 ground)$^{-1}$) of a PFT

$$\Lambda_\alpha = \min(\lambda_{\max}, \max(\lambda_{1,\alpha}, \lambda_{2,\alpha})), \tag{16}$$

if $\mathrm{LAI}_\alpha \le \mathrm{LAI}_{\mathrm{min},\alpha}$:

$\quad \lambda_{1,\alpha} = 0$

if $\mathrm{LAI}_{\mathrm{min},\alpha} < \mathrm{LAI}_\alpha < \mathrm{LAI}_{\mathrm{max},\alpha}$:

$\quad \lambda_{1,\alpha} = \dfrac{\mathrm{LAI}_\alpha - \mathrm{LAI}_{\mathrm{min},\alpha}}{\mathrm{LAI}_{\mathrm{max},\alpha} - \mathrm{LAI}_{\mathrm{min},\alpha}} \lambda_{\mathrm{max}}$

if $\mathrm{LAI}_\alpha \ge \mathrm{LAI}_{\mathrm{max},\alpha}$:

$\quad \lambda_{1,\alpha} = \lambda_{\mathrm{max}}$ $\hspace{3cm}$ (17)

if $\mathrm{LAI}_\alpha > 0.25 \mathrm{LAI}_{\mathrm{min},\alpha}$:

$\quad \lambda_{2,\alpha} = \cosh(0.115(\mathrm{LAI}_\alpha - 0.25\mathrm{LAI}_{\mathrm{min},\alpha})) - 1$

if $\mathrm{LAI}_\alpha \le 0.25 \mathrm{LAI}_{\mathrm{min},\alpha}$:

$\quad \lambda_{2,\alpha} = 0$ $\hspace{3cm}$ (18)

The original formulation of Arora and Boer (2006a) only considered $\lambda_{1,\alpha}$ but here we adjust the parametrization with the addition of $\lambda_{2,\alpha}$, which ensures that a small fraction of NPP is used for spatial expansion even at very low LAI values. This additional constraint allows for improved fractional coverage of grasses in arid regions. Similar to $S_{\mathrm{sap},\alpha}$, $\mathrm{LAI}_{\mathrm{min},\alpha}$ and $\mathrm{LAI}_{\mathrm{max},\alpha}$ are PFT-dependent parameters (see Table 1).

The value of λ_{max} is set to 0.1 so that a maximum of 10 % of daily NPP can be used for spatial expansion. Finally, Λ_α is set to zero for tree PFTs when they are in a full leaf-out mode and all NPP is being used for leaf expansion (see Appendix A5).

2.1.3 Mortality

The PFT-dependent mortality rate (day^{-1}),

$$m_\alpha = m_{\mathrm{intr},\alpha} + m_{\mathrm{ge},\alpha} + m_{\mathrm{bioclim},\alpha} + m_{\mathrm{dist},\alpha}, \hspace{1cm} (19)$$

reflects the net effect of four different processes: (1) intrinsic- or age-related mortality, m_{intr}, (2) growth or stress-related mortality, m_{ge}, (3) mortality associated with bioclimatic criteria, m_{bioclim} and (4) mortality associated with disturbances, m_{dist}.

Intrinsic- or age-related mortality uses a PFT-specific maximum age, A_{max} (Table 1), to calculate an annual mortality rate such that only 1 % of tree PFTs exceed $A_{\mathrm{max},\alpha}$. Intrinsic mortality accounts for processes, whose effect is not explicitly captured in the model including insect damage, hail, wind throw, etc.,

$$m_{\mathrm{intr},\alpha} = 1 - \exp(-4.605/A_{\mathrm{max},\alpha}). \hspace{1cm} (20)$$

Grasses and crops have $m_{\mathrm{intr}} = 0$. The annual growth-related mortality m_{ge} is calculated using growth efficiency of a PFT over the course of the previous year following Prentice et al. (1993) and Sitch et al. (2003) as

$$m_{\mathrm{ge},\alpha} = \frac{m_{\mathrm{ge},\mathrm{max},\alpha}}{1 + k_m g_{\mathrm{e},\alpha}}, \hspace{1cm} (21)$$

where $m_{\mathrm{ge},\mathrm{max}}$ represents the PFT-specific maximum mortality rate when no growth occurs (Table 1). k_m is a parameter set to 0.3 m^2 (gC)$^{-1}$. g_{e} is the growth efficiency of the PFT (gCm^{-2}) calculated based on the maximum LAI ($L_{\alpha,\mathrm{max}}$; m^2 m^{-2}) and the increment in stem and root mass over the course of the previous year (ΔC_{S} and ΔC_{R}; kgCm^{-2}, respectively) (Waring, 1983)

$$g_{\mathrm{e},\alpha} = 1000 \frac{\max(0, (\Delta C_{\mathrm{S},\alpha} + \Delta C_{\mathrm{R},\alpha}))}{L_{\alpha,\mathrm{max}}}. \hspace{1cm} (22)$$

Mortality associated with bioclimatic criteria, m_{bioclim} (0.25 yr^{-1}), is applied when climatic conditions in a grid cell become unfavourable for a PFT to exist and ensures that PFTs do not exist outside their bioclimatic envelopes, as explained in the next section.

The annual mortality rates for m_{intr}, m_{ge} and m_{bioclim} are converted to daily rates and applied at the daily time step of the model, while m_{dist} is calculated by the fire module of the model based on daily area burned for each PFT as summarized in Appendix A9. In practice, the $\frac{\mathrm{d}f_\alpha}{\mathrm{d}t} = -m_{\mathrm{dist},\alpha} f_\alpha$ term of Eq. (5) is implemented right after area burnt is calculated.

2.1.4 Bioclimatic limits and existence

The mortality associated with bioclimatic criteria, m_{bioclim}, ensures that PFTs do not venture outside their bioclimatic envelopes. The bioclimatic criteria that determine PFT existence are listed in Table 2 for tree PFTs. Bioclimatic limits are not used for the C$_3$ and C$_4$ grass PFTs. The bioclimatic limits represent physiological limits to PFT survival that are either not captured in the model or processes that are not sufficiently described by empirical observations to allow their parametrization. Some examples of the latter include a plant's resistance to frost damage and xylem cavitation limits due to moisture stress. The bioclimatic criteria include the minimum coldest month air temperature ($T_{\mathrm{min}}^{\mathrm{cold}}$), the maximum coldest month air temperature ($T_{\mathrm{max}}^{\mathrm{cold}}$), the maximum warmest month air temperature ($T_{\mathrm{max}}^{\mathrm{warm}}$), the minimum number of annual growing degree days above 5 °C (GDD5$_{\mathrm{min}}$), the minimum annual aridity index (ratio of potential evapotranspiration to precipitation; arid$_{\mathrm{min}}$) and the minimum dry season length in a year (dryseason$_{\mathrm{min}}$), where the dry season length represents the number of consecutive months with precipitation less than potential evaporation. The bioclimatic indices are updated on a 25 year timescale ($T = 25$) such that the slowly changing value of a bioclimatic index $X(t+1)$ for time $t + 1$ is updated using its previous year's value $X(t)$ and its value $x(t)$ for the current year as

$$X(t+1) = X(t)e^{-1/T} + x(t)(1 - e^{-1/T}). \hspace{1cm} (23)$$

Equation (23) implies that 63 % of a sudden change in the value of a bioclimatic index Δx is reflected in $X(t)$ in T years $(1 - e^{T(-1/T)} = 1 - e^{-1} = 0.63)$, while 86 % of the change is reflected in $2T$ years $(1 - e^{2T(-1/T)} = 1 - e^{-2} = 0.86)$.

Table 2. Bioclimatic limits used in the competition parametrization of CTEM v. 2.0 as described in Sect. 2.1.4. T_{min}^{cold} and T_{max}^{cold} are the minimum and maximum coldest month temperatures, respectively. T_{max}^{warm} is the maximum warmest month temperature, GDD5$_{min}$ is the minimum growing degree days above 5 °C (GDD5), arid$_{min}$ is the minimum aridity index and dryseason$_{min}$ is the minimum length of the dry season.

Plant functional type short name	T_{min}^{cold} (°C)	T_{max}^{cold} (°C)	T_{max}^{warm} (°C)	GDD5$_{min}$ (°C)	arid$_{min}$ (unitless)	dryseason$_{min}$ (months)
NDL-EVG	–	≤ 18.0	–	≥ 375.0	–	–
NDL-DCD	–	≤ −28.0	≤ 25.0	≥ 600.0	–	–
BDL-EVG	≥ 2.5	–	–	≥ 1200.0	–	–
BDL-DCD-COLD	≥ −35.0	≤ 16.0	–	≥ 300.0	–	–
BDL-DCD-DRY	≥ 4.0	–	–	–	≥ 0.9	≥ 5.5
C3-CROP	–	–	–	–	–	–
C4-CROP	–	–	–	–	–	–
C3-GRASS	–	–	–	–	–	–
C4-GRASS	–	–	–	–	–	–

3 Simulations and observation-based data

3.1 Simulations

We perform three equilibrium simulations for pre-industrial conditions on a Gaussian 96 × 48 grid with a horizontal spatial resolution of approximately 3.75° × 3.75°. A globally uniform CO_2 concentration of 286.37 ppm (corresponding to the year 1861 from Meinshausen et al., 2011) is used. The fractional coverage of CTEM PFTs in these simulations is (i) prescribed (hereafter referred to as the PRES simulation), (ii) based on the CTEM competition parametrization as outlined in Sect. 2.1.1 (hereafter the CTCOMP simulation) and (iii) based on competition between the PFTs using the original L–V formulation (hereafter the LVCOMP simulation). The competition parametrization in the LVCOMP simulation uses a value of $b = 1$ in Eq. (5) similar to the L–V equations used in the TRIFFID (Cox, 2001) and USCM (Brentnall et al., 2005) models. By contrast, the CTCOMP simulation uses a value of $b = 0$.

The simulations are driven with the Climate Research Unit – National Centers for Environmental Prediction (CRU-NCEP) v. 4 climate data (Viovy, 2012). The climate corresponding to the 1901–1940 period is used repeatedly until the model carbon pools reach equilibrium. The climate data are disaggregated from their original 6-hourly temporal resolution to a half-hourly time step. Surface temperature, surface pressure, specific humidity and wind speed are linearly interpolated. Long-wave radiation is uniformly distributed across a 6 h period, and shortwave radiation is diurnally distributed over a day based on a grid cell's latitude and day of year with the maximum value occurring at solar noon. Precipitation is treated following Arora (1997), where the total 6 h precipitation amount is used to determine the number of wet half-hours in a 6 h period. The 6 h precipitation amount is then spread randomly, but conservatively, over the wet half-hourly periods. Soil texture information (percent sand, clay and or-

ganic matter) for the three soil layers is adapted from Zobler (1986). The driving climate data and soil and land cover information are interpolated to 3.75° resolution. All simulations are performed offline, i.e. feedbacks between the vegetation and climate are not possible. Offline simulations allow for more straightforward interpretation of model results and are a necessary first step before proceeding with coupled simulations that incorporate two-way interactions between the land surface and atmosphere.

The simulations with prescribed PFT fractions (PRES) use a reconstruction of land cover for the year 1861 based on the approach described in Wang et al. (2006) (hereafter referred to as W2006) and Arora and Boer (2010). Briefly, the snapshot of land cover in the Global Land Cover 2000 (GLC2000) data set for year 2000 is first mapped on to the CTEM PFTs and then extended back in time to account for changing crop area according to the HYDE v. 3.1 data set (Hurtt et al., 2011). We refer to this as the modified W2006 data set compared to the original W2006 land cover product, which was based on the Ramankutty and Foley (1999) crop data. The crop and grass fractions in each grid cell are divided into C_3 and C_4 sub-types based on the global fractional distribution of C_4 vegetation developed by Still et al. (2003), as described in W2006. Simulations with dynamically determined fractional coverage of CTEM PFTs (CTCOMP and LVCOMP) use only the C_3 and C_4 crop fractions from the modified W2006 data set while competition determines the fractional coverage of CTEM's non-crop PFTs over the remaining fraction. Disturbance, in the form of area burned, is included in all simulations based on CTEM's fire parametrization described in Appendix A9.

3.2 Observation-based data sets

A Moderate-Resolution Imaging Spectroradiometer (MODIS)-derived land cover product (Friedl et al., 2013) and the modified W2006 product are used to evaluate the

simulated fractional coverage of CTEM's seven non-crop PFTs.

The GLC2000 land cover product comprising of 22 land cover types is mapped on to the nine PFTs represented in CTEM by W2006 (their Table 2). W2006 split the broadleaf deciduous PFT into cold deciduous and drought deciduous versions. They made the simple but reasonably realistic assumption that at high latitudes (above 34°) deciduousness is caused primarily by seasonality in temperature while at lower latitudes (below 24°) deciduousness is caused by seasonality in precipitation, with a linear transition in-between. In addition, they separated grass and crops into their C_3 and C_4 sub-types as mentioned above.

The MODIS land cover product contains 17 land cover types and we map it to CTEM PFTs following the mapping scheme of W2006, as closely as possible (see Table S1 in the Supplement). The MODIS product is averaged across the years 2001–2011 and interpolated to the Gaussian 96×48 resolution. We do not attempt to subdivide the fractional coverage of MODIS broadleaf deciduous trees into their cold and drought deciduous versions, nor do we subdivide fractional coverage of MODIS grasses and crops into their C_3 and C_4 sub-types, since our classification scheme is not ground-truthed to the extent the approach used by W2006 has been. As such, the MODIS-based data set is used only for comparisons of global total coverage while the modified W2006 data set is also used for comparison of geographic distribution of CTEM's seven non-crop PFTs.

Both the MODIS-derived product (17 land cover types) and the modified W2006 land cover (based on GLC2000's 22 land cover types) product are subject to errors in categorizing remotely sensed vegetation into broad-scale vegetation types and then their further mapping onto CTEM's nine PFTs. The latter mapping requires assumptions about which of the land cover types contribute to which of the CTEM PFTs and in what proportion, including the bare fraction. For instance, W2006 assign the GLC2000 *tree cover, needle-leaved, deciduous* land cover type to CTEM's needleleaf deciduous tree, grass and bare fraction in proportions of 80, 10 and 10 %, respectively. The subjectiveness inherent in reconstructing observation-based data sets of PFT fractional coverage that can be directly compared to model output therefore implies that these data sets only provide information to constrain simulated results at large continental scales.

We also compare the latitudinal distribution of simulated gross primary productivity (GPP), vegetation biomass and soil carbon with observation-based estimates of these quantities. The observation-based estimate of GPP is from Beer et al. (2010), who analyse the ground-based carbon flux tower observations from about 250 stations using diagnostic models and extrapolate the flux tower observations to the global scale for the 1998–2005 period. We evaluate simulated vegetation biomass against two data sets. The global data set of Ruesch and Holly (2008) is based upon remotely sensed vegetation cover (Global Land Cover
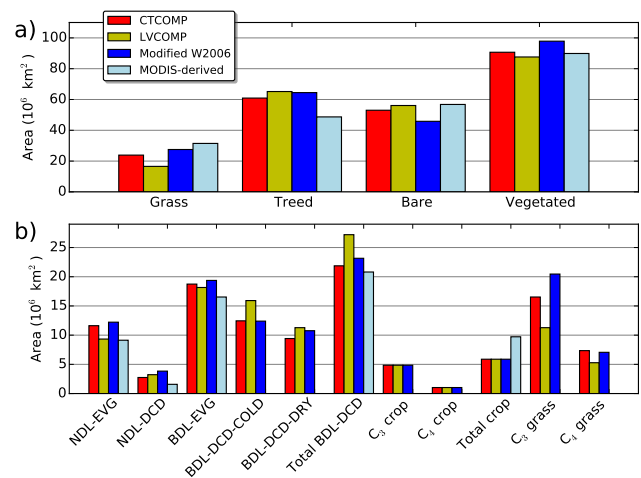

Figure 1. Comparison of observation-based and simulated global areal extent of grass, treed, bare ground and total vegetated area (sum of grass, treed and crop areas) from pre-industrial simulations that use the CTEM competition scheme (modified L–V; CTCOMP) and the unmodified L–V equations (LVCOMP) **(a). (b)** Shows the comparison for areal extent of individual PFTs. The observation-based estimates are from the modified Wang et al. (2006) data set (corresponding to year 1861) and the MODIS-derived product (which corresponds to the present day).

2000 Project, GLC2000) and carbon stocks at the national level. The Saatchi et al. (2011) vegetation biomass data set is available only for tropical and sub-tropical regions between 55° S and 40° N. These data are based on forest height measurements that are used to estimate aboveground biomass; belowground biomass is estimated on the basis of aboveground biomass using tree allometric relationships. The simulated zonally averaged soil carbon is evaluated against the Harmonized World Soils Dataset (HWSD) (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012).

4   Results

4.1   Global values

The performance of the original L–V-based competition parametrization (LVCOMP) and its modified version as implemented in CTEM (CTCOMP) is evaluated at the most basic level in Fig. 1a comparing the simulated coverage of trees and grasses, and the bare and vegetated areas, at the global scale with observation-based estimates of these quantities from the modified W2006 and MODIS-derived data sets.

Overall, the modelled global coverage of trees and grasses, and the global bare and vegetated areas in the CTCOMP and LVCOMP simulations compare reasonably with their observation-based estimates. The global tree cover in the CTCOMP simulation is about 6 % lower than the W2006 data set but 20 % more than the MODIS-derived estimates,
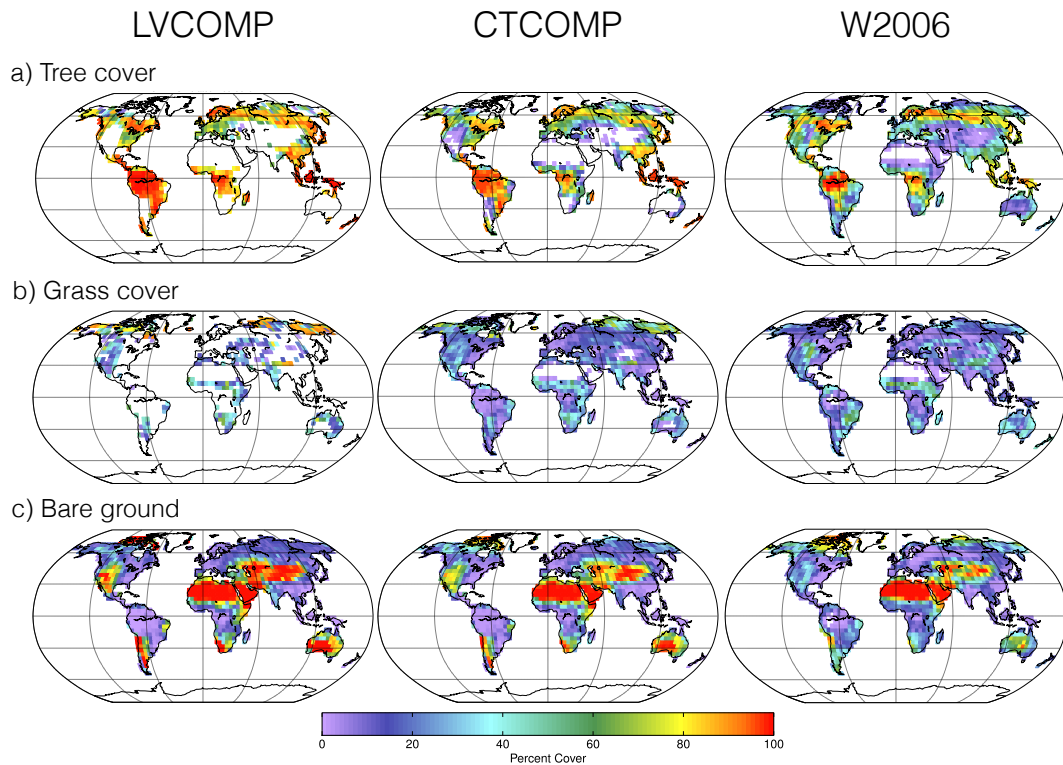
Figure 2. Global tree **(a)**, grass **(b)**, and bare ground **(c)** cover at equilibrium for pre-industrial conditions as simulated by CTEM using the unmodified L–V relations (LVCOMP; left column), the CTEM competition scheme with modified L–V relations (CTCOMP; middle column), and the observation-based fractions from the modified Wang et al. (2006) data set (W2006; right column).

which are for the modern period. Simulated coverage by CTCOMP lies in-between the observation-based estimates based on the modified W2006 and the MODIS-derived products, except for the area covered by grasses, which is lower than both observation-based estimates (13 and 24 %, respectively). LVCOMP simulated global area covered by trees is the highest while coverage of grasses and vegetated area is the lowest compared to the CTCOMP simulation and observation-based estimates. Total vegetation cover in the CTCOMP (90.7×10^6 km^2) and LVCOMP simulations (87.6×10^6 km^2) is slightly too small compared to the modified W2006 data set (97.8×10^6 km^2) but close to the MODIS-derived estimate (89.9×10^6 km^2) for the modern period. Consequently, the simulated bare area (CTCOMP, 53.0×10^6 km^2, and LVCOMP, 56.1×10^6 km^2) is overestimated compared to the modified W2006 estimate (45.8×10^6 km^2) and closer to the MODIS-derived estimate (56.8×10^6 km^2).

Figure 1b compares the simulated global areas at the individual PFT level. This provides a more stringent test of the competition parametrization in contrast to the aggregated comparisons shown in Fig. 1a. The observation-based estimates from the modified W2006 data set are available for individual PFTs. However, we do not attempt to sub-divide the MODIS-derived estimate for global broadleaf deciduous

(BDL-DCD) tree coverage into the dry-deciduous and cold-deciduous sub-types, and grass and crop PFTs into C$_3$ and C$_4$ sub-types, as mentioned earlier.

The global areas of individual PFTs in Fig. 1b from CTCOMP and LVCOMP simulations compare reasonably to the two observation-based estimates. The most dominant and least dominant non-crop PFTs are BDL-DCD trees and needleleaf deciduous (NDL-DCD) trees, respectively, for both simulations and observation-based estimates. Crop areas are specified based on the modified W2006 data set in the LVCOMP and CTCOMP simulations and therefore same for all three cases. The total crop area based on the MODIS data set for the modern period is higher than the value from the modified W2006 data set, which corresponds to 1860, consistent with the increase in crop area over the historical period. As a result MODIS-derived global area covered by all trees (in Fig. 1a) and individual PFTs (in Fig. 1b) is lower than the value based on the modified W2006 product.

As with Fig. 1a, the LVCOMP simulation does somewhat poorly compared to the CTCOMP simulation. The simulated global area for broadleaf deciduous trees is the highest, and C$_3$ grasses is the lowest, in the LVCOMP simulation compared to the CTCOMP simulation and observation-based estimates.
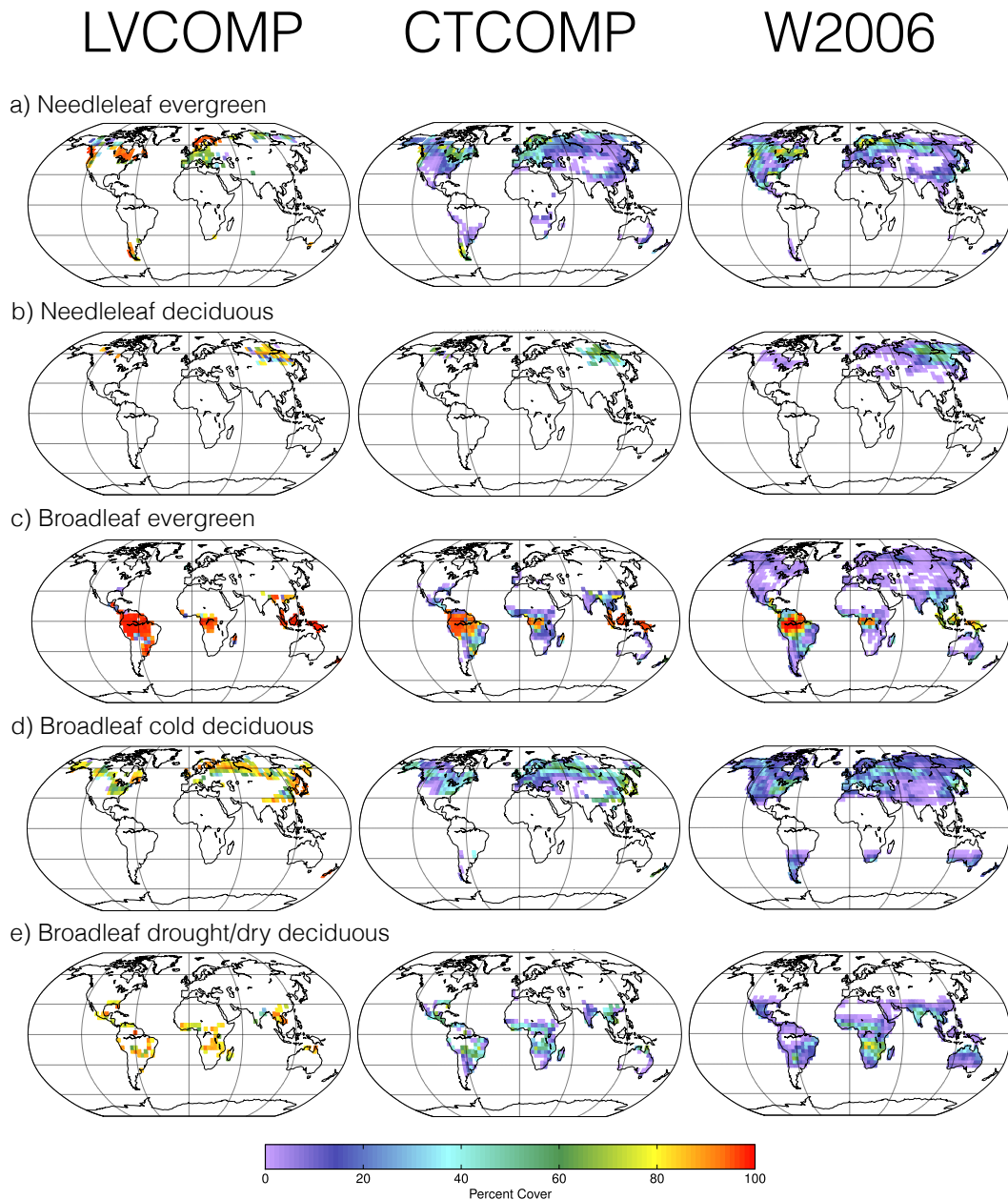
Figure 3. Comparison of simulated fractional coverage of tree PFTs in the LVCOMP (based on the original L–V equations, left column) and CTCOMP (based on modified L–V equations, middle column) simulations with the observation-based estimate from the modified Wang et al. (2006) data set corresponding to year 1861 (right column).

4.2 Geographical distributions

Figure 2 compares the geographical distributions of tree and grass PFTs and the bare fraction in the LVCOMP and CT-COMP simulations with the observation-based estimate from the modified W2006 product. Figures 3 and 4 compare the simulated and observation-based geographical distributions for individual tree and grass PFTs. The goodness of simulated geographical distributions of fractional coverage of PFTs is assessed in Fig. 5, which plots the root mean square

difference (RMSD) against the spatial correlation when results from the LVCOMP and CTCOMP simulations are compared to the modified W2006 product.

4.2.1 Tree cover

The broad spatial distribution of tree cover is simulated reasonably well in the CTCOMP simulation, including the boreal and tropical forests, although the model tends to simulate less tree cover in arid regions, especially the south-western

United States. The spatial correlation between the modelled tree fraction in the CTCOMP simulation and the modified W2006 data set is 0.79 with a RMSD of 22.8 %. Three main factors make simulating tree cover challenging in our modelling framework. First, the sub-grid scale climatic niches are unresolved at the $\sim 3.75°$ grid resolution used in this study. This is especially problematic for grid cells with large variations in climate. For example, this is the case in western Mexico where a strong elevation gradient creates climate niches for trees on the windward side of the Sierra Madre Occidental, while the leeward side of the mountains is very arid. Second, CLASS and CTEM do not presently simulate shrubs as a separate PFT. Shrubs are extensive in both hot (e.g. Australia) and cold (e.g. the Arctic) semi-arid to arid regions. In these regions the model either grows only grasses, or small trees. Finally, the limited number of natural PFTs represented in CTEM (seven) limit our ability to capture the entire spatial range of PFTs, e.g. needleleaf evergreen trees occur naturally both in the Canadian Yukon territory and the south-western United States. The trees found in these two regions have extensive physiological adaptations to their very different habitats. The single needleleaf evergreen PFT in the model, however, has only one set of parameters and as a result the model is unable to represent needleleaf evergreen trees realistically in both regions (as also seen in Fig. 3a).

The spatial distribution of trees in the LVCOMP simulation is similar to that in the CTCOMP simulation but generally shows less tree cover in arid regions and higher cover in predominately forested regions compared to both the W2006 data set and the CTCOMP simulation. The LVCOMP simulation demonstrates similar spatial correlation (0.76) but poorer RMSD (27.5 %) against the modified W2006 data set than the CTCOMP simulation (Fig. 5). The higher simulated tree cover in the LVCOMP simulation, in areas where trees do well, is a characteristic of the L–V formulation, which allows the dominant PFT to cover a large fraction of the grid cell allowing little coexistence (as discussed in detail by Arora and Boer, 2006a). The result of this inability to allow adequate PFT coexistence also leads to less gradation from areas with high tree cover to areas of low tree cover in the LVCOMP simulation compared to the CTCOMP simulation. This amplified expression of dominance, which suppresses PFT coexistence, is also seen in the CLM-DGVM, which uses the competition module of the LPJ dynamic vegetation model (Levis et al., 2004). In the CLM-DGVM simulation reported in Zeng et al. (2008) trees tend to occupy more than 90 % of a grid cell in locations where they are able to exist (their Fig. 7a). The Zeng et al. (2008) simulation is also unable to capture the continuous stretch of boreal forest from Canada's west to east coast.

4.2.2 Grass cover

The spatial coverage of grass in the CTCOMP simulation shows a fairly reasonable agreement compared to modified W2006 data set, especially on the African continent (Fig. 2), although the spatial correlation of 0.38 with the modified W2006 data set is lower than that for trees (Fig. 5).

Grass cover in CTEM, regardless of the competition parametrization (CTCOMP or LVCOMP), is influenced by three main factors: (i) tree coverage, as trees are considered superior to grasses because of their ability to invade them, (ii) moisture availability and (iii) the disturbance regime. Higher disturbance rates act to reduce tree cover and increase grass cover because grasses colonize faster than trees after a disturbance. Frequent disturbance thus reduces the superiority of trees over grasses.

Grass coverage is overestimated in the northern high latitude regions in the CTCOMP simulation (Fig. 2). This is likely due to CTEM not representing shrub PFTs (as mentioned above), allowing greater grass coverage in areas where shrubs should otherwise be extensive. Additionally, CTEM has no moss or lichen PFTs, which are also widespread in those regions and generally known to inhibit grasses. Another constraint is that CTEM represents only one C_3 grass PFT, which does not allow us to accommodate physiological adaptations over the broad climatic envelope in which grasses exist. As grasses cannot outcompete trees, low grass cover is seen in the CTCOMP simulation in areas where CTEM overestimates tree cover, for example the southern regions of Brazil and Uruguay. In areas with very high tree cover (like the western Amazon), the CTCOMP simulation tends to slightly underestimate tree cover (by < 10 %), which allows some grass cover (< 10 %) while the modified W2006 data set shows negligible grass cover.

The grassland extent in the parts of the US Plains and Canadian Prairies is also underestimated in the CTCOMP simulation. The modified W2006 data set shows around 60 % grass cover in this region while CTEM estimates about 30 % grass cover. This underestimate could be due to the following three reasons. First, fire is simulated interactively in CTEM and biases in simulated area burned will influence the PFT dynamics as fire tends to reduce tree cover and increase grass cover. The fire parametrization includes both human (as a function of population density) and natural influences but cultural differences in the human influence are not represented (see Appendix A9). While CTEM simulates more fire in the plains–prairie region (annual fraction burned is around 1–5 %) than indicated by the remotely sensed GFED v. 3 data set (Giglio et al., 2010) (which covers 1996–2009), it is difficult to estimate a realistic amount of annual burned area during the pre-industrial period. Second, the plains region was home to large herds of grazing herbivores prior to the 1880s, which also tend to damage tree seedlings providing beneficial conditions for grass expansion. Herbivory is not represented in CTEM. Finally, the simulated soil moisture in these regions could be overly moist allowing trees to cover a larger fraction than if the soil moisture was more limiting.

Grass cover in the LVCOMP simulation is relegated to areas where trees are not able to effectively colonize, due to
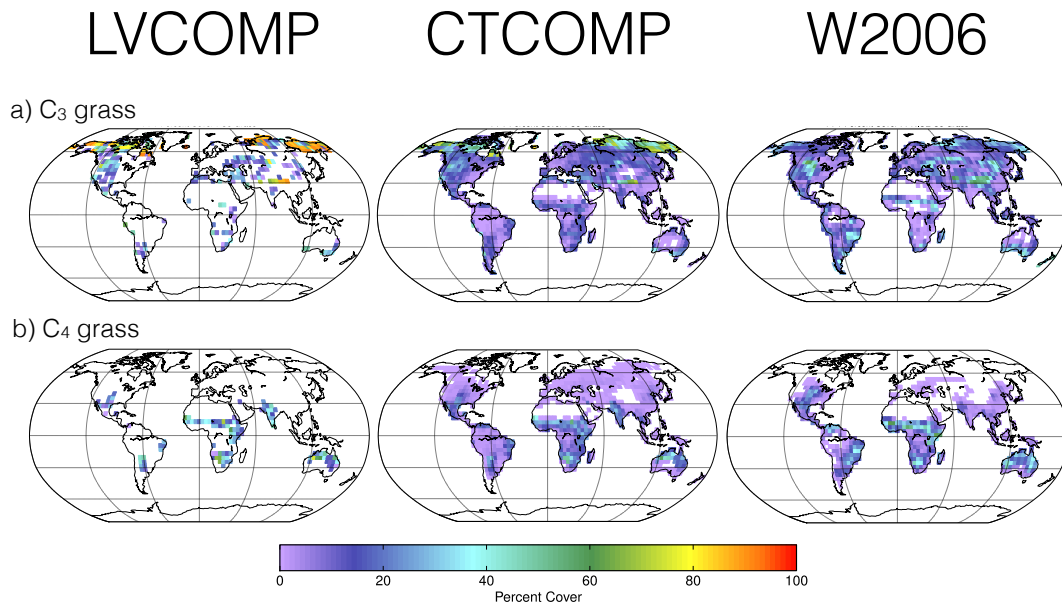
Figure 4. Comparison of simulated fractional coverage of grass PFTs in the LVCOMP (based on the original L–V equations, left column) and CTCOMP (based on modified L–V equations, middle column) simulations with the observation-based estimate from the modified Wang et al. (2006) data set corresponding to year 1861 (right column).

conditions being overly arid, overly cold and/or with high disturbance regimes. In the LVCOMP simulation, even more extensive grass cover is simulated in the high latitudes than in the CTCOMP simulation, approaching 90 % in many regions. The LVCOMP simulation's low grass cover in many other regions reflects the high tree cover estimated in the LVCOMP simulation due to its inability to allow appropriate PFT coexistence (as mentioned earlier) acting to exclude grasses. As a result the grass cover in the LVCOMP simulation is low and patchy and differs greatly from the modified W2006 data set, as is also indicated by a low spatial correlation and higher RMSD in Fig. 5.

4.2.3   Bare ground

Bare ground in the modified W2006 data set occurs primarily in arid, mountainous and arctic regions (Fig. 2c), as is normally expected. Simulated bare fraction in the CTCOMP simulation compares reasonably well with the modified W2006 estimates (spatial correlation of 0.85), especially in Africa and Eurasia although the model simulates a higher bare fraction in the south-western United States, Mexico and Australia. The model generally also simulates a greater bare fraction in other areas like the Andes mountains of South America and the Kalahari desert of Africa. The overestimation of bare fraction in the CTCOMP simulation in these regions reflects the processes mentioned earlier, which affect the simulated distribution of trees and grass. As well, the coarse spatial resolution ($\sim 3.75°$) of our study prevent us from capturing sub-grid scale niches. For example, in mountainous regions, large elevation gradients cre-
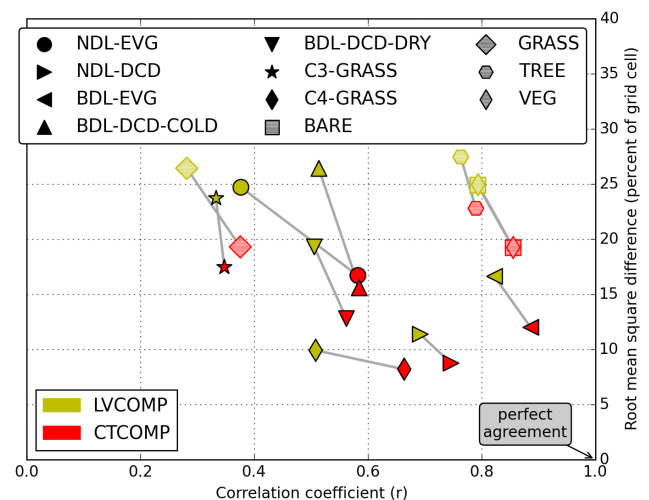

Figure 5. Scatter plot of the model simulated vegetation cover root mean square difference and correlation coefficient as compared to the observation-based estimate derived from the modified Wang et al. (2006) data set. Perfect agreement between the model and the observation-based values would yield a RMSD of zero and a correlation coefficient of 1. All simulations by CTCOMP show closer agreements with the W2006 data set than the corresponding LVCOMP simulations (each simulation pair between LVCOMP and CTCOMP for each PFT/aggregate is linked with a grey line).

ate climatic niches that allow vegetation to exist whereas using mean climate averaged over a large grid cell may prevent vegetation existence in the model. Other regions, such as the arid and semi-arid interior regions of Australia and

Mexico have sub-grid scale features that allow for the formation of landscapes with banded (Tongway and Ludwig, 2001) or patchy (Tongway and Ludwig, 1994) vegetation. These sub-grid scale processes are not resolved in our modelling framework, which assumes uniform soil and climate conditions within each grid cell. The LVCOMP simulation shows a similar pattern to that in the CTCOMP simulation, but with slightly more bare area especially in the south-western United States and Mexico and poorer RMSD and spatial correlation (Fig. 5).

4.3   Individual PFTs

The comparison of geographical distributions of simulated tree, grass cover and bare ground cover with observation-based estimates shows that, while some limitations remain, the competition parametrization of CTEM performs reasonably realistically. A more stringent test of the model performance is the comparison of geographical distributions of individual tree and grass PFTs against observation-based estimates, as shown in Fig. 3 for tree PFTs and Fig. 4 for grass PFTs.

4.3.1   Needleleaf evergreen trees

The modelled global total areal extent of the needleleaf evergreen (NDL-EVG) tree PFT in the CTCOMP simulation is around 5 % less than the modified W2006 estimate and about 21 % more than the modern-day MODIS-derived estimate (Fig. 1b). Spatially, the simulated fractional cover of NDL-EVG trees compares reasonably to W2006 in the boreal regions (Fig. 3 row a). The modelled fractional coverage of NDL-EVG trees is, however, overly extensive in central and western Europe as well as the southern tip of South America with too little NDL-EVG in Mexico, parts of the boreal forest of Canada, and in a band across Eurasia. A model limitation is that it is difficult to have NDL-EVG trees exist as far south as Mexico, as observations suggest, but not encroach into southern Africa (we incorrectly simulate a small fraction of NDL-EVG for a few grid cells in that region). Much of this over- and underestimation can be attributed to the use of a single PFT for NDL-EVG trees that are in reality physiologically distinct depending on their geographical location (as discussed in Sect. 4.1). The issue of physiological distinction between needleleaf evergreen trees growing in different regions is also illustrated by Reich et al. (2014), who show that needle longevity reduces as mean annual temperature increases and that this is an adaptive, rather than a genetic, response. In contrast, the NDL-EVG trees in the model have a constant needle lifespan regardless of their geographic location. The limitations of the use of a single NDL-EVG tree PFT in CTEM has also been demonstrated by Peng et al. (2014), who implemented CTEM at a regional scale for the province of British Columbia, Canada. Peng et al. (2014) found that while the use of a single NDL-EVG tree PFT

yields reasonable LAI and GPP in coastal British Columbia, which experiences mild temperate climate, the same PFT parameters result in lower than observed LAI and GPP in the continental interior of the province, which experiences a colder and drier climate.

The simulated distribution of NDL-EVG trees in the LVCOMP simulation differs greatly from the modified W2006 estimate with much lower coverage of NDL-EVG trees except in certain regions where the simulated coverage of NDL-EVG trees is high (> 60 %). The global simulated coverage of NDL-EVG trees in the LVCOMP simulation is about 25 % less than the modified W2006 estimate although it compares well to the modern-day MODIS-derived estimate. Comparing the RMSD and spatial correlation between the modified W2006 data set and the model results shows a reduced RMSD and a higher correlation, both of which imply an improved performance, in the CTCOMP compared to the LVCOMP simulation (Fig. 5).

4.3.2   Needleleaf deciduous trees

The observation-based distribution of needleleaf deciduous (NDL-DCD) trees is primarily limited to areas in eastern Siberia with some additional areas of low coverage in North America (< 10 %) (Fig. 3 row b). This pattern predominantly reflects the distribution of larches (*Larix* spp.), a common NDL-DCD tree species found in much of Canada, Scandinavia, and Russia, with their primary habitat being the boreal forests of Russia. To prevent NDL-DCD trees from occupying expansive regions in both Eurasia and North America, bioclimatic limits are used to limit their range. As a result, NDL-DCD in both CTCOMP and LVCOMP simulations exist in a closely defined area with little gradation from the regions where NDL-DCD exist to regions where they do not, reflecting the effect of the common bioclimatic limits. Within the climatically favourable regions for NDL-DCD, the simulated fractional cover in the CTCOMP simulation is similar to the modified W2006 data set while the simulated fractional coverage in the LVCOMP simulation is too high. The total area covered by NDL-DCD trees in the CTCOMP simulation is ∼ 28 % lower than in the modified W2006 data set but ∼ 43 % higher than the modern-day MODIS-derived estimate (Fig. 1b). The total area covered by NDL-DCD trees in the LVCOMP simulation is ∼ 16 % lower than the modified W2006 estimate but over twice the MODIS-derived estimate. There is a slight reduction in RMSD and small increase in spatial correlation, when compared to the modified W2006 values, in the CTCOMP compared the LVCOMP simulation (Fig. 5).

4.3.3   Broadleaf evergreen trees

While CTEM parametrizes the broadleaved evergreen (BDL-EVG) PFT as tropical trees (through bioclimatic limits; see Sect. 2.1.4), the W2006 data set also includes evergreen

shrubs when mapping the GLC2000 land cover to the nine CTEM PFTs (see Table 2 of Wang et al., 2006). As a result, the modified W2006 data set shows small areas of BDL-EVG outside the tropical regions. Excluding regions outside of the tropics, the modelled distribution of BDL-EVG trees in the CTCOMP simulation compares reasonably well to the modified W2006 data set with expansive coverage along the equator. The model generally simulates higher than observed coverage of BDL-EVG around Indonesia and the western margin of the Amazon and parts of South-east Asia. In the LVCOMP simulation, BDL-EVG trees have a smaller range and almost full coverage in regions where they are simulated to exist with little gradation from areas with high to areas with low coverage, which is characteristic of the unmodified L–V formulation. The global coverage of BDL-EVG trees in the modified W2006 data set is 19.4 million km^2 with 16.7 million km^2 in the tropics (between 30° S and 30° N). The simulated global coverage of the BDL-EVG trees in the CTCOMP and LVCOMP simulations is somewhat larger at 18.7 and 18.2 million km^2, respectively. The improved spatial distribution of BDL-EVG trees in the CTCOMP compared to the LVCOMP simulation is also seen in Fig. 5.

The western margin of the Amazon River basin shows the limitations of the $\sim 3.75°$ spatial resolution used in our study. The observations show a gradation from high coverage in the interior of the basin to lower values near the Pacific coast, reflecting the rise of the Andes mountains and the change in climate across them (Fig. 3 right column of row c). In our modelling framework, however, climate averaged over the large $\sim 3.75°$ grid cells yields amenable conditions for BDL-EVG trees to exist across the mountain range. The model also does not take into account the fraction of a grid cell that may be covered by rocks and stony outcrops, which would make it impossible for the natural PFTs to expand into those areas. Consideration of the *rocky* fraction of a grid cell will likely improve model simulated tree cover in mountainous regions.

4.3.4 Cold and drought deciduous broadleaf trees

In CTEM, broadleaf deciduous (BDL-DCD) trees are divided into cold (BDL-DCD-COLD) and drought/dry (BDL-DCD-DRY) sub-types. The bioclimatic indices for BDL-DCD-COLD and BDL-DCD-DRY trees are assigned so that their geographical distributions are broadly limited to regions where their deciduousness is primarily controlled by temperature and soil moisture, respectively. For the observation-based data set, this distinction is more arbitrary. W2006 used latitudinal limits of > 34 and $< 24°$ for characterizing deciduousness due to temperature and soil moisture, respectively, with a linear gradation in-between, as mentioned earlier (also visible in Fig. 3). As a result, in the modified W2006 data set, the cold and drought deciduous sub-types of the BDL-DCD PFT can coexist between the latitudes of > 24 and $< 34°$. A bioclimatic index that has been evaluated at a few sites to predict the trigger of deciduousness could present a better

predictor and will be investigated in future work. The bioclimatic indices and the methodology used in CTEM preclude the two deciduous versions from coexisting. While it is fairly obvious that temperature-based bioclimatic indices will be needed for determining the geographical distribution of BDL-DCD-COLD trees, we found that temperature-based indices alone are not sufficient for determining the geographical distribution of BDL-DCD-DRY trees, as would also be intuitively expected. In addition to minimum coldest month temperature, we therefore also use the annual aridity index and the length of the dry season for determining the geographical distribution of BDL-DCD-DRY trees.

The simulated geographical distribution of BDL-DCD-COLD in the CTCOMP simulation is broadly similar to the modified W2006 data set, although some differences remain. Compared to the modified W2006 data set, the model simulates larger BDL-DCD-COLD extent in western Canada (and lower NDL-EVG as a result) and smaller BDL-DCD-COLD coverage in south-eastern Europe (with comparably more NDL-EVG). These areas demonstrate that the competition between the NDL-EVG and BDL-DCD-COLD PFTs is not perfectly modelled in these regions. The arbitrary latitudinal limits used by W2006 imply that BDL-DCD-COLD trees exist as far north as Australia's Gold Coast in the eastern state of Queensland, which is certainly not realistic given its tropical climate. The model does not simulate any broad-scale existence of BDL-DCD-COLD trees between 24 and 34° S, although it reasonably estimates their coverage in New Zealand. The simulated expanse of BDL-DCD-COLD trees is much smaller in the LVCOMP simulation and most regions, where it is present, generally show high coverage (60–90 %). The global total areal coverage of BDL-DCD-COLD in the CTCOMP simulation is very similar to the modified W2006 data set while it is $\sim 22 \%$ higher in the LVCOMP simulation, despite its restricted spatial extent.

The simulated geographical distribution of BDL-DCD-DRY trees in the CTCOMP simulation is also broadly similar to the modified W2006 data set with some differences over India, and in southern Africa near Botswana and Zambia. CTEM coverage in these regions is lower by approximately 20 % than the modified W2006 estimate. As with other tree PFTs, the simulated areal extent of BDL-DCD-DRY trees in the LVCOMP simulation is more limited compared to the CTCOMP simulation and the modified W2006 data set, but of a higher percent cover in the grid cells where BDL-DCD-DRY trees do exist. The global total areal coverage of BDL-DCD-DRY trees in the CTCOMP and LVCOMP simulations are $\sim 12 \%$ lower and $\sim 5 \%$ higher than the modified W2006 estimate, respectively. In Fig. 5, the RMSD between the simulated and modified W2006 values shows a large decrease in the CTCOMP compared to the LVCOMP simulation.

The total BDL-DCD coverage, consisting of both the cold and dry deciduous sub-types, in the CTCOMP simulation is similar to the estimates based on MODIS and modified

W2006 products but much higher in the LVCOMP simulation (Fig. 1b).

4.3.5 C_3 and C_4 grasses

Both C_3 and C_4 sub-types of grasses are represented in CTEM. The geographical distributions of C_3 and C_4 grasses are determined solely on competitive advantages provided by their different photosynthetic pathways and parameter values that determine their differing responses to environmental conditions; bioclimatic indices are not used to limit the range of C_3 and C_4 grasses (see Table 2). The fractional coverage of grasses is, of course, influenced to a large extent by the fractional coverage of trees, which are considered hierarchically superior to grasses. The C_3 and C_4 grasses therefore compete for the remaining fraction of a grid cell that the trees are unable to occupy.

Figure 4 compares the simulated geographical distribution of C_3 and C_4 grasslands with the observation-based modified W2006 data set, but with the caveat that the W2006 data set itself is based on the part-modelled and part-observation-based product of Still et al. (2003), which estimates the C_4 fraction of total vegetation. The model is successfully able to simulate the coexistence of C_3 and C_4 grasses, especially in the CTCOMP simulation. The simulated global areal coverage of C_3 grass in the CTCOMP simulation is $\sim 19\%$ less than the modified W2006 data set and this is especially noticeable in areas like the US Plains and the Brazilian highlands (Fig. 4). Low grassland estimation reflects either an overly high tree cover (as is the case in the Brazilian highlands, see Fig. 2a) or too extensive areas of bare ground (as is the case in the US Plains; see Fig. 2c).

Reasons for overly extensive bare ground include either the inadequacies of using a single grass PFT globally for each photosynthetic pathway (C_3 and C_4) or simulated soil moisture that is too dry in those regions to allow grasses to effectively colonize. Appropriate grass cover is, however, simulated for other regions including southern Australia and the Tibetan plateau in the CTCOMP simulation. The model overestimates grass cover in the high latitudes regions. This overestimation possibly reflects: (i) the lack of sensitivity to high latitude climate of the single C_3 grass PFT used in the model, or (ii) the lack of shrubs, mosses, and lichens in CTEM, which allows the expansion of grasses into areas that are otherwise shrublands or moss and lichen covered, as mentioned earlier. An additional possibility is that CTEM does not presently take in sub-grid scale soil depth, which may be important in parts of the Arctic and would limit plant cover. The simulated grasslands in the LVCOMP simulation are more isolated with even higher coverage in parts of the Arctic region. In addition, the original L–V formulation does not capture the African savannas and the grasslands of South America.

The simulated distribution of C_4 grass in the CTCOMP simulation compares favourably to the modified W2006 data set with a similar geographical distribution and fractional cover. Small differences include an underestimation in the Brazilian highlands (due to an overestimation of the simulated tree cover) and an overestimation over India (due here to an underestimation of simulated tree cover). The total global area covered by C_4 grass also compares well to the modified W2006 estimate (Fig. 1), while the simulated global coverage in the LVCOMP simulation is low and patchy.

4.4 Primary terrestrial carbon pools and fluxes

The differences between the simulations with prescribed fractional coverage of PFTs (PRES) and the simulations in which fractional coverage of PFTs are dynamically simulated (CTCOMP and LVCOMP) are also the result of somewhat different parameter values between the two. The simulations with dynamically determined fractional coverage of PFTs provide an additional constraint, which the model must meet, i.e. the observation-based estimates of the fractional coverage of PFTs. We found that the default parameter values used for simulations with prescribed fractional coverage of PFTs did not yield optimum comparison with observation-based estimates of fractional coverage of PFTs. Changes in parameter values were required because the simulations with dynamically determined fractional coverage of PFTs include additional processes such as mortality generating bare ground (see Sect. 2.1.3) and spatial expansion consuming a fraction of NPP that then reduces vegetation productivity. Bare ground fraction does not need to be generated when fractional coverage of PFTs is prescribed. In the PRES simulation mortality of grasses is implicitly represented through the normal turnover of the leaves and root components (Table 1). The additional constraint provided by observation-based fractional coverage of PFTs provides the opportunity to improve model parameters. Parameter changes when spatial competition between PFTs is explicitly modelled include a slight reduction in base respiration rates for NDL-EVG and BDL-EVG with an increase for BDL-DCD-DRY (Table A2). These values were adjusted to optimize the CTCOMP results against observation-based data sets. The maximum cold and drought stress loss rates, BDL-EVG leaf lifespan, as well as the temperature threshold for determining the cold stress scalar, were optimized in a similar manner for several PFTs (Tables A5 and A6). The PFT-specific base allocation fractions for leaf, stem, and roots (Table A4) were revised to be in line with values derived from Luyssaert et al. (2007) and Litton et al. (2007) and then further optimized against the observation-based data sets. These changes in parameter values are listed in Appendix tables alongside the values used in PRES simulations (specifically Tables 1, A2, A4, A5 and A6). While, in the CTCOMP simulation, these changes in parameter values allow simulation of reasonable geographical distributions of modelled PFTs and generally improve
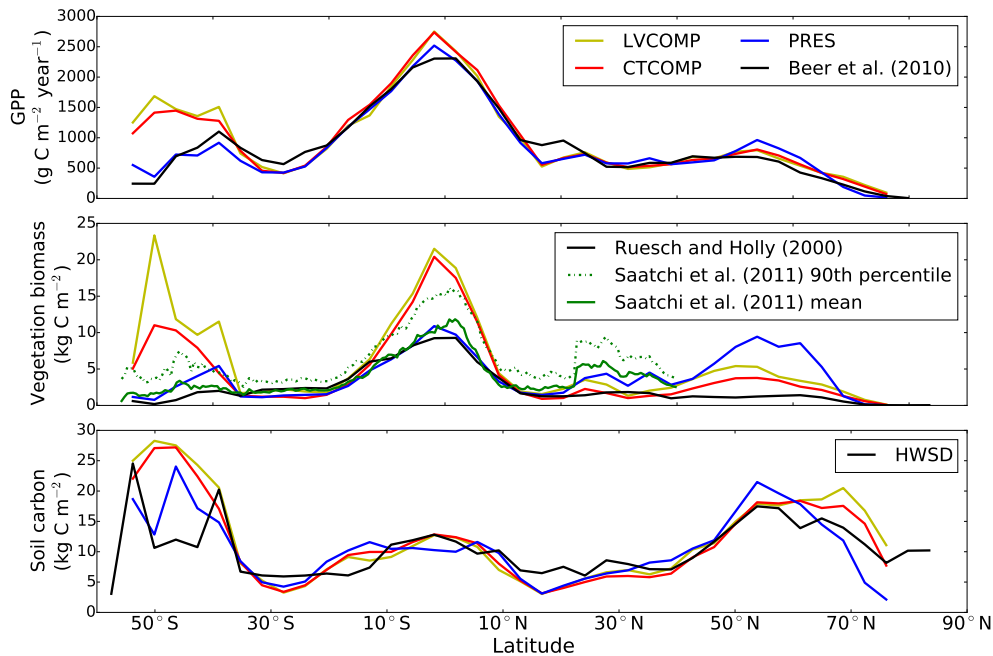
Figure 6. Zonal mean GPP (top), vegetation biomass (middle) and soil carbon (bottom) for CTEM simulations using (i) prescribed PFT fractional cover from the modified Wang et al. (2006) data set (PRES), (ii) the CTEM competition scheme, which uses modified L–V equations (CTCOMP) and (iii) simulations that use unmodified L–V equations (LVCOMP) compared to observation-based estimates. The observation-based estimate for soil carbon is based on the Harmonized World Soils Database (HWSD) version 1.2 (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012).

aspects of the model, they also adversely affect some aspects as discussed later in Sect. 4.4.2.

4.4.1 Global and zonally averaged values

Table 3 compares the global values of primary terrestrial carbon pools and fluxes from the PRES, CTCOMP and LVCOMP simulations with each other, and other model- and observation-based estimates. Figure 6 shows the simulated zonally averaged values for vegetation and soil carbon, and GPP from the three simulations together with available observation-based estimates. Generally, the simulated global values for primary terrestrial carbon pools and fluxes in the CTCOMP and LVCOMP simulations are similar to those in the PRES simulation, and values from all simulations compare reasonably to previously reported model and observation-based estimates (Table 3 and Fig. 6).

GPP is slightly higher in the CTCOMP and LVCOMP simulations compared to the PRES simulation. The primary reason for this increase is the changes in PFT specific parameters, rather than the spatial distribution of PFTs itself. Performing the PRES simulation, with the same parameter set as in the CTCOMP and LVCOMP simulations, yields a global GPP value of 140.8 Pg C yr^{-1}. The zonally averaged GPP in the CTCOMP and LVCOMP simulations compares well to that simulated in the PRES simulation and the observation-

based estimate of Beer et al. (2010) (Fig. 6) with an overestimate below 40° S, where there is relatively little land.

Similar to GPP, NPP in the CTCOMP and LVCOMP simulations is 9.1 and 6.2 % higher, respectively, compared to the PRES simulation (Table 3). Evapotranspiration is also slightly higher in the CTCOMP and LVCOMP simulations (3.2 and 4.8 %, respectively) than in the PRES simulation, although values from all three simulations compare well with model and observation-based estimates summarized in Trenberth et al. (2011).

The simulated vegetation biomass and soil carbon mass are the lowest in the CTCOMP simulation amongst the three simulations. The LVCOMP simulation conversely yields the highest vegetation biomass and soil carbon mass. The PRES simulation has intermediate values of vegetation biomass and soil carbon. Simulated litter mass is similar across all the three simulations. Global values of vegetation, soil carbon and litter mass from all simulations compare reasonably with other estimates.

Compared to the PRES simulation, the zonally averaged vegetation biomass in the CTCOMP and LVCOMP simulations is higher in the high southern latitudes (> 40° S) and the equatorial region (10° S–10° N) but less in the high northern latitudes (> 40° N) (Fig. 6, middle row). Compared to the observation-based estimates of Ruesch and Holly (2008) for the whole globe and Saatchi et al. (2011) for the tropics, the simulated vegetation biomass in the CTCOMP and

Table 3. Comparison of primary model quantities from the three CLASS–CTEM simulations with observation-based estimates and other model estimates. The CTCOMP simulation uses a modified form of the Lotka–Volterra equations for simulating competition between PFTs. The LVCOMP simulation uses the unmodified Lotka–Volterra equations. The PRES simulation uses prescribed fractional coverage of PFTs based on the modified Wang et al. (2006) data set.

Variable	CTCOMP	LVCOMP	PRES	Preindustrial values from modelling studies	Other modern estimates
GPP ($PgC\,yr^{-1}$)	128.4	125.6	124.1	134.0 (Gerber et al., 2004)	ca. 125 (Zhao et al., 2006)[a], 123 ± 8[b] (Beer et al., 2010)
NPP ($PgC\,yr^{-1}$)	68.1	66.0	61.9	64.0 (Sitch et al., 2003), 50–70 (Friedlingstein et al., 2006)	59.9 (Ajtay et al., 1979), 62.6 (Saugier et al., 2001), 56.6 (Running et al., 2004)
Evapotranspiration ($10^3\,km^2\,yr^{-1}$)	86.7	88.1	83.9		83.9 ± 9.9[c] (Trenberth et al., 2011)
Vegetation biomass (PgC)	559.1	693.7	645.2	932 (Sitch et al., 2003)	446[d] (Ruesch and Holly, 2008)
Soil carbon mass (PgC)	1382.3	1430.5	1419.5	1670 (Sitch et al., 2003)	1400–1600 (Schlesinger, 1977), 1395 (Post et al., 1982), 1348 (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012)[d]
Litter mass (PgC)	118.2	116.4	115.1	171 (Sitch et al., 2003)	90 (Ajtay et al., 1979)
Fire emissions ($PgC\,yr^{-1}$)	2.7	2.5	2.5	2.2[e] (Mieville et al., 2010), 1.5–2.7 (Mouillot et al., 2006)	1.92 (Mu et al., 2011)
Annual area burnt ($10^4\,km^2\,yr^{-1}$)	425.3	604.7	560.7	510[f] (Mouillot and Field, 2005)	342 (Mu et al., 2011), 294–374 (Giglio et al., 2006),

[a] MODIS-derived LAI driven with NCEP reanalysis. [b] Estimate for modern period including the effect of anthropogenic land use and elevated CO_2 concentrations. [c] Based on eight reanalyses for the period 2002–2008, except ERA-40, which was for the 1990s (Trenberth et al., 2011). [d] Interpolated to the same resolution and land mask as CTEM. [e] For the year 1900 using the burnt area estimate of Mouillot and Field (2005). [f] For the year 1900.

LVCOMP simulations is better simulated at high latitudes than in the PRES simulation but their simulated vegetation biomass in the equatorial region is too high. Site-level observations from the Large-scale Biosphere–Atmosphere (LBA) experiment in Amazonia, however, indicate total vegetation biomass values closer to those simulated in the CTCOMP and LVCOMP simulations (e.g. Keller et al., 2001). The far southern latitudes, including the southern tip of South America and New Zealand, have higher tree cover in the CTCOMP and LVCOMP simulations, which leads to an overestimated vegetation biomass in these regions compared to the PRES simulation and observation-based estimate of Ruesch and Holly (2008). Similarly the high northern latitude tree cover is somewhat less than that in the PRES simulation and this brings the amount of estimated vegetation biomass lower in the CTCOMP and LVCOMP simulations. We believe, however, that the estimates of vegetation biomass from Ruesch and Holly (2008) are too low at high latitudes as discussed by Peng et al. (2014) for British Columbia, Canada.

The zonal distribution of soil carbon (Fig. 6 bottom row) is broadly similar to the observation-based estimate from the HWSD (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012) for all three simulations.

Annual fire emissions are the highest in the CTCOMP simulation even though it has the lowest annual burned area (24.1 % less than in the PRES simulation) while the LVCOMP has the largest burned area (7.8 % more than in the PRES simulation) but the same amount of emissions as the PRES simulation (Table 3). The lowest annual burned area in the CTCOMP simulation, however, compares best with contemporary observation-based estimates (Table 3). Fire in

CTEM influences the vegetation biomass differently in the simulations with prescribed (PRES) and dynamically modelled (CTCOMP and LVCOMP) fractional coverage of PFTs. In the PRES simulation, biomass burning associated with fire reduces vegetation biomass density since the fractional coverage is not allowed to change, essentially causing thinning of the vegetation biomass. In the CTCOMP and LVCOMP simulations, fire additionally creates bare ground, which reduces PFT fractional cover and is subsequently available for colonization.

4.4.2 Geographical distributions

The geographical distribution of GPP, vegetation biomass, soil carbon mass and burned area from the PRES, CTCOMP and LVCOMP simulations are compared in Figs. 7 and 8 with available observation-based estimates.

The geographical distribution of GPP (Fig. 7a) shows somewhat higher productivity in the tropics for the CTCOMP and LVCOMP simulations compared to PRES simulation, as also seen in the zonally averaged values. Simulated GPP is also marginally higher in the CTCOMP and LVCOMP simulations in the Arctic region than in the PRES simulation due to the overestimated grass cover. Modelled GPP in all three simulations (LVCOMP, CTCOMP, PRES) is generally lower than observation-based estimates in arid regions including the Atacama desert, the Australian outback and the deserts of the central Asian region.

The geographical distribution of vegetation biomass between the three simulations (Fig. 7b) differs the most in the mixed forests of the north-eastern United States and east-
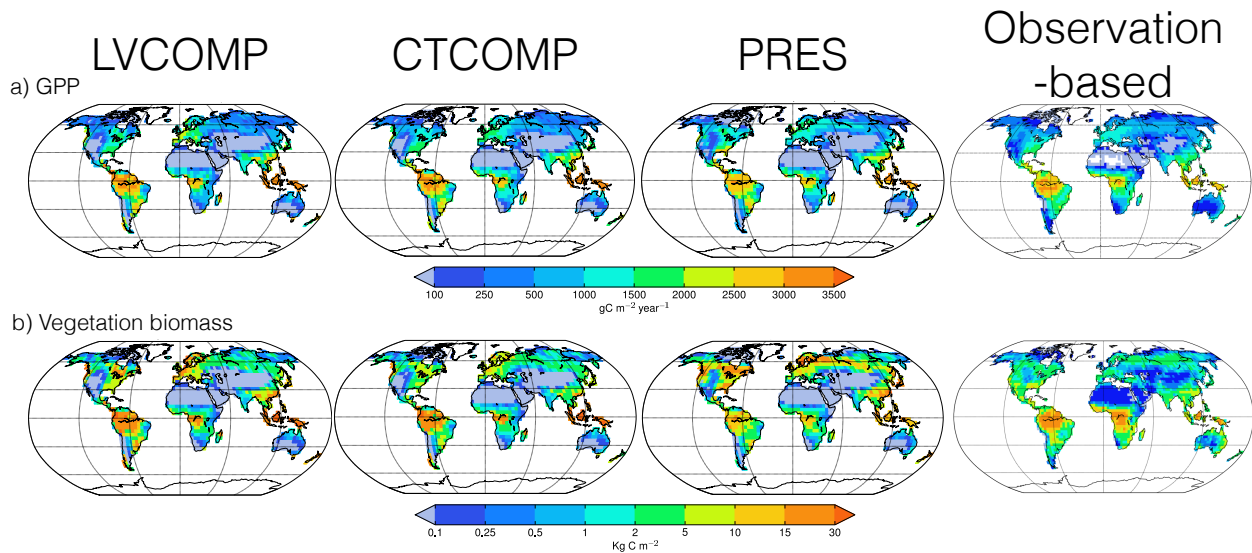
Figure 7. Geographical distribution of modelled GPP **(a)**, and vegetation biomass **(b)** from simulations that use dynamically modelled fractional coverage of PFTs based on the original L–V equations (LVCOMP, leftmost column), and the modified L–V equations (CTCOMP, second column). The third column shows simulation results that use prescribed PFT fractional cover based on the modified Wang et al. (2006) data set (PRES). The fourth column has observation-based data sets of GPP (Beer et al., 2010) and vegetation biomass (Ruesch and Holly, 2008).

ern Canada with a lower biomass in the CTCOMP and LV-COMP simulations compared to the PRES simulation. Parts of Scandinavia and north-western Russia show higher vegetation biomass in the PRES simulation compared to the CT-COMP and LVCOMP simulations. The revised model parameters (see tables in the Appendix) yield lower and more realistic vegetation biomass in some high-latitude regions in the CTCOMP simulation, while in the PRES simulation the vegetation biomass in these high-latitude regions is unrealistically as high as in the tropics. However, the model performance in the CTCOMP and LVCOMP simulations is adversely affected in the tropics where the simulated vegetation biomass is high compared to the PRES simulation and observation-based estimates (Saatchi et al., 2011). As with simulated GPP, modelled vegetation biomass is lower than observation-based estimates in arid regions.

As with the zonally averaged values, the geographical distribution of soil carbon (Fig. 8a) is broadly similar in the PRES, CTCOMP and LVCOMP simulations. Differences exist in areas where the grassland extent differs between the simulations (e.g. the US Plains/Canadian Prairies and the high Arctic) because in CTEM grasses are characterized by low soil decomposition rates, which result in high belowground carbon amounts. Compared to the geographical distribution of observation-based estimates of soil carbon, modelled values are generally higher over the forested areas including the Amazonian region, central Africa, and southeastern China.

The geographical distribution of annual burned area is broadly similar in the PRES, CTCOMP and LVCOMP sim-

ulations. All three simulations show higher area burned in savanna regions of the tropics (Fig. 8b) and lower values in extra tropical regions. CTEM, however, generally underestimates area burned in the savanna regions compared to observation-based estimates (e.g. GFED v. 3 Giglio et al., 2010). The simulations differ as area burned is especially sensitive to grassland extent. Larger fractional coverage of grasses in a grid cell implies a higher fire spread rate (see Appendix A9) and lower grid-averaged root-zone soil moisture content (since grasses are shallow rooted compared to trees), both of which increase area burned.

Overall the impact of dynamically modelling fractional coverage of PFTs in the CTCOMP simulation yields the largest differences for simulated vegetation biomass and annual area burned by fire, while other simulated primary carbon pools and fluxes remain similar to those in the PRES simulation and to observation-based estimates.

5 Discussion and conclusions

Modelling vegetation spatial dynamics explicitly in DGVMs overcomes the limitations inherent in prescribing a vegetation cover that is unable to respond to changes in climate, atmospheric CO_2 concentration and other forcings. Current approaches used in global-scale DGVMs include (i) the use of unmodified L–V equations (e.g. as in TRIFFID Cox, 2001 and USCM Brentnall et al., 2005), (ii) simple approaches that assume the best performing PFT can occupy more space (typically ranked by their NPP, e.g. LPJ; Sitch et al., 2003)
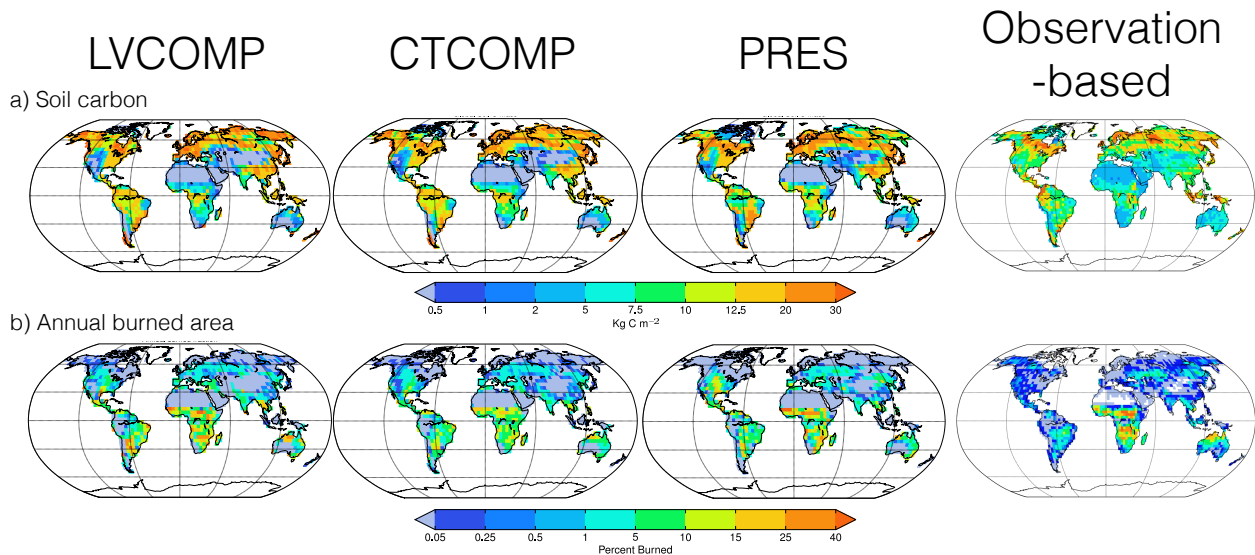
Figure 8. Geographical distribution of modelled soil carbon mass **(a)**, and percentage of annual burned area **(b)** based on simulations that use dynamically modelled fractional coverage of PFTs based on the original L–V equations (LVCOMP, first column) and the modified L–V equations (CTCOMP, second column). The third column shows the results for the simulation that uses prescribed PFT fractional cover based on the modified Wang et al. (2006) data set simulation (PRES). The fourth column has observation-based data sets for soil carbon mass (HWSD v. 1.2; FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012) and burned area (GFED v. 3.1; Giglio et al., 2006).

and (iii) approaches that account for competitiveness via both NPP and mortality (e.g. JSBACH; Brovkin et al., 2009). Most of these approaches suffer from an amplified expression of dominance with the most dominant PFT occupying an unrealistically large fraction of a grid cell, thus allowing little coexistence. Examples can be seen in Cramer et al. (2001) (see Fig. 2 for IBIS, HYBRID, LPJ and TRIFFID) where forests tend to dominate regions, which should otherwise be savannas, and Zeng et al. (2008) (their Fig. 7a) where the CLM-DGVM, which uses the competition module of the LPJ-DGVM, simulates ≥ 90 % coverage of trees where they exist in the model. In addition to these parametrizations of competitive interactions, global-scale models use bioclimatic limits to impose constraints on where PFTs can attempt colonization.

The strength of bioclimatic limits imposed within models varies greatly. At one end of the spectrum there are biogeographic models where present-day geographical distributions of PFTs are used to derive climate envelopes within which PFTs can exist. While this approach gives information about a PFT's areal extent under present-day conditions, the information is essentially binary, i.e. within a grid cell the PFTs exist or they do not. This biogeography approach neither provides any information about the fraction of a grid cell that the PFTs occupy nor how the PFTs' competitive success might evolve under changing environmental conditions. At the other end of the spectrum are parametrizations that require no bioclimatic information to limit species extent, i.e. each PFT's geographic extent is solely derived from its physiological responses and competitive interactions with other

PFTs arising from modelled processes. We are not aware of any competition parametrizations used within a global-scale DGVM that does not require the use of bioclimatic constraints. There are, however, recent attempts at developing models without bioclimatic limitations at the regional-scale. Fisher et al. (2015) recently implemented the ED concept into the CLM-DGVM to model competition between needleleaf evergreen and broadleaf cold deciduous trees based on their leaf traits in the eastern United States. Their attempt at prognostically determining biome boundaries for these two PFTs has some success, but only for particular model traits and structural assumptions. Attempts like these are promising but much work needs to be done to bring this concept to the global scale, across the full suite of PFTs, and with reasonable computational cost. At present, most models fall somewhere in-between these two endpoints with the aim of removing imposed bioclimatic constraints and allowing model physiologic processes to determine PFT geographical extents. The inability to globally model realistic geographic distributions of PFTs without the use of bioclimatic constraints is not surprising given our limited understanding of plant physiological processes that control PFT distributions and their competitive advantages.

We believe the use of bioclimatic constraints is reasonably moderate in CTEM. With the exception of the needleleaf deciduous PFT, the model uses fairly relaxed bioclimatic constraints for its tree PFTs. We do not use any bioclimatic limits for the C_3 and C_4 grass PFTs. The grasses' simulated distribution, including the competition between them and their coexistence in the tropics, is the result solely of the physio-
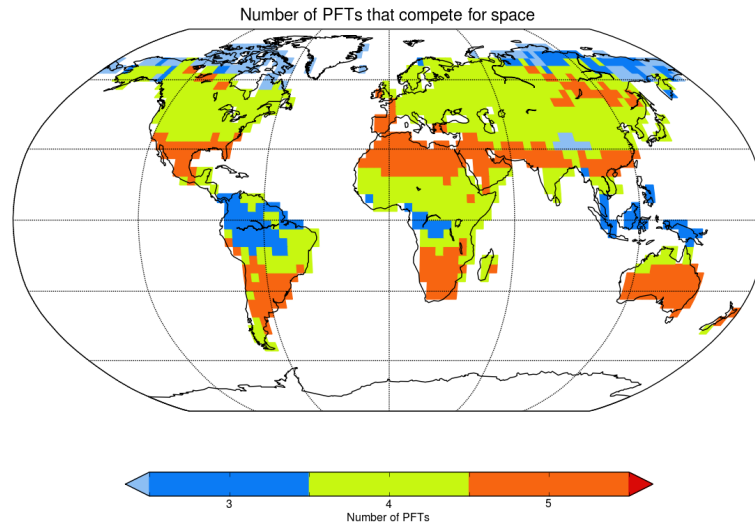
Figure 9. Plot showing the maximum number of PFTs in the pre-industrial equilibrium CTCOMP and LVCOMP simulations that can compete for space within a grid cell based upon each PFT's bioclimatic limits. The total number of natural non-crop PFTs represented in CTEM is seven.

logical processes that are explicitly modelled. The result is that the number of natural, non-crop PFTs that can exist and attempt to colonize a grid cell varies from two (in extreme climates like the high Arctic or alpine where trees cannot survive) to a maximum of five, out of a possible total of seven (Fig. 9). For these two to five PFTs that can exist in every grid cell, the model uses a modified version of the L–V equations, as introduced in Arora and Boer (2006b), to explicitly model competition between them.

Equilibrium offline simulations corresponding to the year 1861, using the original and modified L–V equations show that the model is able to capture the broad geographical distributions of tree and grass cover, and the bare fraction, when compared to the observation-based modified W2006 data set, especially when using the modified L–V equations. The global areas covered by tree and grass PFTs, the bare fraction, and the individual PFTs also compare reasonably to the observation-based estimates of the modified W2006 data set and those derived from a MODIS product (Friedl et al., 2013). Indeed, Fig. 5 demonstrates that the CTCOMP simulation, which uses modified version of the L–V equations, is better able to reproduce the observation-based PFT distributions of the modified W2006 data set for all PFTs, than the LVCOMP simulation, which uses the original L–V equations. The LVCOMP simulation is shown to suffer from an amplified expression of dominance, which results in high tree cover in regions where trees can exist and a sparse grass cover except in the Arctic region where grasses do not face competition from trees. This suggests that other models using the unmodified L–V equations would likely also suffer from this deficiency.

Some limitations are evident in our simulations. Overall, the simulated tree and grass cover is low in hot arid regions

and the simulated grass cover is overly high in cold arid regions. As a result the model generates more bare fraction in Australia, the Andes mountains of South America and the Kalahari desert of Africa, and too little bare fraction in the high Siberian Arctic. The simulated geographical distribution of individual PFTs appears reasonable although limitations remain here too. The model simulates a small fractional coverage of NDL-EVG trees in warm regions of southern Africa, Australia and South America that is not consistent with the observation-based estimate based on the modified W2006 data set. The gradation in the fractional coverage of PFTs from regions of high to low coverage is simulated much better when the modified version of the L–V equations is used, but the simulated fractional coverage is still not as graded as in the modified W2006 data set. These limitations are due primarily to (i) the coarse resolution used in our study, which does not allow resolution of climatic niches, (ii) the absence of shrub and moss/lichen PFTs and (iii) the small number of natural PFTs (seven) that are represented in CTEM. These constraints limit the model's ability to capture distributions of plants within the same broad functional group but that exist in geographically and climatically distinct regions, and will be the focus of future model development.

Modelling competition between PFTs in CTEM required adjusting some of the model parameters in order to realistically simulate the fractional coverage of its seven natural, non-crop PFTs. These parameter changes removed the positive bias in simulated vegetation biomass in certain high-latitude regions but yield too high vegetation biomass in the tropics. Overall, simulated global values of GPP, NPP, vegetation biomass, soil carbon and area burned in the CTCOMP simulation compare reasonably with observation-based and

other model estimates (Table 3). The zonal distributions of GPP and soil carbon also show reasonable comparison with observation-based estimates although larger differences are seen for vegetation biomass with higher simulated values in the tropics, as mentioned above (Fig. 6).

Despite its limitations, the behaviour of the competition module that uses the modified L–V equations, in the CLASS–CTEM modelling framework, is sufficiently realistic to yield a tool with which to study the impact of changes in climate and atmospheric CO_2 concentration on the global distribution of vegetation.

Appendix A: Description of model components and parametrizations

A1 The model structure

The basic model structure of CTEM includes three live vegetation components (leaf (L), stem (S) and root (R)) and two dead carbon pools (litter or detritus (D) and soil carbon (H)). The amount of carbon in these pools (C_L, C_S, C_R, C_D, C_H, $kg\,C\,m^{-2}$) is tracked prognostically through the fluxes in and out of them. The rate change equations for carbon in these pools are summarized in Sect. A7 after the processes leading to the calculation of fluxes in and out of these pools are introduced in the following sections.

A2 Photosynthesis and canopy conductance

A2.1 Net photosynthesis

All biogeochemical processes in CTEM are simulated at a daily time step except gross photosynthetic uptake and associated calculation of canopy conductance, which are simulated on a half hour time step with CLASS. The photosynthesis module of CTEM calculates the net canopy photosynthesis rate, which, together with atmospheric CO_2 concentration and vapour pressure or relative humidity, is used to calculate canopy conductance. This canopy conductance is then used by CLASS in its energy and water balance calculations.

The photosynthesis parametrization is based upon the approach of Farquhar et al. (1980) and Collatz et al. (1991, 1992) as implemented in SiB2 (Sellers et al., 1996) and MOSES (Cox et al., 1999) with some minor modifications as described in Arora (2003). Arora (2003) outlines four possible configurations for the model based on choice of a *big-leaf* or *two-leaf* (sunlight and shaded leaves) mode and stomatal conductance formulations based on either Ball et al. (1987) or Leuning (1995). The Ball et al. (1987) formulation uses relative humidity while Leuning (1995) uses vapour pressure deficit in calculation of canopy conductance. While the model remains capable of all four possible configurations, in practice, the model is usually run using the big-leaf parametrization with the stomatal conductance formulation of Leuning (1995), which is the configuration described here. The original description of the CTEM photosynthesis parametrization in Arora (2003) did not include discussion of all the PFTs simulated by CTEM, which we expand upon here and also include changes to the parametrization since version 1.0.

The gross leaf photosynthesis rate, G_o, depends upon the maximum assimilation rate allowed by the light (J_e), Rubisco (J_c) and transport capacity (J_s). The limitation placed on G_o by the amount of available light is calculated as ($mol\,CO_2\,m^{-2}\,s^{-1}$)

$$J_e = \begin{cases} \varepsilon(1-\nu)I\left[\frac{c_i-\Gamma}{c_i+2\Gamma}\right], & C_3 \text{ plants} \\ \varepsilon(1-\nu)I, & C_4 \text{ plants,} \end{cases} \quad (A1)$$

where I is the incident photosynthetically active radiation (PAR; $mol\,photons\,m^{-2}\,s^{-1}$), ν is the leaf scattering coefficient, with values of 0.15 and 0.17 for C_3 and C_4 plants, respectively, and ε is the quantum efficiency ($mol\,CO_2\,(mol\,photons)^{-1}$; values of 0.08 and 0.04 are used for C_3 and C_4 plants, respectively). c_i is the partial pressure of CO_2 in the leaf interior (Pa) and Γ is the CO_2 compensation point (Pa) (described below).

The Rubisco enzyme limited photosynthesis rate, J_c, is given by

$$J_c = \begin{cases} V_m\left[\frac{c_i-\Gamma}{c_i+K_c(1+O_a/K_o)}\right], & C_3 \text{ plants} \\ V_m, & C_4 \text{ plants,} \end{cases} \quad (A2)$$

where V_m is the maximum catalytic capacity of Rubisco ($mol\,CO_2\,m^{-2}\,s^{-1}$), adjusted for temperature and soil moisture, as described below. K_o and K_c are the Michaelis–Menten constants for O_2 and CO_2, respectively. O_a is the partial pressure (Pa) of oxygen.

The transport capacity (J_s) limitation determines the maximum capacity to transport the products of photosynthesis for C_3 plants, while for C_4 plants it represents CO_2 limitation

$$J_s = \begin{cases} 0.5V_m, & C_3 \text{ plants} \\ 2\times10^4\,V_m\frac{c_i}{p}, & C_4 \text{ plants} \end{cases} \quad (A3)$$

where p is surface atmospheric pressure (Pa).

V_m is calculated as

$$V_m = \frac{V_{max}f_{25}(2.0)S_{root}(\theta)\times10^{-6}}{[1+\exp0.3(T_c-T_{high})][1+\exp0.3(T_{low}-T_c)]}, \quad (A4)$$

where T_c is the canopy temperature (°C) and T_{low} and T_{high} are PFT-dependent lower and upper temperature limits for photosynthesis (Table A1). f_{25} is the standard Q_{10} function at 25 °C ($f_{25}(Q_{10}) = Q_{10}^{(0.1(T_c-25))}$) and V_{max} is the PFT-dependent maximum rate of carboxylation by the enzyme Rubisco ($mol\,CO_2\,m^{-2}\,s^{-1}$; Table A1). The constant 10^{-6} converts V_{max} from units of $\mu mol\,CO_2\,m^{-2}\,s^{-1}$ to $mol\,CO_2\,m^{-2}\,s^{-1}$.

The influence of soil moisture stress is simulated via $S_{root}(\theta)$, which represents a soil moisture stress term formulated as

$$S_{root}(\theta) = \sum_{i=1}^{g} S(\theta_i)r_i, \quad (A5)$$

$$S(\theta_i) = [1 - \{1 - \phi_i\}]^\varrho, \quad (A6)$$

Table A1. Parameters used in the photosynthesis module of CTEM v. 2.0. T_{low} and T_{high} are the lower and upper temperature limits, respectively, used to determine the rate of carboxylation of the enzyme Rubiso (Eq. A4). V_{max} is the maximum rate of carboxylation by Rubisco (Eq. A4). ϱ is a parameter describing the PFT sensitivity to soil moisture stress (Eq. A6). k_n is the extinction coefficient describing the profile of the nitrogen and time mean PAR along the depth of the canopy (Eq. A16).

Plant functional type short name	T_{low} (°C)	T_{high} (°C)	V_{max} ($\mu\text{mol}\,CO_2\,\text{m}^{-2}\,\text{s}^{-1}$)	k_n (unitless)	ϱ (unitless)
NDL-EVG	−5	34	62	0.50	2
NDL-DCD	−5	34	47	0.50	2
BDL-EVG	0	45	48	0.50	4
BDL-DCD-COLD	0	37	57	0.50	2
BDL-DCD-DRY	0	37	40	0.50	2
C3-CROP	−3	42	55	0.40	2
C4-CROP	5	42	40	0.48	2
C3-GRASS	−1	40	75	0.46	2
C4-GRASS	10	50	15	0.44	2

where $S_{\text{root}}(\theta)$ is calculated by weighting $S(\theta_i)$ with the fraction of roots, r_i, in each soil layer i and ϱ is a PFT-specific sensitivity to soil moisture stress (unitless; Table A1). ϕ_i is the degree of soil saturation (soil wetness) given by

$$\phi_i(\theta_i) = \max\left[0, \min\left(1, \frac{\theta_i - \theta_{i,\text{wilt}}}{\theta_{i,\text{field}} - \theta_{i,\text{wilt}}}\right)\right], \tag{A7}$$

where θ_i is the volumetric soil moisture ($\text{m}^3\,\text{water}\,(\text{m}^3\,\text{soil})^{-1}$) of the ith soil layer and $\theta_{i,\text{field}}$ and $\theta_{i,\text{wilt}}$ the soil moisture at field capacity and wilting point, respectively.

The CO_2 compensation point (Γ) is the CO_2 partial pressure where photosynthetic uptake equals the leaf respiratory losses (used in Eqs. A1 and A2). Γ is zero for C_4 plants but is sensitive to oxygen partial pressure for C_3 plants as

$$\Gamma = \begin{cases} \dfrac{O_a}{2\sigma}, & C_3 \text{ plants} \\ 0, & C_4 \text{ plants}, \end{cases} \tag{A8}$$

where σ is the selectivity of Rubisco for CO_2 over O_2 (unitless), estimated by $\sigma = 2600 f_{25}(0.57)$. The CO_2 (K_c) and O_2 (K_o) Michaelis–Menten constants used in Eq. (A2) are determined via

$$K_c = 30 f_{25}(2.1), \tag{A9}$$

$$K_o = 3 \times 10^4 f_{25}(1.2). \tag{A10}$$

Given the light (J_e), Rubsico (J_c) and transportation capacity (J_s) limiting rates, the leaf-level gross photosynthesis rate, G_o ($\text{mol}\,CO_2\,\text{m}^{-2}\,\text{s}^{-1}$), is then determined following a minimization based upon smallest roots of the following two quadratic equations

$$J_p = \frac{(J_c + J_e) \pm \sqrt{(J_c + J_e)^2 - 4\beta_1(J_c + J_e)}}{2\beta_1}, \tag{A11}$$

$$G_o = \frac{(J_p + J_s) \pm \sqrt{(J_p + J_s)^2 - 4\beta_2(J_p + J_s)}}{2\beta_2}, \tag{A12}$$

where β_1 is 0.95 and β_2 is 0.99. When soil moisture stress is occurring, both the J_s and J_c terms are reduced since the V_m term (Eq. A4) includes the effect of soil moisture stress through the $S(\theta)$ term and this reduces the leaf-level gross photosynthesis rate.

The current version of CTEM does not include nutrient constraints on photosynthesis and, as a result, increasing atmospheric CO_2 concentration leads to unconstrained increase in photosynthesis. In natural ecosystems, however, down regulation of photosynthesis occurs due to constraints imposed by availability of nitrogen, as well as phosphorus. To capture this effect, CTEM uses a nutrient limitation term, based on experimental plant growth studies, to down regulate the photosynthetic response to elevated CO_2 concentrations (Arora et al., 2009). The parametrization, and its rationale, are fully described in Arora et al. (2009) but the basic relations are summarized here. The leaf-level gross photosynthetic rate is scaled by the down-regulation term, Ξ_N, to yield the nutrient limited leaf level gross photosynthetic rate as

$$G_{o,\text{N-limited}} = \Xi_N G_o, \tag{A13}$$

$$\Xi_N = \frac{1 + \gamma_{gd} \ln(c_a/c_0)}{1 + \gamma_g \ln(c_a/c_0)}, \tag{A14}$$

where c_a is the atmospheric CO_2 concentration in ppm, c_0 is the pre-industrial CO_2 concentration (285.0 ppm), γ_g is 0.95 (unitless; see Arora et al., 2009). A value of γ_{gd} lower than γ_g ensures that Ξ_N gradually decreases from its pre-industrial value of one as c_a increases to constrain the rate of increase of photosynthesis with rising atmospheric CO_2 concentrations. CTEM v. 2.0 uses a γ_{gd} value of 0.30 (unitless).

Finally, the leaf-level gross photosynthesis rate, $G_{o,\text{N-limited}}$ is scaled up to the canopy-level, G_{canopy}, by considering the exponential vertical profile of radiation along the depth of the canopy as

$$G_{\text{canopy}} = G_{o,\text{N-limited}} f_{\text{PAR}}, \tag{A15}$$

$$f_{\text{PAR}} = \frac{1}{k_n}(1 - \exp^{-k_n \text{LAI}}), \tag{A16}$$

which yields the gross primary productivity (G_{canopy}, GPP). k_n is the extinction coefficient that describes the nitrogen and time-mean photosynthetically absorbed radiation (PAR) profile along the depth of the canopy (Table A1) (Ingestad and Lund, 1986; Field and Mooney, 1986), and LAI (m^2 leaf $(m^2$ ground)$^{-1}$) is the leaf area index.

The net canopy photosynthetic rate, $G_{\text{canopy,net}}$ ($\text{mol CO}_2 \, m^{-2} \, s^{-1}$), is calculated by subtracting canopy leaf maintenance respiration costs (R_{mL}; see Sect. A3.1) as

$$G_{\text{canopy,net}} = G_{\text{canopy}} - R_{\text{mL}}. \tag{A17}$$

A2.2 Coupling of photosynthesis and canopy conductance

When using the Leuning (1995) approach for photosynthesis–canopy conductance coupling, canopy conductance (g_c; $\text{mol} \, m^{-2} \, s^{-1}$) is expressed as a function of the net canopy photosynthesis rate, $G_{\text{canopy, net}}$, as

$$g_c = m \frac{G_{\text{canopy,net}} p}{(c_s - \Gamma)} \frac{1}{(1 + V/V_o)} + b\text{LAI} \tag{A18}$$

where p is the surface atmospheric pressure (Pa), the parameter m is set to 9.0 for needle-leaved trees, 12.0 for other C_3 plants and 6.0 for C_4 plants, parameter b is assigned the values of 0.01 and 0.04 for C_3 and C_4 plants, respectively. V is the vapour pressure deficit (Pa) and the parameter V_o is set to 2000 Pa for trees and 1500 Pa for crops and grasses. The partial pressure of CO_2 at the leaf surface, c_s, is found via

$$c_s = c_{\text{ap}} - \frac{1.37 G_{\text{canopy,net}} p}{g_b}. \tag{A19}$$

Here, c_{ap} is the atmospheric CO_2 partial pressure (Pa) and g_b is the aerodynamic conductance estimated by CLASS ($\text{mol} \, m^{-2} \, s^{-1}$). The intra-cellular CO_2 concentration required in Eqs. (A1)–(A3) is calculated as

$$c_i = c_s - \frac{1.65 G_{\text{canopy,net}} p}{g_c}. \tag{A20}$$

Since calculations of $G_{\text{canopy,net}}$ and c_i depend on each other, the photosynthesis-canopy conductance equations need to be solved iteratively. The initial value of c_i used in calculation of $G_{\text{canopy,net}}$ is the value from the previous time step or, in its absence, c_i is assumed to be $0.7 c_{\text{ap}}$.

Canopy (g_c) and aerodynamic (g_b) conductance used in above calculations are expressed in units of $\text{mol CO}_2 \, m^{-2} \, s^{-1}$

but can be converted to the traditional units of m s^{-1} as follows

$$g_c(\text{m s}^{-1}) = 0.0224 \frac{T_c}{T_f} \frac{p_0}{p} g_c(\text{mol} \, m^{-2} \, s^{-1}), \tag{A21}$$

where p_0 is the standard atmospheric pressure (101 325 Pa) and T_f is freezing temperature (273.16 K).

A3 Respiration

Version 2.0 of CTEM calculates autotrophic respiratory fluxes (R_a) from the living vegetation components and heterotrophic respiratory fluxes (R_h) from the dead litter and soil carbon pools based on the approach described in Arora (2003) but with some modifications.

A3.1 Maintenance and growth respiration

Autotrophic respiration ($\text{mol CO}_2 \, m^{-2} \, s^{-1}$) is composed of maintenance, R_m, and growth respirations, R_g,

$$R_a = R_m + R_g. \tag{A22}$$

Maintenance respiration accounts for carbon consumed by processes that keep existing plant tissues alive and is a function of environmental stresses. Maintenance respiration is calculated on a half-hourly time step (with photosynthesis) for the leaves, R_{mL}, and at a daily time step for the stem, R_{mS}, and root, R_{mR}, components

$$R_m = R_{\text{mL}} + R_{\text{mS}} + R_{\text{mR}}. \tag{A23}$$

Maintenance respiration is generally strongly correlated with nitrogen content (Reich et al., 1998; Ryan, 1991). The current version of CTEM does not explicitly track nitrogen in its vegetation components. Therefore, we adopt the approach of Collatz et al. (1991, 1992) in which the close relation between maximum catalytic capacity of Rubisco, V_m, and leaf nitrogen content is used as a proxy to estimate leaf maintenance respiration,

$$R_{\text{mL}} = \varsigma_L V_m f_{25}(Q_{10d,n}) f_{\text{PAR}}, \tag{A24}$$

where ς_L is set to 0.015 and 0.025 for C_3 and C_4 plants, respectively, f_{PAR} scales respiration from the leaf to the canopy level, similar to Eq. (A15), and the $f_{25}(Q_{10d,n})$ function accounts for different temperature sensitivities of leaf respiration during day (d) and night (n). Pons and Welschen (2003) and Xu and Baldocchi (2003) suggest lower temperature sensitivity for leaf respiration during the day compared to night, and therefore we use values of $Q_{10d} = 1.3$ and $Q_{10n} = 2.0$ for day and night, respectively.

Maintenance respiration from the stem and root components is estimated based on PFT-specific base respiration rates (ς_S and ς_R specified at 15 °C, $\text{kg C (kg C)}^{-1} \, \text{yr}^{-1}$; Table A2) that are modified to account for temperature response

Table A2. PFT-specific base respiration rates $(\mathrm{kg\,C\,(kg\,C)^{-1}\,yr^{-1}})$ for the stem (ς_S) and root (ς_R) components (Eq. A25) used in the autotrophic respiration module of CTEM v. 2.0. Some parameter values differ between the model versions when PFTs' fractional coverage is specified and when they are dynamically modelled using competition between PFTs. If modified, the parameter values for the dynamically modelled version are shown in parentheses.

Plant functional type short name	ς_S	ς_R
NDL-EVG	0.090 (0.070)	0.500
NDL-DCD	0.055	0.285
BDL-EVG	0.060 (0.050)	0.650 (0.400)
BDL-DCD-COLD	0.034	0.225
BDL-DCD-DRY	0.030 (0.035)	0.055 (0.150)
C3-CROP	0.037	0.160
C4-CROP	0.037	0.160
C3-GRASS	–	0.100
C4-GRASS	–	0.100

Table A3. PFT-specific base respiration rates $(\mathrm{kg\,C\,(kg\,C)^{-1}\,yr^{-1}})$ for the litter (ς_D) and soil carbon (ς_H) pools (Eq. A31) and the humification factor $(\chi$, Eq. A37) used in the heterotrophic respiration module of CTEM v. 2.0.

Plant functional type short name	ς_D	ς_H	χ
NDL-EVG	0.4453	0.0260	0.42
NDL-DCD	0.5986	0.0260	0.42
BDL-EVG	0.6339	0.0208	0.53
BDL-DCD-COLD	0.7576	0.0208	0.48
BDL-DCD-DRY	0.6957	0.0208	0.48
C3-CROP	0.6000	0.0350	0.10
C4-CROP	0.6000	0.0350	0.10
C3-GRASS	0.5260	0.0125	0.42
C4-GRASS	0.5260	0.0125	0.42

following a Q_{10} function. Maintenance respiration from stem and root components, $R_{\mathrm{m\{S,R\}}}$, is calculated as

$$R_{\mathrm{m},i} = 2.64 \times 10^{-6}\varsigma_i l_{v,i} C_i f_{15}(Q_{10}), \quad i = \mathrm{S, R}, \quad \text{(A25)}$$

where $l_{v,i}$ is the live fraction of stem or root component, i.e. the sapwood, and C_i is the stem or root carbon mass $(\mathrm{kg\,C\,m^{-2}})$. The constant 2.64×10^{-6} converts units from $\mathrm{kg\,C\,m^{-2}\,yr^{-1}}$ to $\mathrm{mol\,CO_2\,m^{-2}\,s^{-1}}$. The live sapwood fraction, $l_{v,i}$, for stem or root component is calculated following the CENTURY model (Parton et al., 1996) as

$$l_{v,i} = \max(0.05, \min[1.0, \exp^{-0.2835 C_i}]), \quad i = \mathrm{S, R.} \quad \text{(A26)}$$

The Q_{10} value used in Eq. (A25) is not assumed to be constant but modelled as a function of temperature following Tjoelker et al. (2001) as

$$Q_{10} = 3.22 - 0.046 \left(\frac{15.0 + T_{\{S,R\}}}{1.9}\right), \quad \text{(A27)}$$

where $T_{\{S,R\}}$ is stem or root temperature (°C). Stem temperature is assumed to be the same as air temperature while root temperature is based on the soil temperature weighted by the fraction of roots present in each soil layer (Arora and Boer, 2003). The calculated Q_{10} value is additionally constrained to be between 1.5 and 4.0.

Growth respiration, R_{g} $(\mathrm{mol\,CO_2\,m^{-2}\,s^{-1}})$, is estimated as a fraction $(\epsilon_{\mathrm{g}} = 0.15)$ of the positive gross canopy photosynthetic rate after maintenance respiration has been accounted for

$$R_{\mathrm{g}} = \epsilon_{\mathrm{g}}\max[0, (G_{\mathrm{canopy}} - R_{\mathrm{m}})]. \quad \text{(A28)}$$

Finally, net primary productivity (NPP) is calculated as

$$\mathrm{NPP} = G_{\mathrm{canopy}} - R_{\mathrm{m}} - R_{\mathrm{g}}. \quad \text{(A29)}$$

A3.2 Heterotrophic respiration

Heterotrophic respiration, R_{h} $(\mathrm{mol\,CO_2\,m^{-2}\,s^{-1}})$, in CTEM is based on respiration from the litter (which includes contributions from the stem, leaf and root components), $R_{\mathrm{h,D}}$, and soil carbon, $R_{\mathrm{h,H}}$, pools,

$$R_{\mathrm{h}} = R_{\mathrm{h,D}} + R_{\mathrm{h,H}}. \quad \text{(A30)}$$

Heterotrophic respiration is regulated by soil temperature and moisture and is calculated on a daily time step. The original heterotrophic respiration scheme is described in Arora (2003) while the modified parametrization used in CTEM v. 2.0 is detailed in Melton et al. (2015) and is briefly described here. Respiration from the litter and soil carbon pools takes the following basic form

$$R_{\mathrm{h},i} = 2.64 \times 10^{-6}\varsigma_i C_i f_{15}(Q_{10}) f(\Psi)_i,$$
$$i = \mathrm{D, H.} \quad \text{(A31)}$$

The soil carbon and litter respiration depends on the amount of carbon in these components $(C_{\mathrm{H}}$ and C_{D}; $\mathrm{kg\,C\,m^{-2}})$ and on a PFT-dependent respiration rate specified at 15 °C $(\varsigma_{\mathrm{H}}$ and ς_{D}; $\mathrm{kg\,C\,(kg\,C)^{-1}\,yr^{-1}}$; Table A3). The constant 2.64×10^{-6} converts units from $\mathrm{kg\,C\,m^{-2}\,yr^{-1}}$ to $\mathrm{mol\,CO_2\,m^{-2}\,s^{-1}}$.

The effect of soil moisture is accounted for via dependence on soil matric potential $(f(\Psi))$, described later. The temperature dependency of microbial soil respiration rates has been estimated by several different formulations, ranging from simple Q_{10} (exponential) to Arrhenius-type formulations (see review by Lloyd and Taylor, 1994). In CTEM, soil temperature influences heterotrophic respiration through a temperature-dependent Q_{10} function $(f_{15}(Q_{10}))$. The value of Q_{10} itself is assumed to be a function of temperature following a hyperbolic tan function:

$$Q_{10} = 1.44 + 0.56 \tanh[0.075(46.0 - T_i)],$$
$$i = \mathrm{D, H,} \quad \text{(A32)}$$

where $T_{\{D,H\}}$ is the temperature of either the litter or soil carbon pool (°C). The parametrization is a compro-

mise between the temperature-independent Q_{10} commonly found in many terrestrial ecosystem models (e.g. Cox, 2001) and the temperature-dependent Q_{10} of Kirschbaum (1995). While a constant Q_{10} yields an indefinitely increasing respiration rate with increasing temperature, the formulation of Kirschbaum (1995) gives a continuously increasing Q_{10} under decreasing temperature, which leads to unreasonably high soil and litter carbon pools at high latitudes in CTEM. The CTEM parametrization avoids these issues with a Q_{10} value of about 2.0 for temperatures less than 20 °C, while a decreasing value of Q_{10} at temperatures above 20 °C ensures that the respiration rate does not increase indefinitely.

The temperature of the litter pool is a weighted average of the temperature of the top soil layer (T_1) and the root temperature (T_R) as litter consists of leaf, stem, and root litter ($T_D = 0.7T_1 + 0.3T_R$). The temperature of the soil carbon pool is calculated as the mean soil temperature in the rooting zone based upon the fraction of roots in each soil layer and their temperature. The carbon in each soil layer is not explicitly tracked but assumed to adopt an exponential distribution (similar to roots; e.g. Fig. 4 of Jobbágy and Jackson, 2000).

The response of heterotrophic respiration to soil moisture is formulated through soil matric potential (Ψ; MPa). While soil matric potential values are usually negative, the formulation uses absolute values to allow its logarithm to be taken. Absolute values of soil matric potential are high when soil is dry and low when it is wet. The primary premise of soil moisture control on heterotrophic respiration is that heterotrophic respiration is constrained both when the soils are dry (due to reduced microbial activity) and when they are wet (due to impeded oxygen supply to microbes) with optimum conditions in-between. The exception is the respiration from the litter component, which is assumed to be continually exposed to air, and thus never oxygen deprived, even when soil moisture content is high ($0.04 > |\Psi| \geq |\Psi_{sat}|$, where Ψ_{sat} is the soil matric potential at saturation). The soil moisture dependence thus varies between 0 and 1 with matric potential as follows:

for $0.04 > |\Psi| \geq |\Psi_{sat}|$

$$f(\Psi)_H = 1 - 0.5 \frac{\log(0.04) - \log|\Psi|}{\log(0.04) - \log|\Psi_{sat}|} \tag{A33}$$

$$f(\Psi)_D = 1;$$

for $0.06 \geq |\Psi| \geq 0.04$

$$f(\Psi)_{\{D,H\}} = 1; \tag{A34}$$

for $100.0 \geq |\Psi| > 0.06$

$$f(\Psi)_{\{D,H\}} = 1 - 0.8 \frac{\log|\Psi| - \log(0.06)}{\log(100) - \log(0.06)}; \tag{A35}$$

for $|\Psi| > 100.0$

$$f(\Psi)_{\{D,H\}} = 0.2. \tag{A36}$$

Heterotrophic respiration for bare ground is treated separately in CTEM. The carbon contributions to the bare ground litter and soil carbon pools come via processes such as creation of bare ground due to fire, competition between PFTs and land use change. The heterotrophic respiration is sensitive to temperature and moisture in the same manner as vegetated areas using Eqs. (A31)–(A36). The base respiration rates of $\varsigma_{D,bare}$ and $\varsigma_{H,bare}$ are set to 0.5605 and 0.02258 kg C (kg C)$^{-1}$ yr^{-1}, respectively.

The amount of humidified litter, which is transferred from the litter to the soil carbon pool ($C_{D \to H}$) is modelled as a fraction of litter respiration ($R_{h,D}$) as

$$C_{D \to H} = \chi R_{h,D} \tag{A37}$$

where χ (Table A3) is the PFT-dependent humification factor and varies between 0.4 and 0.5. For crops, χ is set to 0.1 to account for reduced transfer of humidified litter to the soil carbon pool which leads to loss in soil carbon when natural vegetation is converted to croplands. Over the bare ground fraction χ is set to 0.45.

With heterotrophic respiration known, net ecosystem productivity (NEP) is calculated as

$$NEP = G_{canopy} - R_m - R_g - R_h. \tag{A38}$$

A4 Allocation

Positive NPP is allocated daily to the leaf, stem and root components, which generally causes their respective biomass to increase, although the biomass may also decrease depending on the autotrophic respiration flux of a component. Negative NPP generally causes net carbon loss from the components. While CTEM offers the ability to use both specified constant or dynamically calculated allocation fractions for leaves, stems and roots, in practice the dynamic allocation fractions are primarily used. The formulation used in CTEM v. 2.0 differs from that for CTEM v. 1.0 as described in Arora and Boer (2005a) only in the parameter values.

The dynamic allocation to the live plant tissues is based on the light, water and leaf phenological status of vegetation. The preferential allocation of carbon to the different tissue pools is based on three assumptions: (i) if soil moisture is limiting, carbon should be preferentially allocated to roots for greater access to water, (ii) if LAI is low, carbon should be allocated to leaves for enhanced photosynthesis and finally (iii) carbon is allocated to the stem to increase vegetation height and lateral spread of vegetation when the increase in LAI results in a decrease in light penetration.

The vegetation water status, W, is determined as a linear scalar quantity that varies between 0 and 1 for each PFT and calculated by weighting the degree of soil saturation ($\phi_i(\theta_i)$, Eq. A7) with the fraction of roots in each soil layer

$$W = \phi_{root} = \sum_{i=1}^{g} \phi_i(\theta_i) r_i. \tag{A39}$$

The light status, L, is parametrized as a function of LAI and nitrogen extinction coefficient, k_n (PFT-dependent; see Table A1), as

$$L = \begin{cases} \exp(-k_n \text{LAI}), & \text{trees and crops} \\ \max\left(0, 1 - \frac{\text{LAI}}{4.5}\right), & \text{grasses.} \end{cases} \quad (A40)$$

For PFTs with a stem component (i.e. tree and crop PFTs), the fractions of positive NPP allocated to stem (a_{fS}), leaf (a_{fL}) and root (a_{fR}) components are calculated as

$$a_{fS} = \frac{\epsilon_S + \omega_a(1 - L)}{1 + \omega_a(2 - L - W)}, \quad (A41)$$

$$a_{fR} = \frac{\epsilon_R + \omega_a(1 - W)}{1 + \omega_a(2 - L - W)}, \quad (A42)$$

$$a_{fL} = \frac{\epsilon_L}{1 + \omega_a(2 - L - W)} = 1 - a_{fS} - a_{fR}. \quad (A43)$$

The base allocation fractions for each component (leaves – ϵ_L, stem – ϵ_S, and roots – ϵ_R) are PFT-dependent (Table A4) and sum to 1, i.e. $\epsilon_L + \epsilon_S + \epsilon_R = 1$. The parameter ω_a, which varies by PFT (Table A4), determines the sensitivity of the allocation scheme to changes in W and L. Larger values of ω_a yield higher sensitivity to changes in L and W.

Grasses do not have a stem component (i.e. $a_{fS} = 0$) and the allocation fractions for leaf and root components are given by

$$a_{fL} = \frac{\epsilon_L + \omega_a L}{1 + \omega_a(1 + L - W)}, \quad (A44)$$

$$a_{fR} = \frac{\epsilon_R + \omega_a(1 - W)}{1 + \omega_a(1 + L - W)}. \quad (A45)$$

The above equations ensure that the allocation fractions add up to one ($a_{fL} + a_{fR} + a_{fS} = 1$).

The dynamic allocation fractions are superseded under three conditions. First, during the leaf onset for crops and deciduous trees, all carbon must be allocated to leaves ($a_{fL} = 1$, $a_{fS} = a_{fR} = 0$). Second, the proportion of stem plus root biomasses to leaf biomass must satisfy the relationship:

$$C_S + C_R = \eta C_L^\kappa, \quad (A46)$$

where C_S, C_R and C_L are the carbon in the stem, root and leaves, respectively. The parameter η is PFT-specific (Table A4) and parameter κ has a value of 1.6 for trees and crops and 1.2 for grasses. Both parameters are based on the Frankfurt Biosphere Model (FBM) (Lüdeke et al., 1994). This constraint (Eq. A46) is based on the physical requirement of sufficient stem and root tissues to support a given leaf biomass. As grasses have no stem component, Eq. (A46) determines their root to shoot ratio (i.e. the ratio of below-ground to aboveground biomass). The final condition ensures that a minimum realistic root to shoot ratio is maintained for all PFTs (lr$_{min}$, listed in Table A4). Root mass is required for nutrient and water uptake and support for the aboveground biomass. If the minimum root to shoot ratio is not being maintained, carbon is allocated preferentially to roots.

A5   Leaf phenology

The leaf phenology parametrization used in CTEM v. 1.0 is described in detail by Arora and Boer (2005a). Changes between version 1.0 and 2.0 are limited to parameter values and the parametrization is briefly described here. There are four different leaf phenological states in which vegetation can be at a given instant: (i) no leaves or dormant, (ii) maximum growth, (iii) normal growth and (iv) leaf fall or harvest. PFTs may go through only some, or all, of these phenological states depending on their deciduousness. A broadleaf cold deciduous tree, for example, transitions through all these four states in a year. In winter, the broadleaf cold deciduous trees are in the no leaves/dormant state; favourable climatic conditions in spring trigger leaf growth and the tree enters the maximum leaf growth state when all the NPP is allocated to leaves to accelerate leaf out; when the LAI reaches a threshold (described below) the tree enters the normal leaf growth state and NPP is also allocated to stem and root components; finally the arrival of autumn triggers leaf fall and the trees go into the leaf fall mode where no carbon is allocated to leaves (but it continues for roots and stems). When all the leaves have been shed, the trees go into the no leaves or dormant state again and the cycle is repeated the next year. The evergreen tree PFTs and the grass PFTs do not enter the leaf fall state and maintain a leaf canopy as long as environmental conditions are favourable. Although drought and cold stress cause accelerated leaf loss compared to the normal leaf turnover from these PFTs, they do not explicitly go into the leaf fall mode where the intent is to lose all leaves in a specified amount of time.

The leaf phenological state transitions are dependent upon environmental conditions. In particular, the transition from no leaves/dormant state to the maximum growth state is based on the carbon-gain approach. CTEM uses *virtual* leaves to assess favourable meteorological conditions for leaf out. The virtual leaves photosynthesize and respire in a manner similar to normal leaves except the carbon gain or loss is not taken into account in vegetation's carbon balance. A positive net leaf photosynthesis rate ($G_{canopy,net}$, Eq. A17) for the virtual leaves over seven consecutive days indicates the arrival of favourable growth conditions and triggers leaf onset and the associated transition from the no leaves/dormant state to the maximum leaf growth state, when the entire positive NPP is allocated to leaves ($a_{fL} = 1$, $a_{fS} = a_{fR} = 0$). When LAI reaches LAI$_{thrs}$ then the vegetation switches to the normal growth mode and positive NPP is allocated to all three vegetation components – leaves, stem and roots ($a_{fL}, a_{fS}, a_{fR} > 0$). LAI$_{thrs}$ is calculated as

$$\text{LAI}_{thrs} = L_f \left[\text{SLA} \left(\frac{C_S + C_R}{\eta} \right)^{1/\kappa} \right]. \quad (A47)$$

The PFT-specific L_f term (see Table A5) calculates LAI$_{thrs}$ to be typically between 40 and 50 % of the maximum LAI that

Table A4. PFT-specific base allocation fractions for leaf (ϵ_L), stem (ϵ_S) and root (ϵ_R) components (Eqs. A41 to A43), the parameter ω_a, which determines the sensitivity of allocation to light and water status, the parameter η (Eq. A46), which determines proportion of stem plus wood biomass to leaf biomass and the minimum root : shoot ratio lr_{min} used in the allocation module of CTEM v. 2.0. Some parameter values differ between the model versions when PFTs' fractional coverage is specified and when they are dynamically modelled using competition between PFTs. If modified, the parameter values for the dynamically modelled version are shown in parentheses.

Plant functional type short name	ϵ_L	ϵ_S	ϵ_R	ω_a	η	lr_{min}
NDL-EVG	0.20 (0.19)	0.15 (0.40)	0.65 (0.41)	0.8	10.0	0.16
NDL-DCD	0.06 (0.45)	0.05 (0.34)	0.89 (0.21)	0.5	30.8	0.16
BDL-EVG	0.35 (0.39)	0.05 (0.21)	0.60 (0.40)	0.8	31.0	0.16
BDL-DCD-COLD	0.35 (0.50)	0.10 (0.35)	0.55 (0.15)	0.8 (0.45)	50.0	0.16
BDL-DCD-DRY	0.25 (0.30)	0.10	0.65 (0.60)	0.8	70.0	0.32
C3-CROP	0.80	0.15	0.05	0.05	7.0	0.16
C4-CROP	0.80	0.15	0.05	0.05	7.0	0.16
C3-GRASS	0.01 (0.10)	–	0.99 (0.90)	1.0	3.0	0.5
C4-GRASS	0.01 (0.10)	–	0.99 (0.90)	1.0	3.0	0.5

a given amount of stem and root biomass can support (based on the terms in the square brackets and Eq. (A46). SLA is the specific leaf area (Eq. A56). This rule for transition from a maximum to a normal growth state is also used for evergreen tree PFTs and grass PFTs. Similar to LAI$_{thrs}$, the LAI of virtual leaves is 7.5 % of the maximum LAI a given amount of root and stem biomass can support for tree and crop PFTs and 2.5 % for grass PFTs. In addition, the LAI of virtual leaves is constrained to be, at least, 0.3 m^2 m^{-2} for tree PFTs and 0.2 m^2 m^{-2} for crop and grass PFTs.

The transition from the normal growth state to the leaf fall state is triggered by unfavourable environmental conditions and shorter day length. Broadleaf deciduous trees transition to the leaf fall state when either: (i) day length is less than 11 h and the rooting zone temperature drops below 11.15 °C or (ii) when the rooting zone temperature drops below 8 °C regardless of the day length. Needleleaf deciduous tress begin leaf fall after seven consecutive days with daily mean air temperature below −5 °C. Leaf fall occurs over a period of 15 days. In the leaf fall state, the vegetation continues carbon allocation to its root and stem components, but not to leaves ($a_{fL} = 0$, $a_{fS} + a_{fR} = 1$). Evergreen trees and grasses do not enter the leaf fall state and neither do the broadleaf drought deciduous trees. The implication for the latter PFT is that if the climate changes and the dry season becomes shorter, then the trees will keep their leaves on for a longer period of time since broadleaf drought deciduous trees lose leaves due to soil moisture stress (described below).

The model vegetation is able to transition between the different leaf phenological states in response to changing conditions. For example, a leaf out in spring for broadleaf cold deciduous trees can be interrupted by a cold event when the vegetation goes into a leaf fall state until the return of more favourable conditions.

Leaf litter generation is caused by normal turnover of leaves (Ω_N, day^{-1}) and also due to cold (Ω_C, day^{-1}) and

drought (Ω_D, day^{-1}) stress, both of which contribute to seasonality of LAI. For example, the leaf loss associated with drought and reduced photosynthesis during the dry season are the principal causes of the seasonality of LAI for the broadleaf drought deciduous tree PFT.

The conversion of leaf carbon to leaf litter (D_L, kg C m^{-2} day^{-1}) is expressed as

$$D_L = C_L[1 - \exp(-\Omega_N - \Omega_C - \Omega_D)], \tag{A48}$$

where ($\Omega_{N,C,D}$, day^{-1}) are the leaf loss rates associated with normal turnover of leaves and the cold and drought stress. The rate of normal turnover of leaves is governed by PFT-specific leaf lifespan (τ_L, yr) as $\Omega_N = 1/365\tau_L$ (See Table A6 for PFT specific values of τ_L). The leaf loss rate associated with cold stress (Ω_C) is calculated as

$$\Omega_C = \Omega_{C,max} L_{cold}^3, \tag{A49}$$

where $\Omega_{C,max}$ (day^{-1}, Table A5) is the maximum cold stress loss rate. L_{cold} is a scalar that varies between 0 and 1 as

$$L_{cold} = \begin{cases} 1, & T_a < (T_{cold}^{leaf} - 5) \\ 1 - \frac{T_a - (T_{cold}^{leaf} - 5)}{5}, & \\ \quad T_{cold}^{leaf} > T_a > (T_{cold}^{leaf} - 5) \\ 0, & T_a > T_{cold}^{leaf}, \end{cases} \tag{A50}$$

where T_{cold}^{leaf} is a PFT-specific temperature threshold below which a PFT experiences damage to its leaves promoting leaf loss (Table A5) and T_a is the daily mean air temperature (°C). The leaf loss rate due to drought stress is calculated in a similar manner

$$\Omega_D = \Omega_{D,max} (1 - \phi_{root})^3, \tag{A51}$$

where $\Omega_{D,max}$ (day^{-1}, Table A5) is the maximum drought stress loss rate and ϕ_{root} (Eq. A39) is the degree of soil saturation in the rooting zone.

Table A5. PFT-specific parameters used in the phenology module of CTEM v. 2.0: maximum cold ($\Omega_{C,max}$, day^{-1}, Eq. A49) and drought ($\Omega_{D,max}$, day^{-1}, Eq. A51) stress loss rates, T_{cold}^{leaf} (°C) temperature threshold for determining cold stress scalar L_{cold} (Eq. A50) and the fraction L_f (Eq. A47) used to determine threshold LAI for switching from maximum growth to normal growth leaf phenology mode. Also shown in the table are the turnover timescales for the stem (τ_S, yr, Eq. A52) and root (τ_R, yr, Eq. A52) components. Some parameter values differ between the model versions when PFTs' fractional coverage is specified and when they are dynamically modelled using competition between PFTs. If modified, the parameter values for the dynamically modelled version are shown in parentheses.

Plant functional type short name	$\Omega_{C,max}$	$\Omega_{D,max}$	T_{cold}^{leaf}	L_f	τ_S	τ_R
NDL-EVG	0.15 (0.10)	0.0025 (0.006)	−45.0 (−50.0)	0.4	86.0	13.8
NDL-DCD	0.30	0.005	−5.0	0.4	86.0	13.2
BDL-EVG	0.30 (1.5)	0.005 (0.010)	5.0	0.4	80.0	12.7
BDL-DCD-COLD	0.15 (0.40)	0.005 (0.025)	5.0 (8.0)	0.5	80.0	10.9
BDL-DCD-DRY	0.15	0.025 (0.030)	5.0	0.5	76.0	9.8
C3-CROP	0.15	0.005	5.0	0.5	20.0	3.0
C4-CROP	0.15	0.005	5.0	0.5	20.0	3.0
C3-GRASS	0.15	0.050 (0.020)	0.1	0.4	–	3.0
C4-GRASS	0.15	0.050 (0.020)	5.0	0.4	–	3.0

A6 Stem and root turnover

The turnover of stem and root components is modelled via their PFT-dependent specified lifetimes. The litter generation (kg C m^{-2} day^{-1}) associated with turnover of stem (D_S) and root (D_R) components is calculated based on the amount of biomass in the respective components (C_S, C_R; kg C m^{-2}) and their respective turnover timescales (τ_S and τ_R; yr; see Table A5) as

$$D_i = C_i \left[1 - \exp\left(-\frac{1}{365\,\tau_i} \right) \right], \quad i = \text{S, R.} \tag{A52}$$

A7 Rate change equations for carbon pools

With gross canopy photosynthesis rate (G_{canopy}, Eq. A15), maintenance and growth respirations (R_m and R_g, Eqs. A23 and A28), and heterotrophic respiration components ($R_{h,H}$ and $R_{h,D}$, Eq. A30) determined, it is possible to estimate the change in carbon amount of the model's five pools.

When the daily NPP ($G_{canopy} - R_m - R_g$) is positive, carbon is allocated to the plant's live carbon pools and the rate of change is given by

$$\frac{dC_i}{dt} = a_{fi} \left(G_{canopy} - R_m - R_g \right) - D_i - H_i - M_i,$$
$$i = \text{L, S, R} \tag{A53}$$

where a_{fi} is the corresponding allocation fractions for each pool (stem, root and leaves) and D_i is the litter produced from these components as explained in Sect. A5. H_i is the loss associated with fire that releases CO_2 and other trace gases to the atmosphere and M_i is the mortality associated with fire that contributes to the litter pool as explained in Sect. A9.

If the daily NPP is negative ($G_{canopy} < R_m$, $R_g = 0$), the rate of change is given by

$$\frac{dC_i}{dt} = a_{fi} G_{canopy} - R_{m,i} - D_i - H_i - M_i,$$
$$i = \text{L, S, R.} \tag{A54}$$

Negative NPP causes the plant to lose carbon from its live carbon pools due to respiratory costs in addition to the losses due to litter production (D_i) and disturbance (H_i, M_i).

The rate change equations for the litter and soil carbon pools are given by

$$\frac{dC_D}{dt} = D_L + D_S + D_R + M_L + M_R + M_S - H_D$$
$$- C_{D \to H} - R_{h,D}$$
$$\frac{dC_H}{dt} = C_{D \to H} - R_{h,H}, \tag{A55}$$

where $C_{D \to H}$ represents the transfer of humified litter to the soil carbon pool (Eq. A37) and H_D is loss associated with burning of litter associated with fire that releases CO_2 and other trace gases to the atmosphere.

A8 Conversion of biomass to structural vegetation attributes

The time-varying biomass in the leaves (C_L), stem (C_S) and root (C_R) components is used to calculate the structural attributes of vegetation for the energy and water balance calculations by CLASS.

Leaf biomass is converted to LAI using specific leaf area (SLA, m^2 (kg C)$^{-1}$), which itself is assumed to be a function of leaf lifespan (τ_L; Table A6)

$$SLA = \gamma_L \tau_L^{-0.5} \tag{A56}$$
$$LAI = C_L SLA$$

where γ_L is a constant with value equal to $25\,\mathrm{m^2\,(kg\,C)^{-1}\,yr^{0.5}}$.

The vegetation height (H; m) is calculated for tree, crop and grass PFTs as

$$H = \begin{cases} \min\left(10.0 C_S^{0.385}, 45\right) & \text{trees} \\ (C_S + C_L)^{0.385} & \text{crops} \\ 3.5(C_{L,g} + 0.55\, C_{L,b})^{0.5} & \text{grasses,} \end{cases} \tag{A57}$$

where $C_{L,g}$ is the green leaf biomass and $C_{L,b}$ is the brown leaf biomass that is scaled by 0.55 to reduce its contribution to the plant height. CTEM explicitly tracks brown leaf mass for grass PFTs. The turnover of green grass leaves, due to normal aging or stress from drought and/or cold, does not contribute to litter pool directly as the leaves first turn brown. The brown leaves themselves turnover to litter relatively rapidly ($\tau_{L,b} = 0.1\,\tau_L$).

CTEM dynamically simulates root distribution and depth in soil following Arora and Boer (2003). The root distribution takes an exponential form and roots grow and deepen with increasing root biomass. The cumulative root fraction at depth z is given by

$$f_R(z) = 1 - \exp(-\iota z). \tag{A58}$$

Rooting depth (d_R; m), which is defined to be the depth containing 99 % of the root mass, is found by setting z equal to d_R and $f_R = 0.99$, which yields

$$d_R = \frac{-\ln(1 - f_R)}{\iota} = \frac{-\ln(1 - 0.99)}{\iota} = \frac{4.605}{\iota}. \tag{A59}$$

The parameter ι that describes the exponential root distribution is calculated as

$$\iota = \bar{\iota} \left(\frac{\overline{C_R}}{C_R}\right)^{0.8}, \tag{A60}$$

where $\bar{\iota}$ represents the PFT-specific mean root distribution profile parameter and $\overline{C_R}$ the average root biomass derived from Jackson et al. (1996) (Table A6). Equation (A60) yields a lower (higher) value of ι than $\bar{\iota}$ when root biomass C_R is higher (lower) than the PFT-specific mean root biomass $\overline{C_R}$, resulting in a deeper (shallower) root profile than the mean root profile.

The rooting depth d_R is checked to ensure it does not exceed the soil depth. If so, d_R is set to the soil depth and ι is recalculated as $\iota = 4.605/d_R$ (see Eq. A59 for derivation of 4.605 term). The new value of ι is used to determine the root distribution profile adjusted to the shallower depth. Finally, the root distribution profile is used to calculate fraction of roots in each of the model's soil layers.

Table A6. PFT-specific leaf lifespans (τ_L, yr, Eq. A56) and mean root distribution profile ($\bar{\iota}$, dimensionless) and average root biomass ($\overline{C_R}$, $\mathrm{kg\,C\,m^{-2}}$) used to determine root distribution profile in Eq. (A60) for conversion of biomass to structural vegetation attributes in CTEM v. 2.0. Some parameter values differ between the model versions when PFTs' fractional coverage is specified and when they are dynamically modelled using competition between PFTs. If modified, the parameter values for the dynamically modelled version are shown in parentheses.

Plant functional type short name	τ_L	$\bar{\iota}$	$\overline{C_R}$
NDL-EVG	5.0	4.70	1.85
NDL-DCD	1.0	5.86	1.45
BDL-EVG	1.75 (1.5)	3.87	2.45
BDL-DCD-COLD	1.0	3.46	2.10
BDL-DCD-DRY	1.0	3.97	2.10
C3-CROP	1.75	3.97	0.10
C4-CROP	1.75	3.97	0.10
C3-GRASS	1.0	5.86	0.70
C4-GRASS	1.0	4.92	0.70

A9 Fire

CTEM v. 2.0 represents disturbance as both natural and human-influenced fires. The original fire parametrization corresponding to CTEM v. 1.0 is described in Arora and Boer (2005b). The parametrization has since been adapted and used in several other DGVMs (Kloster et al., 2010, 2012; Migliavacca et al., 2013; Li et al., 2012). CTEM v. 2.0 incorporates changes suggested in these studies as well as several new improvements.

Fire in CTEM is simulated using a process-based scheme of intermediate complexity that accounts for all elements of the fire triangle: fuel load, combustibility of fuel, and an ignition source. CTEM represents the probability of a fire occurrence (P_f), for a representative area of $500\,\mathrm{km^2}$ (a_{rep}), as

$$P_f = P_b P_i P_m, \tag{A61}$$

where the right hand side terms represent the fire probabilities that are conditioned on (i) the availability of biomass as a fuel source (P_b), (ii) the combustibility of the fuel based on its moisture content (P_m), and (iii) the presence of an ignition source (P_i). The probability of fire and the subsequent calculations are performed for each PFT present in a grid cell (but the PFT index α is omitted for clarity in Eq. A61). Since the CTEM parametrization is based on one fire per day per representative area, the representative area has to be sufficiently small that the requirement of only one fire per day is reasonable, yet sufficiently large such that it is not possible to burn the entire representative area in 1 day. Based on MODIS observed fire counts in Fig. 1 of Li et al. (2012), $500\,\mathrm{km^2}$ is an appropriate size to not have more than one fire per day and still be a large enough area to be assumed representative of the grid cell as a whole.

The P_b term depends on the aboveground biomass (B_{ag}) available for sustaining a fire (which includes the green and brown leaf mass, stem mass and litter mass, $B_{ag} = C_L + C_S + C_D$). Below a lower threshold of aboveground biomass (B_{low}; $0.2 \, kg \, C \, m^{-2}$; similar to Moorcroft et al., 2001, and Kucharik et al., 2000), fire is not sustained and thus has a probability of 0. Above a biomass of $1.0 \, kg \, C \, m^{-2}$ (B_{high}), P_b is set to 1 as the amount of fuel available is assumed sufficient for fire. P_b is then calculated using the aboveground biomass, B_{ag} ($kg \, C \, m^{-2}$) with a linear variation between the upper and lower thresholds as

$$P_b = \max\left[0, \min\left(1, \frac{B_{ag} - B_{low}}{B_{high} - B_{low}}\right)\right]. \tag{A62}$$

The linear decrease of P_b from B_{high} to B_{low} reflects the fragmentation of fuel that occurs as biomass decreases. Fuel fragmentation impacts upon area burned as it impedes the fire spread rate (Guyette et al., 2002).

The probability of fire based on the presence of ignition sources (P_i) is influenced by both natural (lightning) and anthropogenic agents (either intentional or accidental). An initial lightning scalar, ϑ_F, that varies between 0 and 1 is found as

$$\vartheta_F = \max\left[0, \min\left(1, \frac{F_{c2g} - F_{low}}{F_{high} - F_{low}}\right)\right], \tag{A63}$$

where F_{low} and F_{high} represent lower and upper thresholds of cloud-to-ground lightning strikes (F_{c2g}, flashes $km^{-2} \, month^{-1}$), respectively. Similar to Eq. (A62), below the lower threshold (F_{low}; $0.25 \, flashes \, km^{-2} \, month^{-1}$), ϑ_F is 0 implying lightning strikes are not sufficient to cause fire ignition, above the upper threshold (F_{high}; $10.0 \, flashes \, km^{-2} \, month^{-1}$) ϑ_F is 1, as ignition sources now do not pose a constraint on fire. The amount of cloud-to-ground lightning, F_{c2g}, is a fraction of the total lightning based on the relationship derived by Price and Rind (1993) (approximation of their Eqs. 1 and 2) as

$$F_{c2g} = 0.22 \exp(0.0059 \times |\Phi|) F_{tot}, \tag{A64}$$

where Φ is the grid cell latitude in degrees and F_{tot} is the total number of lightning flashes $km^{-2} \, month^{-1}$ (both cloud-to-cloud and cloud-to-ground). The probability of fire due to natural ignition, $P_{i,n}$, depends on the lightning scalar, ϑ_F, as

$$P_{i,n} = y(\vartheta_F) - y(0)(1 - \vartheta_F) + \vartheta_F[1 - y(1)]$$
$$y(\vartheta_F) = \frac{1}{1 + \exp\left(\frac{0.8 - \vartheta_F}{0.1}\right)}. \tag{A65}$$

Fire probability due to ignition caused by humans, $P_{i,h}$, is parametrized following Kloster et al. (2010) with a dependence on population density, p_d (number of people km^{-2})

$$P_{i,h} = \min\left[1, \left(\frac{p_d}{p_{thres}}\right)^{0.43}\right], \tag{A66}$$

where p_{thres} is a population threshold ($300 \, people \, km^{-2}$) above which $P_{i,h}$ is 1. The probability of fire conditioned on ignition, P_i, is then the total contribution from both natural and human ignition sources

$$P_i = \max[0, \min\{1, P_{i,n} + (1 - P_{i,n})P_{i,h}\}]. \tag{A67}$$

The population data used to calculate probability of fire ignition caused by humans and anthropogenic fire suppression (discussed further down in this section) is based on the HYDE 3.1 data set (Goldewijk et al., 2010)

The probability of fire due to the combustibility of the fuel, P_m, is dependent on the soil moisture in vegetation's root zone and in the litter layer. The root-zone soil wetness (ϕ_{root}, Eq. A39) is used as a surrogate for the vegetation moisture content and the soil wetness of the top soil layer as a surrogate for the litter moisture content. If a grid cell is covered by snow, P_m is set to zero. The probability of fire conditioned on soil wetness in vegetation's rooting zone, $P_{m,V}$, is then

$$P_{m,V} = 1 - \tanh\left[\left(\frac{1.75 \, \phi_{root}}{E_V}\right)^2\right], \tag{A68}$$

where E_V is the extinction soil wetness above which $P_{f,V}$ is reduced to near zero and is set to 0.30.

The probability of fire based on the moisture content in the *duff* layer, $P_{m,D}$, which includes the brown leaf mass (grasses only) and litter mass ($B_{duff} = C_{L,b} + C_D$; $kg \, C \, m^{-2}$), is calculated in a similar way but uses the soil wetness of the first soil layer, (ϕ_1, Eq. A7), as a surrogate for the moisture in the duff layer itself as

$$P_{m,D} = 1 - \tanh\left[\left(\frac{1.75\phi_1}{E_D}\right)^2\right], \tag{A69}$$

where the extinction soil wetness for the litter layer, E_D, is set to 0.50, which yields a higher probability of fire for the litter layer than for the vegetation for the same soil wetness. P_m is then the weighted average of $P_{m,V}$ and $P_{m,D}$ given by

$$P_m = P_{m,V}(1 - f_{duff}) + P_{m,D}f_{duff}$$
$$f_{duff} = \frac{B_{duff}}{B_{ag}} \tag{A70}$$

where f_{duff} is the duff fraction of aboveground combustible biomass.

The area burned (a) is assumed to be elliptical in shape for fires based upon the wind speed and properties of an ellipse

$$a(t) = \pi \frac{l}{2} \frac{w}{2} = \frac{\pi}{2}(v_d + v_u)v_p t^2, \tag{A71}$$

where l (m) and w (m) are the lengths of major and minor axes of the elliptical area burnt; v_d ($km \, h^{-1}$) and v_u ($km \, h^{-1}$) are the fire spread rates in the downwind and upwind directions, respectively; v_p ($km \, h^{-1}$) is the fire spread rate perpendicular to the wind direction and t is the time (h).

The fire spread rate in the downwind direction (v_d) is represented as

$$v_d = v_{d,\text{max}} \, g(u) \, h(\phi_{r,d}) \tag{A72}$$

where $v_{d,\text{max}}$ (km h^{-1}) is the PFT-specific maximum fire spread rate from Li et al. (2012), which is set to zero for crop PFTs (Table A7). The functions $g(u)$ accounts for the effect of wind speed and $h(\phi_{r,d})$ accounts for the effect of rooting zone and duff soil wetness on the fire spread rate, as discussed below.

The wind speed (u; km h^{-1}) is used to determine the length (l) to breadth (w) ratio, L_b, of the elliptical area burned by fire

$$L_b = \frac{l}{w} = \frac{v_d + v_u}{2v_p} = 1 + 10[1 - \exp(-0.06u)] \tag{A73}$$

and its head to back ratio, H_b, following Li et al. (2012), as

$$H_b = \frac{v_d}{v_u} = \frac{L_b + (L_b^2 - 1)^{0.5}}{L_b - (L_b^2 - 1)^{0.5}}, \tag{A74}$$

which help determine the fire spread rate in the direction perpendicular to the wind speed and in the downward direction. Equations (A73) and (A74) are combined to estimate the wind scalar $g(u)$ as

$$g(u) = g(0) \frac{2.0 L_b}{(1 + 1/H_b)}$$
$$\frac{g(u)}{g(0)} = \frac{v_d}{v_p} = \frac{2.0 L_b}{(1 + 1/H_b)}, \tag{A75}$$

which varies between 0.05 and 1. The lower limit is imposed by the $g(0)$ term, which has a value of 0.05 and represents the fire spread rate in the absence of wind ($u = 0$); the upper limit is assigned a maximum value of 1. The fire spread rate in the absence of wind is essentially the spread rate in the direction perpendicular to the wind speed (v_p). The value of the $g(0)$ term is derived by considering the case where the wind speed becomes very large. As $u \to \infty$ then $L_b \to 11$ and $H_b \to 482$, while $g(\infty) = 1$ due to its definition, which yields $g(0) = 0.0455 \approx 0.05$.

The dependence of fire spread rate on the rooting zone and duff soil wetness, $h(\phi_{r,d})$ is represented as

$$h(\phi_{r,d}) = h(\phi_{\text{root}})(1 - f_{\text{duff}}) + h(\phi_1) f_{\text{duff}}$$
$$h(\phi_{\text{root}}) = \left(1 - \min\left(1, \frac{\phi_{\text{root}}}{E_V}\right)\right)^2$$
$$h(\phi_1) = \left(1 - \min\left(1, \frac{\phi_1}{E_D}\right)\right)^2. \tag{A76}$$

Both $h(\phi_{\text{root}})$ and $h(\phi_1)$ gradually decrease from 1 (when soil wetness is 0 and soil moisture does not constrain fire spread rate) to 0 when soil wetness exceeds the respective extinction wetness thresholds, E_V and E_D.

With fire spread rate determined, and the geometry of the burned area defined, the area burned in 1 day, a_{1d} ($\text{km}^2 \, \text{day}^{-1}$), following Li et al. (2012), is calculated as

$$a_{1d} = \frac{\pi v_d^2 t^2}{4 L_b} \left(1 + \frac{1}{H_b}\right)^2$$
$$= \frac{\pi v_d^2 (24^2)}{4 L_b} \left(1 + \frac{1}{H_b}\right)^2 \tag{A77}$$

by setting t equal to 24 h.

The fire extinguishing probability, q, is used to calculate the duration (τ, days) of the fire, which in turn is used to calculated the area burned over the duration of the fire, $a_{\tau d}$. q is represented following Kloster et al. (2010) as

$$q = 0.5 + \frac{\max\left[0, 0.9 - \exp(-0.025 \, p_d)\right]}{2}, \tag{A78}$$

which yields a value of q that varies from 0.5 to 0.95 as population density, p_d (number of people km^{-2}), increases from zero to infinity. Higher population density thus implies a higher probability of fire being extinguished. q represents the probability that a fire will be extinguished on the same day it initiated and the probability that it will continue to the next day is $(1 - q)$. Assuming individual days are independent, the probability that the fire will still be burning on day τ is $(1 - q)^\tau$. The probability that a fire will last exactly τ days, $P(\tau)$, is the product of the probability that the fire still exists at day τ and the probability it will be extinguished on that day hence $P(\tau) = q(1 - q)^\tau$. This yields an exponential distribution of fire duration whose expected value is

$$\overline{\tau} = \mathbb{E}(\tau) = \sum_{\tau=0}^{\infty} \tau \, q(1 - q)^\tau = \frac{1 - q}{q}. \tag{A79}$$

Based on this fire duration and the area burned in 1 day (Eq. A77), the area burned over the duration of the fire ($a_{\tau d}$) (but still implemented in 1 day since the model does not track individual fires over their duration, $\text{km}^2 \, \text{day}^{-1}$) is calculated as

$$a_{\tau d} = \mathbb{E}(a_{1d} \tau^2) = \sum_{\tau=0}^{\infty} a_{1d} \tau^2 q(1 - q)^\tau \tag{A80}$$
$$= a_{1d} \frac{(1 - q)(2 - q)}{q^2}.$$

Finally, and reintroducing the PFT index α, the area burned is extrapolated for a PFT α ($A_{b,\alpha}$, $\text{km}^2 \, \text{day}^{-1}$) to the whole grid cell as

$$A_{b,\alpha} = P_{f,\alpha} \, a_{\tau d,\alpha} \frac{A_g f_\alpha}{a_{\text{rep}}}, \tag{A81}$$

where A_g is area of a grid cell (km^2), f_α the fractional coverage of PFT α and a_{rep} the representative area of $500 \, \text{km}^2$, as

Table A7. PFT-specific parameter values used in the fire module of CTEM v. 2.0. The maximum fire spread rate, $v_{d,max}$ ($km\,h^{-1}$), is used in Eq. (A72). The combustion factors for leaf ($\mho_L$), stem ($\mho_S$), root ($\mho_R$) and litter ($\mho_D$) components are used in Eq. (A83), and the mortality factors ($\Theta_L, \Theta_S, \Theta_R$) in Eq. (A85). A PFT's resistance to stand-replacing fire events is expressed via ζ_r and is used in Eq. (A87). $\mho$, Θ and ζ_r are all expressed as a fraction.

PFT short name	$v_{d,max}$	$\mho_L$	$\mho_S$	$\mho_R$	$\mho_D$	Θ_L	Θ_S	Θ_R	ζ_r
NDL-EVG	0.54	0.21	0.06	0.0	0.15	0.06	0.15	0.03	0.20
NDL-DCD	0.54	0.21	0.06	0.0	0.15	0.06	0.15	0.03	0.20
BDL-EVG	0.40	0.21	0.06	0.0	0.18	0.06	0.15	0.03	0.50
BDL-DCD-COLD	0.40	0.21	0.03	0.0	0.18	0.06	0.09	0.03	0.20
BDL-DCD-DRY	0.40	0.21	0.03	0.0	0.18	0.06	0.09	0.03	0.15
C3-GRASS	0.72	0.24*	–	0.0	0.21	0.03*	–	0.08	0.25
C4-GRASS	0.72	0.24*	–	0.0	0.21	0.03*	–	0.08	0.25

* For brown grass leaves $\mho_L$ is set to 0.24 and Θ_L to 0.02.

mentioned earlier. Area burned over the whole grid cell (A_b, $km^2\,day^{-1}$) is then calculated as the sum of area burned for individual PFTs,

$$A_b = \sum_{\alpha=1}^{N} A_{b,\alpha}. \tag{A82}$$

Fire emits CO_2, other trace gases, and aerosols as biomass is burned while plant mortality and damage due to fire contribute to the litter pool. The emissions of a trace gas/aerosol species j from PFT α, $E_{\alpha,j}$ (g species (m^{-2} grid cell area) day^{-1}) are obtained from a vector of carbon densities $\boldsymbol{C}_\alpha = (C_L, C_S, C_R, C_D)_\alpha$ ($kg\,C\,m^{-2}$) for its leaf, stem, root and litter components, multiplied by a vector of combustion factors $\boldsymbol{\mho}_\alpha = (\mho_L, \mho_S, \mho_R, \mho_D)_\alpha$, which determines what fraction of leaf, stem, root and litter components gets burned, multiplied by a vector of emissions factors $\boldsymbol{\Upsilon}_j = (\Upsilon_L, \Upsilon_S, \Upsilon_R, \Upsilon_D)_j$ (g species ($kg\,C$ dry organic matter)$^{-1}$), and by the area burned $A_{b,\alpha}$ for that PFT.

The dot product of $\boldsymbol{C}_\alpha$, $\boldsymbol{\Upsilon}_j$ and $\boldsymbol{\mho}_\alpha$ thus yields emissions per unit grid cell area of species j from PFT α,

$$E_{\alpha,j} = ((\boldsymbol{C}_\alpha \cdot \boldsymbol{\mho}_\alpha) \cdot \boldsymbol{\Upsilon}_j) \frac{A_{b,\alpha}}{A_g} \frac{1000}{450}, \tag{A83}$$

where the constant 1000 converts $\boldsymbol{C}_\alpha$ from $kg\,C\,m^{-2}$ to $g\,C\,m^{-2}$ and the constant 450 ($g\,C$ (kg dry organic matter)$^{-1}$) converts biomass from carbon units to dry organic matter (Li et al., 2012). The corresponding loss of carbon ($kg\,C\,m^{-2}\,day^{-1}$) from the three live vegetation components (L, S, R) and the litter pool (D) of PFT α is given by

$$H_{\alpha,i} = C_{\alpha,i}\mho_i \left(\frac{A_{b,\alpha}}{A_g}\right) \quad i = L, S, R, D. \tag{A84}$$

The PFT-specific combustion factors for leaf ($\mho_L$), stem ($\mho_S$), root ($\mho_R$) and litter ($\mho_D$) components are summarized in Table A7. Emission factors for all species of trace gases

and aerosols (CO_2, CO, CH_4, H_2, NHMC, NO_x, N_2O, total particulate matter, particulate matter less than $2.5\,\mu m$ in diameter, and black and organic carbon) are based on an updated set by Andreae and Merlet (2001) listed in Tables 3 and 4 of Li et al. (2012).

Litter generated by fire is based on similar mortality factors, which reflect a PFT's susceptibility to damage due to fire $\boldsymbol{\Theta}_\alpha = (\Theta_L, \Theta_S, \Theta_R)_\alpha$ (fraction). The contribution to litter pool of each PFT due to plant mortality associated with fire ($kg\,C\,m^{-2}\,day^{-1}$) is calculated as

$$M_\alpha = (\boldsymbol{C}_\alpha \cdot \boldsymbol{\Theta}_\alpha)\frac{A_{b,\alpha}}{A_g}, \tag{A85}$$

which is the sum of contribution from individual live vegetation pools

$$M_{\alpha,i} = C_{\alpha,i}\Theta_{\alpha,i}\left(\frac{A_{b,\alpha}}{A_g}\right) \quad i = L, S, R. \tag{A86}$$

The carbon loss terms associated with combustion of vegetation components and litter ($H_{\alpha,i}$, $i = L, S, R, D$) and mortality of vegetation components ($M_{\alpha,i}$, $i = L, S, R$) due to fire are used in Eqs. (A53) and (A55), which describe the rate of change of carbon in model's five pools (however, listed there without the PFT subscript α). The PFT-specific mortality factors for leaf (Θ_L), stem (Θ_S) and root (Θ_R) components are listed in Table A7.

When CTEM is run with prescribed PFT fractional cover, the area of PFTs does not change and the fire-related emissions of CO_2, other trace gases and aerosols, and generation of litter act to thin the remaining biomass. When competition between PFTs for space is allowed, fire both thins the remaining biomass and through plant mortality creates bare ground, which is subsequently available for colonization. The creation of bare ground depends on the susceptibility of each PFT to stand replacing fire (ζ_r, fraction) (Table A7) and the PFT area burned. The fire-related mortality rate, m_{dist} (day^{-1}), used in Eq. (19), is then

$$m_{\mathrm{dist},\alpha} = \zeta_{\mathrm{r},\alpha} \frac{A_{\mathrm{b},\alpha}}{f_\alpha A_{\mathrm{g}}}. \tag{A87}$$

After bare ground generation associated with fire, the thinned biomass is spread uniformly over the remaining fraction of a PFT. However, it is ensured that the carbon density of the remaining biomass does not increase to a value above what it was before the fire occurred.

A10 Treatment of agricultural crop PFTs

CTEM explicitly simulates C_3 and C_4 crops using a simple approach. Crops increase their biomass depending on environmental conditions similar to other PFTs; however, unlike other PFTs they are explicitly harvested. Harvesting is initiated when the daily mean air temperature remains below $8\,^\circ\mathrm{C}$ for 5 consecutive days, or when the crop LAI reaches a threshold ($3.5\,\mathrm{m^2\,m^{-2}}$ for C_3 crops and $4.5\,\mathrm{m^2\,m^{-2}}$ for C_4 crops) signifying that the crops have matured (Arora and Boer, 2005a). Harvesting occurs over a period of 15 days and the harvested biomass contributes to the litter pool. The above environmental criteria typically lead to one annual crop cycle in mid- to high-latitude regions while multiple crop cycles may be realized in tropical regions depending on the precipitation seasonality. Harvesting ensures that vegetation biomass does not keep increasing on croplands and, unlike natural vegetation, this prevents croplands from sequestering carbon as atmospheric CO_2 increases. Some carbon may still be sequestered in the soil carbon pool but crops have a higher soil carbon decomposition rate and lower humification factor compared to other PFTs (ς_{H} and χ, respectively, in Table A3) to account for tillage, which prevents soil carbon sequestration in croplands.

A11 Land use change

The land use change (LUC) module of CTEM is based on Arora and Boer (2010) and briefly described here. When the area of crop PFTs changes, CTEM generates LUC emissions. In the simulation where fractional coverage of PFTs is specified, the changes in fractional coverage of crop PFTs are made consistent with changes in the fractional coverage of natural non-crop PFTs. That is, an increase or decrease in the area of crop PFTs is associated with a corresponding decrease or increase in the area of non-crop PFTs. This approach is taken by Wang et al. (2006), which allows one to reconstruct historical land cover given a spatial data set of changes in crop area over the historical period and an estimate of potential natural land cover for the pre-industrial period (as described in Sect. 3). When competition between

PFTs for space is allowed, only the fractional coverage of crop PFTs is specified. Similar to a simulation with prescribed PFT fractions, when the area of crop PFTs increases, the fractional coverage of non-crop PFTs is decreased in proportion to their existing coverage (i.e. the linear approach of Wang et al., 2006). Alternatively, and in contrast to the simulation with prescribed PFT fractions, when the area of crop PFTs decreases then the generated bare fraction is available for recolonization by non-crop PFTs.

A decrease in the area of natural non-crop PFTs, associated with an increase in area of crop PFTs, results in deforested biomass (while the term *deforested* implies clearing of forests, the same processes can occur in grasslands as well and is meant here to imply removal of the biomass). The deforested biomass is divided into three components: (i) the component that is combusted or used for fuel wood immediately after natural vegetated is deforested and which contributes to atmospheric CO_2, (ii) the component left as slash or used for pulp and paper products and (iii) the component that is used for long-lasting wood products. The fractions allocated to these three components depend on whether the PFTs are woody or herbaceous and on their aboveground vegetation biomass density (see Table 1 of Arora and Boer, 2010). To account for the timescales involved, the fraction allocated to slash or pulp and paper products is transferred to the model's litter pool and the fraction allocated to long-lasting wood products is allocated to the model's soil carbon pool. Land use change associated with a decrease in the area of natural vegetation thus redistributes carbon from living vegetation to dead litter and soil carbon pools and emits CO_2 to the atmosphere through direct burning of the deforested biomass. The net result is positive LUC carbon emissions from land to the atmosphere.

When croplands are abandoned, the area of natural PFTs increases. In simulations with prescribed fractional coverage of PFTs this results in a decreased carbon density for all model pools as the same amount of carbon is spread over a larger fraction of the grid cell. This reduced density implies that natural vegetation is able to take up carbon as it comes into equilibrium with the driving climate and atmospheric CO_2 concentration. This creates the carbon sink associated with abandonment of croplands as natural vegetation grows in its place. In simulations with competition between PFTs, the abandoned land is treated as bare ground, which is subsequently available for recolonization, as mentioned above. As natural vegetation expands into bare ground it takes up carbon, again creating the carbon sink associated with abandonment of croplands. The net result is negative LUC carbon emissions as carbon is taken from atmosphere to grow vegetation over the area that was previously a cropland.

Code availability

Fortan code for CLASS–CTEM modelling framework is available on request and upon agreeing to Environment Canada's licensing agreement available at http://collaboration.cmc.ec.gc.ca/science/rpn.comm/license.html. Please contact the first author (joe.melton@canada.ca) to easily and rapidly obtain the model code.

Copyright statement

The works published in this journal are distributed under the Creative Commons Attribution 3.0 License. This license does not affect the Crown copyright work, which is re-usable under the Open Government Licence (OGL). The Creative Commons Attribution 3.0 License and the OGL are interoperable and do not conflict with, reduce or limit each other.

© Crown copyright 2015.

The Supplement related to this article is available online at doi:10.5194/gmd-9-323-2016-supplement.

References

Ajtay, M. J., Ketner, P., and Duvigneaud, P.: Terrestrial primary production and phytomass, in: The Global Carbon Cycle, SCOPE 13, edited by: Bolin, B., Degens, E. T., and Ketner, P., John Wiley & Sons, New York, 129–182, 1979.

Andreae, M. O. and Merlet, P.: Emission of trace gases and aerosols from biomass burning, Global Biogeochem. Cy., 15, 955–966, doi:10.1029/2000GB001382, 2001.

Arora, V.: Land surface modelling in general circulation models: a hydrological perspective, PhD thesis, Department of Civil and Environmental Engineering, University of Melbourne, 1997.

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 representation of variable root distribution in dynamic vegetation models, Earth Interact., 7, 1–19, doi:10.1175/1087-3562(2003)007<0001:AROVRD>2.0.CO;2, 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, 2005a.

Arora, V. K. and Boer, G. J.: Fire as an interactive component of dynamic vegetation models, J. Geophys. Res.-Biogeo., 110, G02008, doi:10.1029/2005JG000042, 2005b.

Arora, V. K. and Boer, G.: Achieving coexistence: Comment on "Modelling rainforest diversity: The role of competition" by Bampfylde et al.(2005), Ecol. Model., 192, 322–324, doi:10.1016/j.ecolmodel.2005.08.001, 2006a.

Arora, V. K. and Boer, G. J.: Simulating competition and coexistence between plant functional types in a dynamic vegetation model, Earth Interact., 10, 1–30, doi:10.1175/EI170.1, 2006b.

Arora, V. K. and Boer, G. J.: Uncertainties in the 20th century carbon budget associated with land use change, Glob. Change Biol., 16, 3327–3348, doi:10.1111/j.1365-2486.2010.02202.x, 2010.

Arora, V. K. and Boer, G. J.: Terrestrial ecosystems response to future changes in climate and atmospheric CO_2 concentration, Biogeosciences, 11, 4157–4171, doi:10.5194/bg-11-4157-2014, 2014.

Arora, V. K., Boer, G. J., Christian, J. R., Curry, C. L., Denman, K. L., Zahariev, K., Flato, G. M., Scinocca, J. F., Merryfield, W. J., and Lee, W. G.: The effect of terrestrial photosynthesis down regulation on the twentieth-century carbon budget simulated with the CCCma Earth System Model, J. Climate, 22, 6066–6088, doi:10.1175/2009JCLI3037.1, 2009.

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.

Arora, V. K., Boer, G. J., Friedlingstein, P., Eby, M., Jones, C. D., Christian, J. R., Bonan, G., Bopp, L., Brovkin, V., Cadule, P., Hajima, T., Ilyina, T., Lindsay, K., Tjiputra, J. F., and Wu, T.: Carbon–concentration and carbon–climate feedbacks in CMIP5 earth system models, J. Climate, 26, 5289–5314, doi:10.1175/JCLI-D-12-00494.1, 2013.

Ball, T. J., Woodrow, I. E., and Berry, J. A.: A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions, in: Progress in Photosynthesis Research, Springer Netherlands, 221–224, doi:10.1007/978-94-017-0519-6_48, 1987.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I., and Papale, D.: Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate, Science, 329, 834–838, doi:10.1126/science.1184984, 2010.

Box, E. O.: Plant functional types and climate at the global scale, J. Veg. Sci., 7, 309–320, doi:10.2307/3236274, 1996.

Brentnall, S., Beerling, D., Osborne, C., Harland, M., Francis, J. E., Valdes, P. J., and Wittig, V. E.: Climatic and ecological determinants of leaf lifespan in polar forests of the high CO_2 Cretaceous 'greenhouse'world, Glob. Change Biol., 11, 2177–2195, doi:10.1111/j.1365-2486.2005.001068.x, 2005.

Brovkin, V., Raddatz, T., Reick, C. H., Claussen, M., and Gayler, V.: Global biogeophysical interactions between forest and climate,

Geophys. Res. Lett., 36, L07405, doi:10.1029/2009GL037543, 2009.

Bugmann, H.: A review of forest gap models, Climatic Change, 51, 259–305, doi:10.1023/A:1012525626267, 2001.

Collatz, G., Ball, J., Grivet, C., and Berry, J. A.: Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer, Agr. Forest Meteorol., 54, 107–136, doi:10.1016/0168-1923(91)90002-8, 1991.

Collatz, G., Ribas-Carbo, M., and Berry, J.: Coupled photosynthesis-stomatal conductance model for leaves of C_4 plants, Funct. Plant Biol., 19, 519–538, doi:10.1071/pp9920519, 1992.

Cox, P.: Description of the TRIFFID dynamic global vegetation model, Tech. Rep. 24, Hadley Centre, 2001.

Cox, P. M., Betts, R. A., Bunton, C. B., Essery, R. L. H., Rowntree, P. R., and Smith, J.: The impact of new land surface physics on the GCM simulation of climate and climate sensitivity, Clim. Dynam., 15, 183–203, doi:10.1007/s003820050276, 1999.

Cramer, W., Bondeau, A., Woodward, F. I., 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.

FAO/IIASA/ISRIC/ISS-CAS/JRC: Harmonized World Soil Database (version 1.2), available at: http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/ (last access: 20 January 2016), 2012.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species, Planta, 149, 78–90, doi:10.1007/BF00386231, 1980.

Field, C. H. and Mooney, H. A.: Photosynthesis–nitrogen relationship in wild plants, in: On the Economy of Plant Form and Function, edited by: Givnish, T. J., 25–55, available at: http://agris.fao.org (last access: 3 September 2014), 1986.

Fisher, R. A., Muszala, S., Verteinstein, M., Lawrence, P., Xu, C., McDowell, N. G., Knox, R. G., Koven, C., Holm, J., Rogers, B. M., Spessa, A., Lawrence, D., and Bonan, G.: Taking off the training wheels: the properties of a dynamic vegetation model without climate envelopes, CLM4.5(ED), Geosci. Model Dev., 8, 3593–3619, doi:10.5194/gmd-8-3593-2015, 2015.

Friedl, M., Strahler, A., Schaaf, C., Hodges, J. C. F., and Salomon, J.: Binary MODIS MOD12C1 0.25 Degree Land Cover Climate Modeler Grid, available at: http://webmap.ornl.gov/wcsdown/wcsdown.jsp?dg_id=10011_1 (last access: 20 January 2016), Department of Geography, Boston University, Boston, Massachusetts, USA, 2013.

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 intercom-

parison, J. Climate, 19, 3337–3353, doi:10.1175/JCLI3800.1, 2006.

Gerber, S., Joos, F., and Prentice, I. C.: Sensitivity of a dynamic global vegetation model to climate and atmospheric CO_2, Glob. Change Biol., 10, 1223–1239, doi:10.1111/j.1529-8817.2003.00807.x, 2004.

Giglio, L., van der Werf, G. R., Randerson, J. T., Collatz, G. J., and Kasibhatla, P.: Global estimation of burned area using MODIS active fire observations, Atmos. Chem. Phys., 6, 957–974, doi:10.5194/acp-6-957-2006, 2006.

Giglio, L., Randerson, J. T., van der Werf, G. R., Kasibhatla, P. S., Collatz, G. J., Morton, D. C., and DeFries, R. S.: Assessing variability and long-term trends in burned area by merging multiple satellite fire products, Biogeosciences, 7, 1171–1186, doi:10.5194/bg-7-1171-2010, 2010.

Gobron, N., Belward, A., Pinty, B., and Knorr, W.: Monitoring biosphere vegetation 1998–2009, Geophys. Res. Lett., 37, L15402, doi:10.1029/2010GL043870, 2010.

Goldewijk, K. K., Beusen, A., and Janssen, P.: Long-term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1, Holocene, 20, 1–9, doi:10.1177/0959683609356587, 2010.

Guyette, R. P., Muzika, R. M., and Dey, D. C.: Dynamics of an anthropogenic fire regime, Ecosystems, 5, 472–486, doi:10.1007/s10021-002-0115-7, 2002.

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., Fisk, J. P., Hibbard, K., Houghton, R. A., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T., Goldewijk, K. K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., van Vuuren, D. P., and Wang, Y. P.: Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, Climatic Change, 109, 117–161, doi:10.1007/s10584-011-0153-2, 2011.

Ingestad, T. and Lund, A.-B.: Theory and techniques for steady state mineral nutrition and growth of plants, Scand. J. Forest Res., 1, 439–453, doi:10.1080/02827588609382436, 1986.

Jackson, R. B., Canadell, J., Ehleringer, J. R., Mooney, H. A., Sala, O. E., and Schulze, E. D.: A global analysis of root distributions for terrestrial biomes, Oecologia, 108, 389–411, doi:10.1007/BF00333714, 1996.

Jobbágy, 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, doi:10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2, 2000.

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.

Keller, M., Palace, M., and Hurtt, G.: Biomass estimation in the Tapajos National Forest, Brazil: examination of sampling and allometric uncertainties – new directions in tropical forest re-

search, Forest Ecol. Manag., 154, 371–382, doi:10.1016/S0378-1127(01)00509-6, 2001.

Kim, Y. and Wang, G.: Impact of vegetation feedback on the response of precipitation to antecedent soil moisture anomalies over North America, J. Hydrometeorol., 8, 534–550, doi:10.1175/JHM612.1, 2007.

Kirschbaum, M. U.: The temperature dependence of soil organic matter decomposition, and the effect of global warming on soil organic C storage, Soil Biol. Biochem., 27, 753–760, doi:10.1016/0038-0717(94)00242-S, 1995.

Kloster, S., Mahowald, N. M., Randerson, J. T., Thornton, P. E., Hoffman, F. M., Levis, S., Lawrence, P. J., Feddema, J. J., Oleson, K. W., and Lawrence, D. M.: Fire dynamics during the 20th century simulated by the Community Land Model, Biogeosciences, 7, 1877–1902, doi:10.5194/bg-7-1877-2010, 2010.

Kloster, S., Mahowald, N. M., Randerson, J. T., and Lawrence, P. J.: The impacts of climate, land use, and demography on fires during the 21st century simulated by CLM-CN, Biogeosciences, 9, 509–525, doi:10.5194/bg-9-509-2012, 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, Global Biogeochem. Cy., 19, GB1015 doi:10.1029/2003GB002199, 2005.

Kucharik, C. J., Foley, J. A., Delire, C., Fisher, V. A., Coe, M. T., Lenters, J. D., Young-Molling, C., Ramankutty, N., Norman, J. M., and Gower, S. T.: Testing the performance of a dynamic global ecosystem model: water balance, carbon balance, and vegetation structure, Global Biogeochem. Cy., 14, 795–825, doi:10.1029/1999GB001138, 2000.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., and Slater, A. G.: Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model, J. Adv. Model. Earth Syst., 3, M03001, doi:10.1029/2011MS000045, 2011.

Leuning, R.: A critical appraisal of a combined stomatal-photosynthesis model for C_3 plants, Plant Cell Environ., 18, 339–355, 1995.

Levis, S., Bonan, G. B., Vertenstein, M., and Oleson, K. W.: Technical Documentation and User's Guide to the Community Land Model's Dynamic Global Vegetation Model, Tech. Rep. TN-459+IA, NCAR, 2004.

Li, F., Zeng, X. D., and Levis, S.: A process-based fire parameterization of intermediate complexity in a Dynamic Global Vegetation Model, Biogeosciences, 9, 2761–2780, doi:10.5194/bg-9-2761-2012, 2012.

Li, R. and Arora, V. K.: Effect of mosaic representation of vegetation in land surface schemes on simulated energy and carbon balances, Biogeosciences, 9, 593–605, doi:10.5194/bg-9-593-2012, 2012.

Litton, C. M., Raich, J. W., and Ryan, M. G.: Carbon allocation in forest ecosystems, Glob. Change Biol., 13, 2089–2109, 2007.

Lloyd, J. and Taylor, J. A.: On the temperature dependence of soil respiration, Funct. Ecol., 8, 315–323, doi:10.2307/2389824, 1994.

Lotka, A. J.: Elements of Physical Biology, Williams and Wilkins, Baltimore, 1925.

Lüdeke, M., Badeck, F., Otto, R., Häager, C., Dönges, S., Kindermann, J., Würth, G., Lang, T., Jäkel, U., Klaudius, A., Ramge, P., Habermehl, S., and Kohlmaier, G. H.: The Frankfurt Biosphere Model: a global process-oriented model of seasonal and long-term CO_z exchange between terrestrial ecosystems and the atmosphere, Clim. Res., 4, 143–166, 1994.

Luyssaert, S. Inglima, I., Jung, M., Richardson, A. D., Reichstein, M., Papale, D., Piao, S. L., Schulze, E.-D., Wingate, L., Matteucci, G., Aragao, L., Aubinet, M., Beer, C., Bernhofer, C., Black, K., G., Bonal, D., Bonnefond, J.-M., Chambers, J., Ciais, P., Cook, B., Davis, K. J., Dolman, A. J., Gielen, B., Goulden, M., Grace, J., Granier, A., Grelle, A.,Griffis, T., Grünwald, T., Guidolotti, G., Hanson, P. J., Harding, R., Hollinger, D. Y., Hutyra, L. R., Kolari, P., Kruijt, B., Kutsch, W., Lagergren, F., Laurila, T., Law, B. E., Le Maire, G., Lindroth, A., Loustau, D., Malhi, Y., Mateus, J., Migliavacca, M., Misson, L., Montagnani, L., Moncrieff, J., Moors, E., Munger, J. W., Nikinmaa, E., Ollinger, S. V., Pita, G., Rebmann, C., Roupsard, O., Saigusa, N., Sanz, M. J., Seufert, G., Sierra, C., Smith, M.-L., Tang, J., Valentini, R., Vesala, T., and Janssens, I. A.: CO_2 balance of boreal, temperate, and tropical forests derived from a global database, Glob. Change Biol., 13, 2509–2537, doi:10.1111/j.1365-2486.2007.01439.x,2007.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and van Vuuren, D. P.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Climatic Change, 109, 213–241, doi:10.1007/s10584-011-0156-z, 2011.

Melton, J. R. and Arora, V. K.: Sub-grid scale representation of vegetation in global land surface schemes: implications for estimation of the terrestrial carbon sink, Biogeosciences, 11, 1021–1036, doi:10.5194/bg-11-1021-2014, 2014.

Melton, J. R., Shrestha, R. K., and Arora, V. K.: The influence of soils on heterotrophic respiration exerts a strong control on net ecosystem productivity in seasonally dry Amazonian forests, Biogeosciences, 12, 1151–1168, doi:10.5194/bg-12-1151-2015, 2015.

Mieville, A., Granier, C., Liousse, C., Guillaume, B., Mouillot, F., Lamarque, J.-F., Grégoire, J.-M., and Pétron, G.: Emissions of gases and particles from biomass burning during the 20th century using satellite data and an historical reconstruction, Atmos. Environ., 44, 1469–1477, doi:10.1016/j.atmosenv.2010.01.011, 2010.

Migliavacca, M., Dosio, A., Kloster, S., Ward, D. S., Camia, A., Houborg, R., Houston Durrant, T., Khabarov, N., Krasovskii, A. A., San Miguel-Ayanz, J., and Cescatti, A.: Modeling burned area in Europe with the Community Land Model, J. Geophys. Res.-Biogeo., 118, 265–279, doi:10.1002/jgrg.20026, 2013.

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, doi:10.1890/0012-9615(2001)071[0557:AMFSVD]2.0.CO;2, 2001.

Mouillot, F. and Field, C.: Fire history and the global carbon budget: A 1×1 fire history reconstruction for the 20th century, Glob. Change Biol., 11, 398–420, 2005.

Mouillot, F., Narasimha, A., Balkanski, Y., Lamarque, J.-F., and Field, C. B.: Global carbon emissions from biomass burn-

ing in the 20th century, Geophys. Res. Lett., 33, L01801, doi:10.1029/2005GL024707, 2006.

Mu, M., Randerson, J. T., van der Werf, G. R., Giglio, L., Kasibhatla, P., Morton, D., Collatz, G. J., DeFries, R. S., Hyer, E. J., Prins, E. M., Griffith, D. W. T., Wunch, D., Toon, G. C., Sherlock, V., and Wennberg, P. O.: Daily and 3-hourly variability in global fire emissions and consequences for atmospheric model predictions of carbon monoxide, J. Geophys. Res.-Atmos., 116, D24303, doi:10.1029/2011JD016245, 2011.

Pacala, S. W. and Tilman, D.: Limiting similarity in mechanistic and spatial models of plant competition in heterogeneous environments, Am. Nat., 143, 222–257, 1994.

Parton, W., Haxeltine, A., Thornton, P., Anne, R., and Hartman, M.: Ecosystem sensitivity to land-surface models and leaf area index, Global Planet. Change, 13, 89–98, doi:10.1016/0921-8181(95)00040-2, 1996.

Peng, Y., Arora, V. K., Kurz, W. A., Hember, R. A., Hawkins, B. J., Fyfe, J. C., and Werner, A. T.: Climate and atmospheric drivers of historical terrestrial carbon uptake in the province of British Columbia, Canada, Biogeosciences, 11, 635–649, doi:10.5194/bg-11-635-2014, 2014.

Pons, T. L. and Welschen, R. A. M.: Midday depression of net photosynthesis in the tropical rainforest tree *Eperua grandiflora*: contributions of stomatal and internal conductances, respiration and Rubisco functioning, Tree Physiol., 23, 937–947, 2003.

Post, W. M., Emanuel, W. R., Zinke, P. J., and Stangenberger, A. G.: Soil carbon pools and world life zones, Nature, 298, 156–159, doi:10.1038/298156a0, 1982.

Prentice, C. I., Sykes, M. T., and Cramer, W.: A simulation model for the transient effects of climate change on forest landscapes, Ecol. Model., 65, 51–70, doi:10.1016/0304-3800(93)90126-D, 1993.

Price, C. and Rind, D.: What determines the cloud-to-ground lightning fraction in thunderstorms?, Geophys. Res. Lett., 20, 463–466, doi:10.1029/93GL00226, 1993.

Ramankutty, N. and Foley, J. A.: Estimating historical changes in global land cover: Croplands from 1700 to 1992, Global Biogeochem. Cy., 13, 997–1027, 1999.

Reich, P. B., Walters, M. B., Tjoelker, M. G., Vanderklein, D., and Buschena, C.: Photosynthesis and respiration rates depend on leaf and root morphology and nitrogen concentration in nine boreal tree species differing in relative growth rate, Funct. Ecol., 12, 395–405, doi:10.1046/j.1365-2435.1998.00209.x, 1998.

Reich, P. B., Roy, R. L., Lu, X., and Oleksyn, J.: Biogeographic variation in evergreen conifer needle longevity and impacts on boreal forest carbon cycle projections, P. Natl. Acad. Sci. USA, 111, 13703–13708, doi:10.1073/pnas.1216054110, 2014.

Ruesch, A. and Holly, K.: New IPCC Tier-1 Global Biomass Carbon Map For the Year 2000, available at: ftp://cdiac.ornl.gov/pub/global_carbon/ (last access: 5 July 2013), Carbon Dioxide Information Analysis Center http://cdiac.ornl.gov (last access: 5 July 2013), Oak Ridge National Laboratory, Oak Ridge, Tennessee, 2008.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., and Hashimoto, H.: A continuous satellite-derived measure of global terrestrial primary production, Bioscience, 54, 547–560, doi:10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2, 2004.

Ryan, M. G.: Effects of climate change on plant respiration, Ecol. Appl., 1, 157–167, doi:10.2307/1941808, 1991.

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., and Morel, A.: Benchmark map of forest carbon stocks in tropical regions across three continents, P. Natl. Acad. Sci. USA, 108, 9899–9904, doi:10.1073/pnas.1019576108, 2011.

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.

Saugier, B., Roy, J., and Mooney, H. A.: Estimations of global terrestrial productivity: converging towards a single number?, in: Terrestrial Global Productivity, edited by: Roy, J., Saugier, B., and Mooney, H. A., Physiological Ecology, Academic Press, San Diego, California, 543–558, 2001.

Schlesinger, W. H.: Carbon balance in terrestrial detritus, Annu. Rev. Ecol. Syst., 8, 51–81, doi:10.1146/annurev.es.08.110177.000411, 1977.

Sellers, P. J., Tucker, C. J., Collatz, G. J., Los, S. O., Justice, C. O., Dazlich, D. A., and Randall, D. A.: A revised land surface parameterization (SiB2) for atmospheric GCMS. Part II: The generation of global fields of terrestrial biophysical parameters from satellite data, J. Climate, 9, 706–737, doi:10.1175/1520-0442(1996)009<0706:ARLSPF>2.0.CO;2, 1996.

Siemann, E. and Rogers, W. E.: Changes in light and nitrogen availability under pioneer trees may indirectly facilitate tree invasions of grasslands, J. Ecol., 91, 923–931, doi:10.1046/j.1365-2745.2003.00822.x, 2003.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, Glob. Change Biol., 9, 161–185, doi:10.1046/j.1365-2486.2003.00569.x, 2003.

Smith, B., Prentice, I. C., and Sykes, M. T.: Representation of vegetation dynamics in the modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space, Global Ecol. Biogeogr., 10, 621–637, doi:10.1046/j.1466-822X.2001.t01-1-00256.x, 2001.

Still, C. J., Berry, J. A., Collatz, G. J., and DeFries, R. S.: Global distribution of C_3 and C_4 vegetation: carbon cycle implications, Global Biogeochem. Cy., 17, 6-1–6-14, doi:10.1029/2001GB001807, 2003.

Tian, H., Melillo, J. M., Kicklighter, D. W., McGuire, A. D., and Helfrich, J.: The sensitivity of terrestrial carbon storage to historical climate variability and atmospheric CO_2 in the United States, Tellus B, 51, 414–452, doi:10.1034/j.1600-0889.1999.00021.x, 1999.

Tilman, D.: Resource competition and community structure, Monogr. Popul. Biol., 17, 1–296, 1982.

Tjoelker, M., Oleksyn, J., and Reich, P.: Modelling respiration of vegetation: evidence for a general temperature-dependent Q_{10}, Glob. Change Biol., 7, 223–230, 2001.

Tongway, D. J. and Ludwig, J. A.: Small-scale resource heterogeneity in semi-arid landscapes, Pac. Conserv. Biol., 1, 201–208, 1994.

Tongway, D. J. and Ludwig, J. A.: Theories on the origins, maintenance, dynamics, and functioning of banded landscapes, in: Banded Vegetation Patterning in Arid and Semiarid Environments, Ecological Studies, Springer, New York, 20–31, doi:10.1007/978-1-4613-0207-0_2, 2001.

Trenberth, K. E., Fasullo, J. T., and Mackaro, J.: Atmospheric moisture transports from ocean to land and global energy flows in reanalyses, J. Climate, 24, 4907–4924, doi:10.1175/2011JCLI4171.1, 2011.

Verseghy, D.: CLASS – The Canadian land surface scheme, Manual/Documentation version 3.6, Climate Research Division, Science and Technology Branch, Environment Canada, 2012.

Viovy, N.: CRU-NCEP Version 4, title of the publication associated with this dataset: CRU-NCEP version 4, available at: http://dods.extra.cea.fr/data/p529viov/cruncep/ (last access: 4 June 2013), 2012.

Volterra, V.: Fluctuations in the abundance of a species considered mathematically, Nature, 118, 558–560, doi:10.1038/118558a0, 1926.

Wang, A., Price, D. T., and Arora, V.: Estimating changes in global vegetation cover (1850–2100) for use in climate models, Global Biogeochem. Cy., 20, GB3028, doi:10.1029/2005GB002514, 2006.

Wang, G., Sun, S., and Mei, R.: Vegetation dynamics contributes to the multi-decadal variability of precipitation in the Amazon region, Geophys. Res. Lett., 38, L19703, doi:10.1029/2011GL049017, 2011.

Waring, R. H.: Estimating forest growth and efficiency in relation to canopy leaf area, Adv. Ecol. Res., 13, 327–354, doi:10.1016/s0065-2504(08)60111-7, 1983.

Xu, L. and Baldocchi, D. D.: Seasonal trends in photosynthetic parameters and stomatal conductance of blue oak (*Quercus douglasii*) under prolonged summer drought and high temperature, Tree Physiol., 23, 865–877, 2003.

Zeng, X., Zeng, X., and Barlage, M.: Growing temperate shrubs over arid and semiarid regions in the Community Land Model-Dynamic Global Vegetation Model, Global Biogeochem. Cy., 22, GB3003, doi:10.1029/2007GB003014, 2008.

Zhao, M., Running, S. W., and Nemani, R. R.: Sensitivity of Moderate Resolution Imaging Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reanalyses, J. Geophys. Res.-Biogeo., 111, G01002, doi:10.1029/2004JG000004, 2006.

Zhu, D., Peng, S. S., Ciais, P., Viovy, N., Druel, A., Kageyama, M., Krinner, G., Peylin, P., Ottlé, C., Piao, S. L., Poulter, B., Schepaschenko, D., and Shvidenko, A.: Improving the dynamics of Northern Hemisphere high-latitude vegetation in the OR-CHIDEE ecosystem model, Geosci. Model Dev., 8, 2263–2283, doi:10.5194/gmd-8-2263-2015, 2015.

Zobler, L.: A world soil file for global climate modelling, NASA Technical Memorandum 87802, 1986.