Robust Ecosystem Demography (RED): a parsimonious approach to modelling vegetation dynamics in Earth System Models

Arthur P. K. Argles[1], Jonathan R. Moore[1], Chris Huntingford[2], Andrew J. Wiltshire[3], Chris D. Jones[3], and Peter M. Cox[1]

[1]College of Engineering, Mathematics, and Physical Sciences, University of Exeter, Exeter EX4 4QF, UK
[2]Centre for Ecology and Hydrology, Wallingford OX10 8BB, UK
[3]Met Office Hadley Centre, Fitzroy Road, Exeter EX1 3PB, UK

Correspondence: aa760@exeter.ac.uk, P.M.Cox@exeter.ac.uk, J.Moore3@exeter.ac.uk

Abstract. A significant proportion of the uncertainty in climate projections arises from uncertainty in the representation of land carbon uptake. Dynamic Global Vegetation Models (DGVMs) vary in their representations of regrowth and competition for resources, which results in differing responses to changes in atmospheric CO_2 and climate. More advanced cohort-based patch models are now becoming established in the latest DGVMs. These typically attempt to simulate the size-distribution of trees

5 as a function of both tree-size (mass or trunk diameter) and age (time since disturbance). This approach can capture the overall impact of stochastic disturbance events on the forest structure and biomass, but at the cost of needing to update a probability density function in two-dimensions. Here we present the *Robust Ecosystem Demography (RED)*, in which the pdf is collapsed on to the single dimension of tree mass. RED is designed to retain the ability of more complex cohort DGVMs to represent forest demography, while also being parameter sparse and analytically soluble. The population of each Plant Functional Type

10 (PFT) is partitioned into mass classes with a fixed baseline mortality along with an assumed power-law scaling of growth-rate with mass . The analytical equilibrium solutions of RED allow the model to be calibrated against observed forest cover using a single parameter - the ratio of mortality to growth for a tree of a reference mass (μ_0). We show that RED can thus be calibrated to the ESA LC_CCI (European Space Agency Land Cover Climate Change Initiative) coverage dataset for nine PFTs. Using Net Primary Productivity and litter outputs from the UK Earth System Model (UKESM), we are able to diagnose the spatially

15 varying disturbance rates consistent with this observed vegetation map. The analytical form for RED circumnavigates the need to spin-up the numerical model, making it attractive for application in Earth System Models (ESMs). This is especially so given that the model is also highly parameter-sparse.

1 Introduction

A key requirement of Earth System Science is to estimate how much carbon the land surface will take-up in the decades ahead.

20 This is an important component of the total carbon budget consistent with avoiding global warming thresholds, such as 2K. Unfortunately, projections of future land carbon storage still span a wide-range.

Using an ensemble of Dynamic Global Vegetation Models (DGVMs), the Global Carbon Project estimates that the global land absorbed carbon at a rate of $3.2 \pm 0.8 \, \mathrm{GtC \, yr^{-1}}$ in 2018, which is approximately 30% of anthropogenic emissions (Le Quéré

et al., 2018). The cumulative uptake of land carbon across the Coupled Model Intercomparision Model Project 5 (CMIP5) under a common emissions scenario (RCP8.5) produced a range between $-185\,\mathrm{PgC}$ (source) to $758\,\mathrm{PgC}$ (sink) by the end of the twenty-first century. This uncertainty arose in part from differences in how land-use change (LUC) was modelled, with eight of the eleven ESMs within CMIP5 including LUC (Friedlingstein et al., 2014). However, even when focusing within CMIP5

5   ESMs including LUC the results are uncertain. Under RCP8.5, model runs had a range of $34-205\,\mathrm{PgC}$ net emitted carbon by the land surface at the end of the century (Brovkin et al., 2013), which can be attributed to uncertainties in both regrowth and CO_2 fertilisation. Furthermore, models have been shown to have varying responses in climate-induced land cover change (Davies-Barnard et al., 2015; Pugh et al., 2018).

DGVMs employ a variety of methods to simulate vegetation which contributes to the differences found at the Earth system
10   level (Fisher et al., 2018). Complex models such as the SEIB-DGVM (Sato et al., 2007) stochastically represent individuals on a meter by meter grid which is scaled-up to hundreds of kilometres squared. The benefit of this is that size-dependent physiology and spatial heterogeneity can be explicitly represented. However, multiple ensemble-members are needed to construct meaningful forest statistics, which makes such models computationally expensive to run at large scales. Top-down DGVMs, where all individuals of a Plant Functional Type (PFT) are essentially treated as of the same size, are significantly simpler
15   and more computationally efficient. However, processes that are dependent on size cannot be represented, and forest regrowth times are often over-estimated. In the JULES-TRIFFID coupled model (Best et al., 2011), regrowth from disturbances is often unrealistically long, resulting in fast-growing PFTs becoming dominant in dynamical runs with fire and LUC (Burton et al., 2019). Compromises between the complexity of individual and top-down DGVMs exists as a class of tree *Cohort* models. In the ED model (Moorcroft et al., 2001; Medvigy et al., 2009) the population is partitioned between patch disturbance and
20   biomass classes allowing for the scaling of process to be represented in both age and size. ED2 can realistically model forests around the world (boreal, rainforest and temperate) (Medvigy et al., 2009; Fisher et al., 2018). However parameterisation of competition within cohort DGVMs can result in a wide spread of outcomes when simulating climate change (Fisher et al., 2010; Scheiter et al., 2013). Such patch models can also be numerically unwieldy, as new patches (or 'tiles') are created after each disturbance event.

25   This paper represents a simplified cohort model (*Robust Ecosystem Demography (RED)*) which updates the number of trees in each mass class, but does not separately track tree-age or patch-age. This simplification significantly reduces the number of free parameters in the model, but still enables it to fit forest inventory data in North America (Moore et al., 2018) and South America (Moore et al., 2019).

2 Description of the Model

2.1 Theory

The underlying theoretical model for RED is a continuity equation, for each PFT and spatial location, which describes the time-evolution of the number density $n(m)$ of plants per unit area of mass m:

$$\frac{\partial n}{\partial t} + \frac{\partial}{\partial m} n\, g = -\gamma\, n. \tag{1}$$

Here $g(m)$ is the growth-rate and $\gamma(m)$ is the mortality rate of a plant of mass m. In general, g and γ could take any form, but we make simplifying assumptions for these functions consistent with observed $n(m)$ from forest inventory data (Moore et al., 2018, 2019). Simply, we assume that γ is independent of plant mass, and that $g(m)$ is a power-law of plant mass:

$$g = g_0 \left(\frac{m}{m_0} \right)^{\phi_g}, \tag{2}$$

Here g_0 is the growth rate of a plant with the reference mass, m_0. A value of $\phi_g = 0.75$ is assumed by default, consistent with (Niklas and Spatz, 2004). We also follow (Niklas and Spatz, 2004) in assuming the scaling of plant canopy area a with plant mass:

$$a = a_0 \left(\frac{m}{m_0} \right)^{\phi_a}, \tag{3}$$

where $\phi_a = 0.5$ by default.

Solutions for n can be integrated over mass to derive the total plant number, $N = \int n\, dm$, the total growth-rate, $G = \int g\, n\, dm$, the total biomass, $M = \int m\, n\, dm$, and the fractional area covered $\nu = \int a\, n\, dm$.

2.2 Discrete Mass Classes

We wish to produce a model of vegetation demography that can be updated numerically and which explicitly conserves vegetation carbon. In order to do this we integrate Eq. (1) over finite mass ranges:

$$\frac{\partial N_i}{\partial t} + F_i - F_{i-1} = -\gamma\, N_i, \tag{4}$$

where i denotes the ith mass class; F_i is the flux of plants growing out of the ith mass class and into the $(i+1)^{th}$ mass class; F_{i-1} is the flux of plants growing out of the $(i-1)^{th}$ mass class and into the ith mass class; and N_i is the number of plants per unit area in the ith mass class. For clarity, Eq. (4) is deliberately presented as continuous in time at this stage, as the focus in this subsection is on discretization of the mass profile. The fully numerical version of RED, which includes discretization of time, is described in Section 2.4 and 2.5. In order to conserve carbon (see below) the flux F_i must take the form:

$$F_i = \frac{N_i\, g_i}{(m_{i+1} - m_i)}, \tag{5}$$

where m_i is the mean mass of a plant in the ith mass class, and g_i is the growth-rate per plant of the ith mass class [kgC yr^{-1} plant^{-1}].

For large-scale application in ESMs, a primary concern is to ensure that the total vegetation carbon obeys carbon balance (i.e. only changes due to the net impact of total growth minus total mortality). The total vegetation carbon in each mass class is $M_i = m_i N_i$. The update equation for M_i is therefore Eq. (4) multiplied by m_i:

$$\frac{\partial M_i}{\partial t} + m_i \left(F_i - F_{i-1} \right) = -\gamma M_i. \tag{6}$$

5 The total carbon in the vegetation, M, is the sum of the carbon in each of the mass classes:

$$M = \sum_i M_i. \tag{7}$$

Thus the update equation for the total carbon is:

$$\frac{\partial M}{\partial t} + \sum_i m_i \left(F_i - F_{i-1} \right) = -\gamma M, \tag{8}$$

which can be rewritten as:

10 $$\frac{\partial M}{\partial t} + \sum_i F_i \left(m_i - m_{i+1} \right) = -\gamma M. \tag{9}$$

Now substituting Eq. (5) into Eq. (9) gives:

$$\frac{\partial M}{\partial t} = \sum_i N_i g_i - \gamma M. \tag{10}$$

The first term on the righthand-side of this equation is the total carbon uptake due to growth, and the second-term represents the total carbon loss due to mortality, which is the required carbon conservation equation.

15 ## 2.3 Seedling production and gap competition

To solve Eq. (4) we also require a lower boundary condition which represents the rate at which seedlings of mass m_0 are introduced into the cohort. Here we assume that a fixed fraction α of the total assimilate available to a PFT (P), is devoted to producing new seedlings, with the remainder G being allocated to the growth of existing plants. In addition, we assume that only those seedlings growing in 'gaps' will survive. The net incoming flux of seedlings of mass m_0 is therefore:

20 $$F_0 = \frac{\alpha P}{m_0} s = \frac{\alpha}{(1 - \alpha)} \frac{G}{m_0} s, \tag{11}$$

where s is the fractional gap area available for seedlings. The definition of s is assumed to differ by PFT to reflect an underlying tree-shrub-grass dominance hierarchy, as shown schematically in Figure 1.

Equation (11) assumes a random overlap between seedlings and the existing vegetation. This lower boundary condition is the only place within RED where there is significant competition. Minimum overlap, which is broadly consistent with 'perfect 25 plasticity' (Strigul et al., 2008), is assumed once seedlings have been injected into the cohort according to Eq. (11).
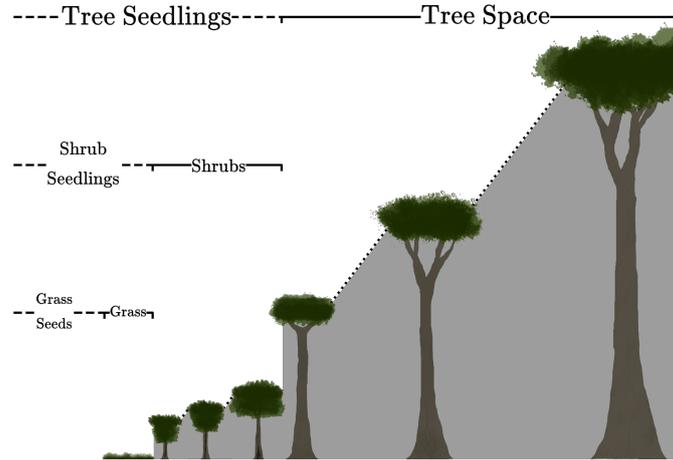
Figure 1. Illustration depicting the hierarchical PFT functional group regime within RED. Trees shade Trees, Shrubs and Grasses. Shrubs shade Shrubs and Grasses, while grasses only shade grasses.

The space available to the seedlings of the kth PFT is calculated from the area fractions of the PFTs to which it is subdominant:

$$s_k = 1 - \sum_l c_{kl}\nu_l, \tag{12}$$

where ν_l is the area fraction of the lth PFT, and c_{kl} is the competition coefficient for the impact of PFT l on PFT k. If PFT l
5 is within the same plant functional group (trees, shrubs or grasses) as PFT k, or dominant over it, $c_{kl} = 1$. If PFT k is dominant over PFT l, $c_{kl} = 0$ (Figure 1). This is similar to the competition regime in the TRIFFID model (Cox, 2001), and allows for the co-existence between inter-functional groups of PFTs. The hierarchy also enables the simulation of succession during regrowth. Faster growing species of grasses will not be able to expand into space occupied by trees and shrubs, unless there is space created by disturbance. A summary of the competition coefficients is given in table 1.

Table 1. The competition coefficients values for different Plant Functional Groups. A more detailed example of this is given with specific PFTs in table 2.

	c_{kl}	Trees	Shrubs	Grasses
	Trees	1	0	0
k	Shrubs	1	1	0
	Grasses	1	1	1

2.4 Coupling to Earth System Models

RED updates plant size distributions, biomass, and fractional areal coverage for an arbitrary number of PFTs at each spatial location, and can be driven by variables provided by a land carbon cycle model, an Earth System Model, or observations (see Figure 2). For each PFT, the minimum required input is a time-series of net carbon assimilate (P), defined as the difference

5 between Net Primary Productivity (Π_{N}), and local litter production due to turnover of leaves, stems and roots (Λ_{l}):

$$P = \Pi_{\mathrm{N}} - \Lambda_{\mathrm{l}}. \tag{13}$$

Where available, additional mortality due to disturbance events such as droughts, fires and anthropogenic deforestation (γ_{d}) can be added to the baseline mortality rates (γ_{b}), for each PFT:

$$\gamma = \gamma_{\mathrm{b}} + \gamma_{\mathrm{d}}. \tag{14}$$

These γ_{d} values can in principle be both PFT-dependent and mass-dependent (e.g. to capture forestry practices).
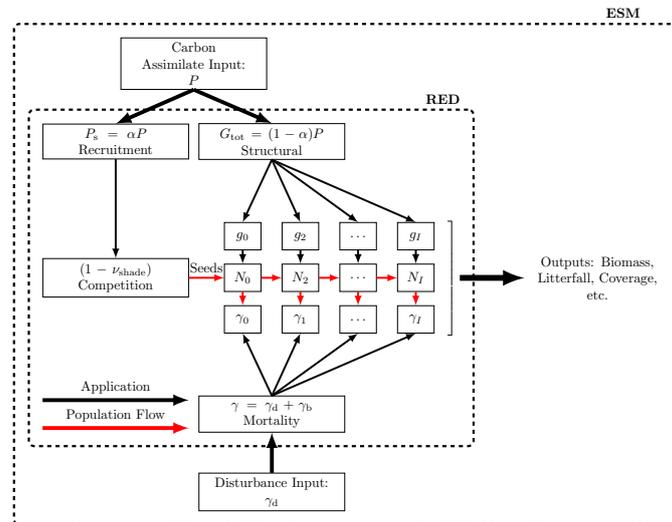
Figure 2. Schematic of RED coupled to an ESM on land carbon cycle model. RED is driven by a time-series of net carbon assimilate, P, which is then split between seedling production, αP, and the growth of existing plants, $G = (1 - \alpha) P$. The seedling flux is limited by the available free space, s. Additional mortality rates diagnosed from disturbance models, γ_{d}, can be added on to an assumed baseline mortality, γ_{b}, as a function of both PFT and mass-class.

The input values of net assimilate for each PFT (P), define the total structural growth-rate, $G = (1 - \alpha) P$, and the seedling flux F_0 (via Eq. (11)), using PFT-specific values of the parameter α (see table 2). The definition of the total structural growth-rate at a given time-step:

15 $$G = \sum_i N_i g_i, \tag{15}$$

can be combined with the growth-scaling given by Eq. (2), to derive the reference growth-rate, g_0, from the net assimilate, P, which is a driving input:

$$g_0 = \frac{(1-\alpha)\,P}{\sum_i N_i \left(\frac{m_i}{m_0}\right)^{\phi_g}}. \tag{16}$$

This in turn enables the growth-rate of each mass class to be calculated using Eq. (2). For each PFT, the number of plants in
5 mass class (N_i) is updated using a discretised form of Eq. (4):

$$N_i^{(j+1)} = N_i^{(j)} + \Delta t \left(F_{i-1}^{(j)} - F_i^{(j)} - \gamma^{(j)} N_i^{(j)} \right). \tag{17}$$

where Δt is the RED timestep (typically 1 month), and the superscript $^{(j)}$ denotes the jth timestep. The lower boundary
seedling flux is calculated from Eq. (11) using Eq. (12). We do not impose a zero-flux condition out of the upper mass class,
under the assumption that there will be enough mass classes to ensure that this flux is negligible. However, to ensure carbon
10 conservation on the land we add any plants that grow out of the upper mass class into a demographic litterfall term for each
PFT, which is a RED output. This demographic litterfall term, Λ_d, keeps track of the carbon lost from the vegetation due to
competition, mortality and the carbon in any such plants that grow out of the largest resolved mass-class (class I):

$$\Lambda_{\mathrm{d}} = \alpha\,P\,(1-s) + \sum_i \gamma_i\,M_i + g_I\,N_I. \tag{18}$$

The first term on the righthand-side of this equation represents carbon loss due to the shading of seedlings; the second term
15 represents mortality of the resolved mass classes (which may include disturbance events); and the third term, which is normally
very small, is the loss of vegetation carbon due to plants growing beyond the modelled mass classes. In order to initiate
regrowth from bare soil, RED also assumes a minimum effective fractional area of each PFT. Where the net assimilate would
be sufficiently negative to take the vegetation fraction below this minimum, the minimum value is maintained by subtraction
from the demographic litter. The demographic litterfall term therefore represents the net addition litter production consistent
20 with the prescribed net assimilate flux, the disturbance rate, and the change in vegetation carbon modelled by RED. When
coupling to an ESM or land carbon model, the demographic litterfall term (Λ_d) should be added to the input local litterfall (Λ_l)
(as used in Eq. (13)), to calculate the total litterfall flux into the soil/litter system.

2.5 Steady-State

The steady state of the continuum model defined by Eq. (1) and Eq. (2) can be solved analytical for each PFT (Moore et al.,
25 2018, 2019). The continuum analytical solutions for the equilibrium mass distribution ($n_{\mathrm{eq}}(m)$), the total plant number (N_{eq}),
biomass (M_{eq}), growth-rate (G_{eq}) and fractional area (ν_{eq}) are summarised in Appendix B. The shape of the mass distribution
and each of these parameters depend on the ratio of plant mortality to growth, which we choose to define for the reference
mass class m_0:

$$\mu_0 = \frac{\gamma\,m_0}{g_0}. \tag{19}$$

In order to initialise the numerical RED model in a drift-free initial state, we also derive the steady-state of the discrete model (Equation (17)), which will differ from the continuum model for a finite number of mass classes. The equilibrium solution of Eq. (17) defines a recursive relationship for the number of plants N_i in each mass-class:

$$N_i = \lambda_i N_{i-1}, \tag{20}$$

5   where

$$\lambda_i = \frac{(m_{i-1}/m_0)^{\phi_g} m_0/(m_i - m_{i-1})}{(m_i/m_0)^{\phi_g} m_0/(m_{i+1} - m_i) + \mu_0}. \tag{21}$$

Thus for the discrete model the shape of the mass-number distribution also depends on the mortality-to-growth parameter, μ_0. Repeated action of Eq. (20) gives an equation for the total plant number in terms of the number of plants in the lowest mass-class, N_0:

10    $$N_{\text{eq}} = \sum_{i=0}^{I} N_i = N_0 X_N. \tag{22}$$

Where X_N:

$$X_N = 1 + \sum_{i=1}^{I} \prod_{j=1}^{i} \lambda_j, \tag{23}$$

and I is the top mass-class. Where i and j are class indices over the sum product. Likewise, we can calculate:

1. the total structural growth at equilibrium:

15    $$G_{\text{eq}} = \sum_{i=0}^{I} N_i g_i = N_0 g_0 X_G. \tag{24}$$

where Eq. (2) implies:

$$X_G = 1 + \sum_{i=1}^{I} \left(\frac{m_j}{m_0}\right)^{\phi_g} \prod_{j=1}^{i} \lambda_j. \tag{25}$$

2. the total biomass at equilibrium:

$$M_{\text{eq}} = \sum_{i=0}^{I} N_i m_i = N_0 m_0 X_M, \tag{26}$$

20   where:

$$X_M = 1 + \sum_{i=1}^{I} \frac{m_i}{m_0} \prod_{j=1}^{i} \lambda_j. \tag{27}$$
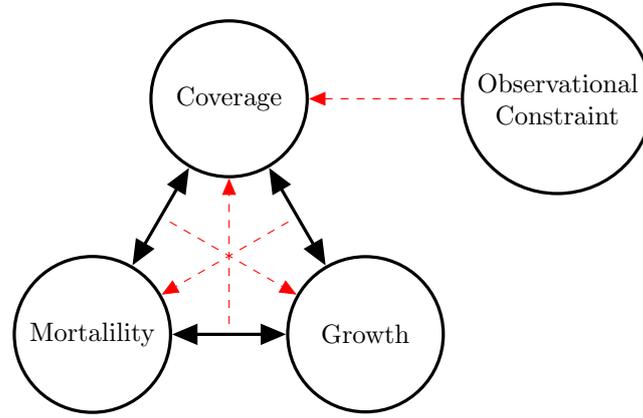
Figure 3. We can use an observation to constrain (red dashed arrow) μ_0 giving us an equilibrium coverage. If the addition of total growth is known (black double-headed arrow), we can rearrange μ_0 to find an equilibrium mortality rate (or vice-versa).

3. the fractional area covered by the PFT at equilibrium:

$$\nu_{\mathrm{eq}} = \sum_{i=0}^{I} N_i\, a_i = N_0\, a_0\, X_\nu, \tag{28}$$

where Eq. (3) implies:

$$X_\nu = 1 + \sum_{i=1}^{I} \left(\frac{m_i}{m_0}\right)^{\phi_a} \prod_{j=1}^{i.} \lambda_j. \tag{29}$$

5   The equations above therefore define the equilibrium state of the discrete system for given values of N_0 and μ_0. The value of μ_0 can be estimated from forest demographic data where this is available (Moore et al., 2018, 2019). However, for global applications we rarely have more observations than the fractional coverage of each PFT. Under these circumstances, we use the condition that at equilibrium the rate of injection of seedlings (Equation (11)) must balance the rate of loss of plants due to mortality (γN_{eq}):

$$\gamma N_{\mathrm{eq}} = \frac{\alpha}{(1-\alpha)}\, \frac{G}{m_0}\, s. \tag{30}$$

Now substituting in Eq. (22), Eq. (24) and Eq. (12) yields a balance equation for the kth PFT:

$$\left(\frac{\alpha}{1-\alpha}\right)\left(1 - \sum_l c_{kl}\, \nu_l\right) = \mu_0\, \frac{X_N}{X_G} = f(\mu_0), \tag{31}$$

where X_N and X_G are given by Eq. (23) and Eq. (25) respectively. As the righthand-side of this equation depends only on prescribed constants and μ_0, Eq. (31) can be inverted (by numerical iteration) to estimate μ_0 for observed values of the PFT 15   fractions (ν_l) and an assumed value of α (see Table 2). Once the value of μ_0 has been derived in this manner, it can be used to calculate X_ν via Eq. (29), and therefore N_0 by inversion of Eq. (28):

$$N_0 = \frac{\nu_{\mathrm{eq}}}{a_0\, X_\nu}. \tag{32}$$

Equations (31) and (32) therefore allow us to define an initial equilibrium state (N_i) which is consistent with observed area fractions of each PFT (Figure 3). Furthermore, when paired with an estimate of the net carbon assimilate (from a model or observations), the μ_0 estimate can be converted into a map of the implied mortality (γ) by PFT. We demonstrate this capability as part of the global tests of RED described in the next section.

3 Modelling Results

For these tests, the numerical RED model is set-up to use the 9 PFTs which are currently used in JULES (Harper et al., 2018). This enables us to directly use driving data - time-series of the rate of net assimilation (P) - from a previous UKESM model simulation that includes JULES (Sellar et al., in prep). RED is integrated forward using a one month time-step and successive mass-classes that differ by a multiplicative constant ξ, so that $m_i = \xi\, m_{i-1}$. The value of ξ was chosen to optimally fit the analytical equilibrium solutions assuming 10 mass classes for trees, 8 mass classes for shrubs and 1 mass class for grasses, assuming $\mu_0 = 0.25$ (see Appendix B2). Other PFT-specific parameters are assumed as summarised in Table 2.

Table 2. The PFT list and their corresponding parameters (m_0, a_0, h_0), seedling fraction (α) and competition coefficient ($c_{\mathrm{pft,j}}$). The growth allometry of trees and shrubs across size is assumed to follow (Niklas and Spatz, 2004) ($\phi_g = 0.75$, $\phi_a = 0.5$, $\phi_h = 0.25$). The competition coefficients given describe which PFT functional group shades the current PFT, if $c_{\mathrm{pft,j}} = 1$, the PFT is shaded, otherwise it is not (Table 1).

Long-name	Abbrev	Classes	Scaling (ξ)	α	$m_0(\mathrm{kgC})$	$a_0(\mathrm{m}^2)$	$h_0(\mathrm{m})$	Tree	Shrub	Grass
Broadleaf Evergreen Tree Tropical	BET-Tr	10	2.32	0.10	1.00	0.50	3.00	1	0	0
Broadleaf Evergreen Tree Temperate	BET-Te	10	2.32	0.10	1.00	0.50	3.00	1	0	0
Broadleaf Deciduous Tree	BDT	10	2.35	0.10	1.00	0.50	3.00	1	0	0
Needleleaf Evergreen Tree	NET	10	2.35	0.10	1.00	0.50	3.00	1	0	0
Needleleaf Deciduous Tree	NDT	10	2.32	0.10	1.00	0.50	3.00	1	0	0
Cool Season Grasses	C3	1	1.50	0.60	0.10	0.25	0.05	1	1	1
Tropical Grasses	C4	1	1.50	0.60	0.15	0.25	0.05	1	1	1
Evergreen Shrub	ESh	8	2.80	0.35	0.15	0.25	3.00	1	1	0
Deciduous Shrub	DSh	8	2.80	0.35	0.50	0.25	3.00	1	1	0

3.1 Local: Simulating Succession

We begin by demonstrating the vegetation succession simulated by RED in an idealised spin-up from bare-soil (a minimum vegetation fraction of 0.001 occupying the lowest class.), in this case for a grid-box at the edge of the Amazonian rainforest
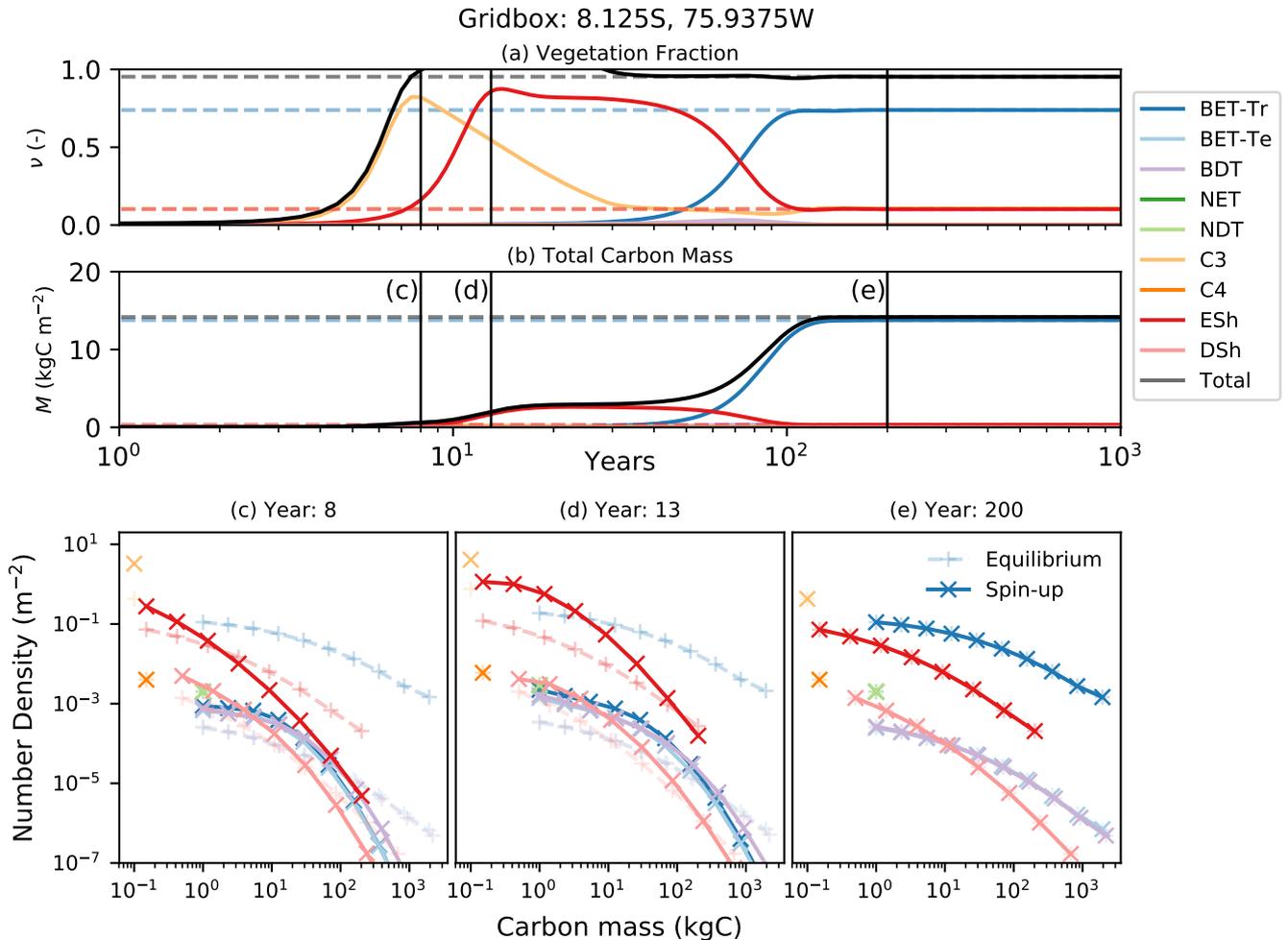
Figure 4. Dynamical run of two RED simulations at grid-box level, dashed transparent lines initialised with the equilibrium solutions, the other (solid lines), initialises from bare soil. Both use the same constant assimilate rate from the UKESM dataset per coverage. (a) is the total vegetation fractions among each of the PFTs. (b) is the corresponding biomass. (c), (d) and (e) provide snapshots of the number density distribution of the PFTs across mass classes at different times. Lines marked as **+** are the equilibrium runs while **X** indicates the spin-up run.

(Figure 4). Under these circumstances, the diagnosed initial state is indeed the long-term equilibrium state, as evidenced by the horizontal dashed lines in panels a and b of Figure 4.

Faster growing grass PFTs dominate the grid-box within the first twelve years, before being replaced by evergreen shrubs which shade the grass seedlings. Eventually, Broad-leaf Evergreen Tropical trees replace much of the shrub and grass, on a timescale determined in large-part by the parameter α and the reference mass-class m_0. With the parameters used here, the vegetation fraction reaches close to its equilibrium value after about 20 years (panel a), but full spin-up of the biomass takes

around 150 years (panel b). Having a fixed rate of growth per coverage results in a changing g_0, as the allometiric scaling for growth and coverage differs.

The modelled evolution of the plant number versus mass distribution for each PFT is shown in panel c (after 6 years), panel d (after 13 years) and panel (after 100 years), with the eventual demographic equilibrium profiles shown by the dashed-lines. It is clear that grass PFTs are close to their demographic equilibrium after only 6 years, but tree PFTs need more than 100 years to reach equilibrium.

The dashed-lines in Figure 4 represent a dynamical RED simulation from the diagnosed demographic equilibrium state. This state is derived using the methodology described in Section 2.5, with one significant change. The competition rules given by Eq. (12) and Table 1 result ultimately in equilibria which have a single dominate PFT in each class of co-competing types (trees, shrubs, grasses). To avoid drifts associated with the competitive exclusion of the subdominant PFTs in each vegetation class, we choose to initialise the dominant PFT to have the total area fraction of all the PFTs in that vegetation class.

3.2 Global: Simulation of Current Vegetation Map

In this section we use a similar procedure to diagnose the map of PFT-specific mortality rates consistent with the current observed vegetation state, and rates of net assimilation (P) provided by UKESM (Sellar et al., in prep). The observed maps of PFTs are provided by the ESA LC_CCI dataset for 2008-2012 (Poulter et al., 2015), projected onto the 9 JULES PFTs (Figure 5). Maps of the prescribed annual mean values of the rate of net assimilation (P) are shown in Figure 6.
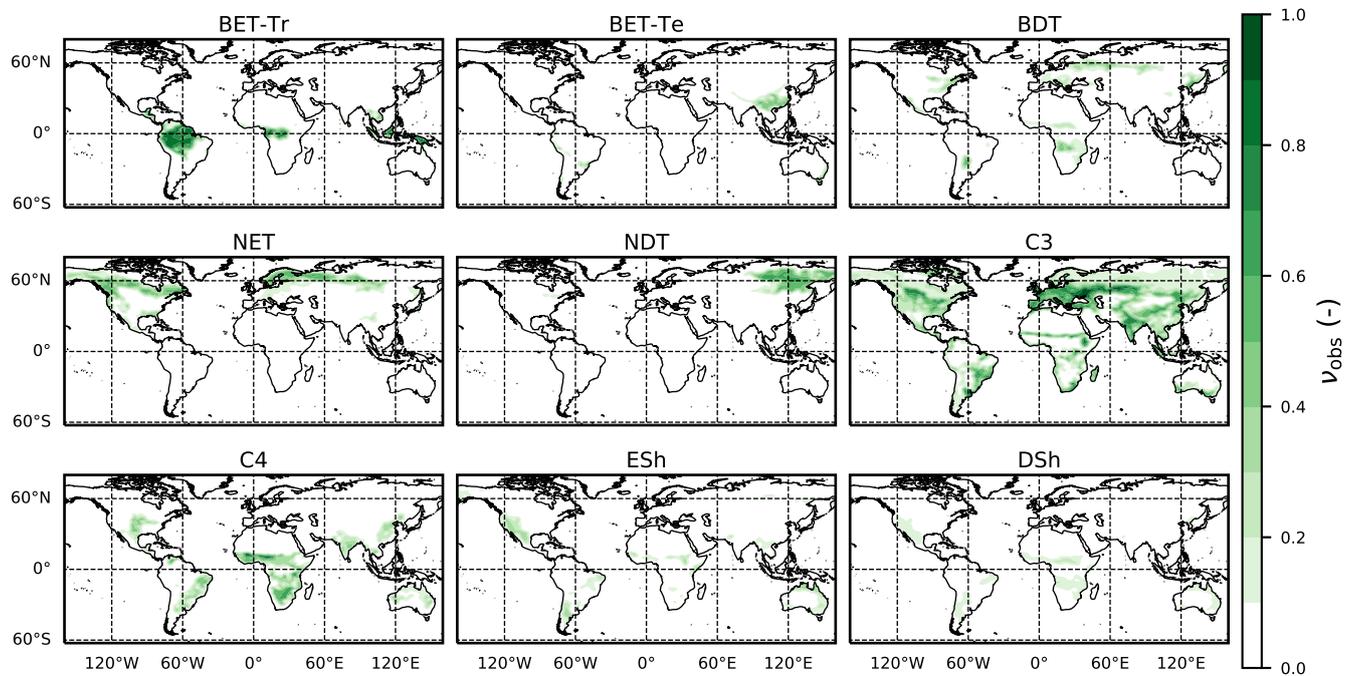
Figure 5. Observed dataset of the PFT fraction for the nine JULES PFTs (Harper et al., 2016), as given by the title abbreviations listed in Table 2.
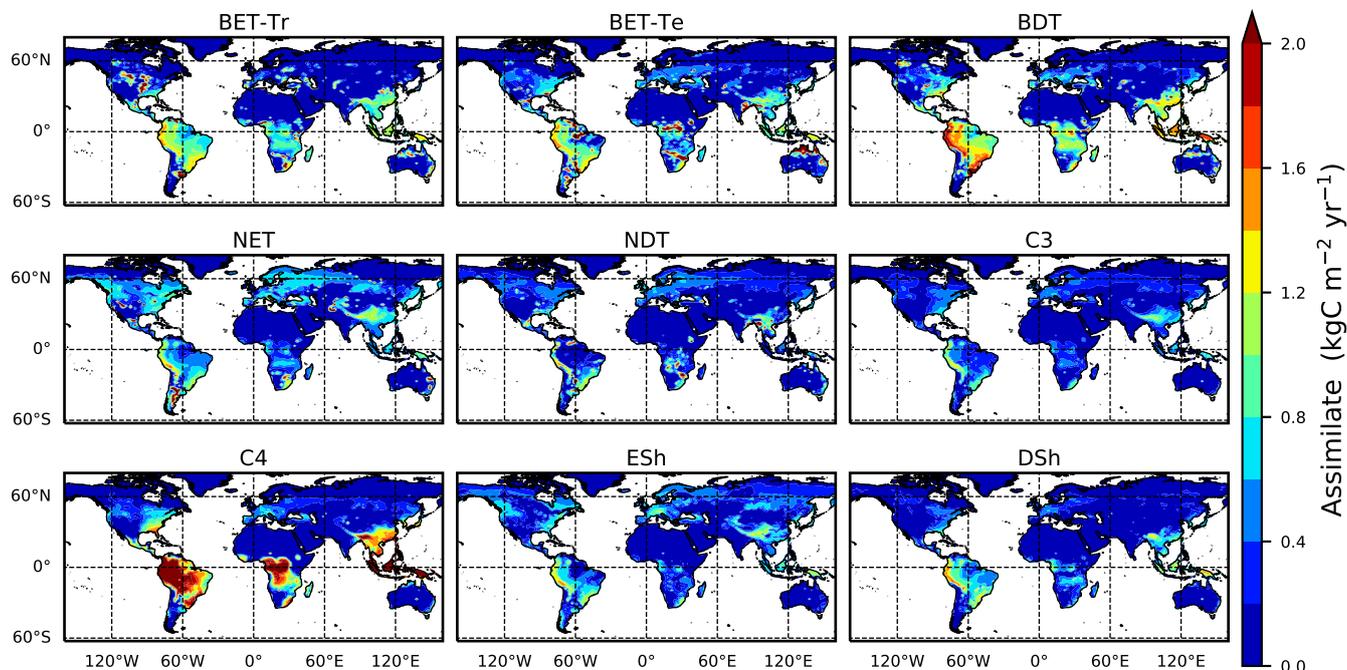
Figure 6. Driving assimilate (Equation (13)) from UKESM, averaged between 2000-2010. The average is constructed by truncating any negative growth rates to zero.

We use the procedure outlined in Section 2.5 to estimate spatially-varying values of μ_0 for each PFT, using Eq. (31), and then Eq. (32) to estimate N_0. This method successfully reproduces the ESA map of dominant PFT to good accuracy, as shown in Figure 7 and Table 3.
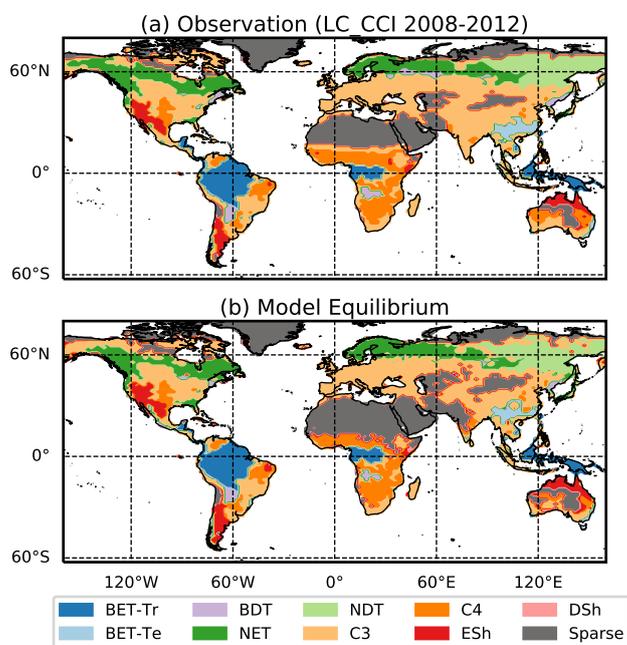
Figure 7. (a) Map of observed PFT cover and (b) RED equilibrium fractions assigned to the nine PFTs Harper et al. (2016). Colour-coded are the PFTs possessing the largest faction within the grid-box. Sparse area is defined as total vegetation coverage of less than 10%.

Table 3. Goodness of fits for the RED equilibrium coverages to that of the ESA LC_CCI dataset across the PFTs. r represents the Persons Correlation Coefficient. The RMSE has been weighted by the grid-box area to the mean grid-box area to account for latitudinal variations of grid-box areas.

PFT	r	RMSE
BET-Tr	0.990	0.030
BET-Te	0.935	0.030
BDT	0.783	0.053
NET	0.905	0.051
NDT	0.928	0.033
C3	0.895	0.129
C4	0.818	0.088
ESh	0.854	0.051
DSh	0.525	0.049

There are a few reasons why the model equilibrium fractions are not an identical fit with the observed fractions. Firstly, the two datasets are not necessarily consistent - there are a few places (Central Asia, Sahel) were the average UKESM assimilates used is zero, not aligning with the positive coverage from the ESA dataset. Secondly, areas of mixed PFTs within the same vegetation class, as previously stated, will have an adjusted RED equilibrium fraction where the dominant PFT equilibrium

5 will be the sum of the vegetation class.
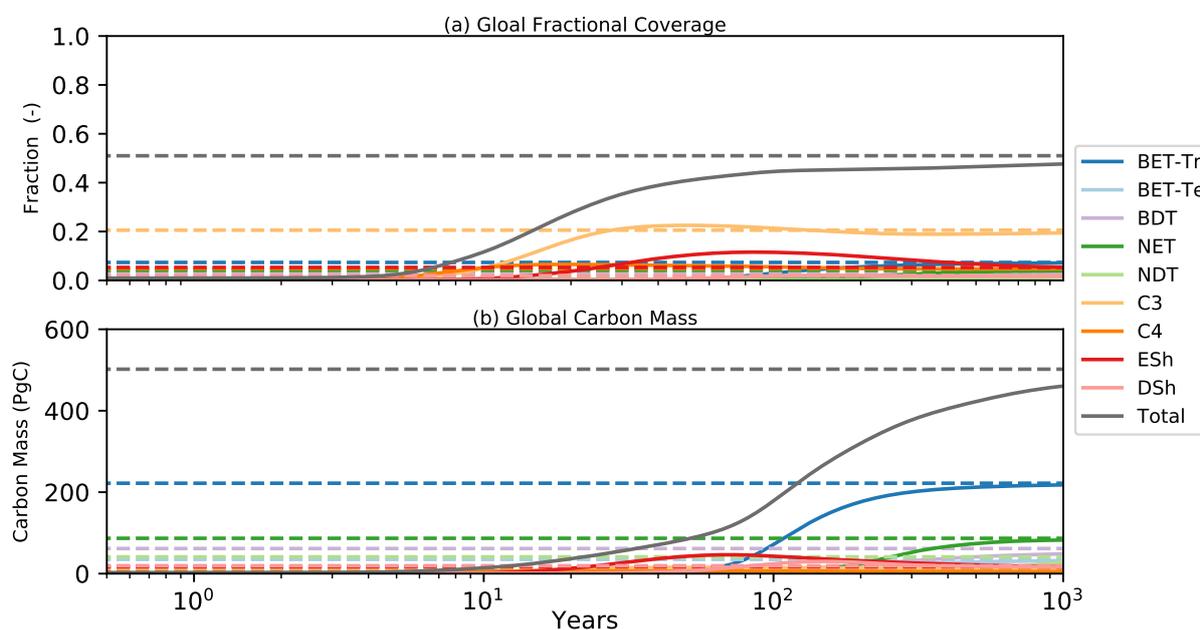
Figure 8. As for figure 4, Solid lines are spin-up from bare soil, dashed lines are the equilibrium instillation run. Panel (a) represents the fractional global coverage relative to the total land area; panel (b) represents the total biomass of the vegetation.

As stated within the previous section, Figure 8, demonstrates that the initialisation scales globally. There are two RED simulations, one being started from bare soil the other from the diagnosed equilibrium. Using a constant assimilate rate (Figure 6) and the mortality distribution (Figure 10), we see convergence towards the two runs in global coverage (Panel a) and global biomass (Panel b). The equilibrium initialised run also demonstrates the practical advantage RED posses in avoiding spinning

10 up. In instances that have a low spin-up growth rates, this saves significant time.

3.3 Global: Diagnosed Plant Mortality Rates

As we now have diagnosed values of μ_0 and N_0, along with prescribed values of P, we can also diagnose the mean plant mortality rate γ, for each location and for each PFT, from Eq. (19) :

$$\gamma = \frac{\mu_0 \, g_0}{m_0},\tag{33}$$

5   where g_0 is given by Eq. (16) combined with Eq. (20) and Eq. (21). The maps of γ values, derived in this way, are shown in Figure 9.
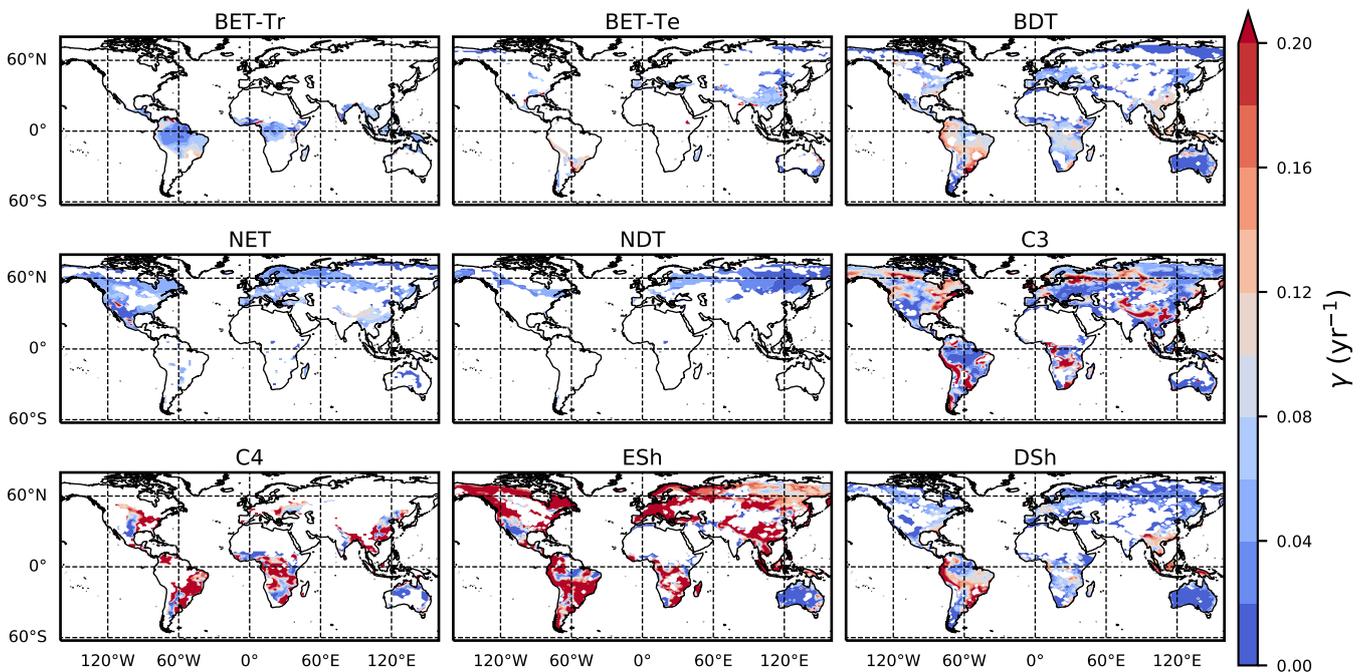
Figure 9. Diagnosed maps of mortality rates γ for each PFT, as required for consistency with the ESA observations and the UKESM growth rates. White areas correspond with zero coverage and/or zero growth - mortality is assumed infinite within RED.

The mortality rate derived is very dependent on the overall coverage and the total assimilate. Having a high coverage with a low growth rate will result in RED compensating through having a low mortality rate (and vice-versa). This explains why some mortality rates of PFTs seem to posses large variations (Figure 10). Furthermore, the choice of α (Equation (16)) and
10   the m_0 is also influential when it comes to the value of γ. Under the assumption that high coverages are close to a 'healthy' environment for a PFT. We can take a sub-sample of the grid-boxes that are within the top quartile of non-zero coverages ($\nu_{\mathrm{eq}} > 0.01$) (Table 4). The median μ_0 value diagnosed from the top quartile of BET-Tr of $0.232^{+0.008}_{-0.007}$ (Table 3), is close to the values calculated by our previous paper (Moore et al., 2019) of approximately 0.235 for all of South America. The value within the paper 0.198 is converted to dry carbon mass through Eq. (B.9). Carbon dry mass constitutes approximate half of the

total dry mass (Thomas and Martin, 2012). In addition to parametrisations (such as α), some of the differences between these μ_0 values can arise from the discretization of the model, as the discretised form will underestimate the diagnosed μ_0 values to meet the same observation - when compared with the continuous form (Figure B1.a). A possible usage of RED might be to diagnose a μ_0, fix a baseline mortality rate from surveys and diagnose the required carbon assimilate to match an additional

5   observation (Figure 3). Potentially providing a future constraint on ESM growth rates for PFTs.
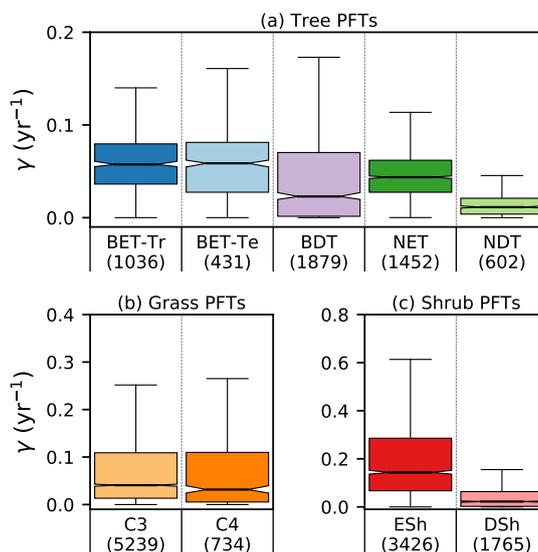
Figure 10. Diagonsed mortality rates for (a) trees, (b) grasses and (c) shrubs for non-zero coverage ($\nu_{eq} > 0.01$). Hatches within the box represent the confidence bounds of the median. The confidence bounds are estimated using a bootstrap method. Bracketed numbers represent the number of grid-points.

There have been multiple site-level assessments of the rates of stand mortality within pan-tropical forests - typically the background rate is between $1 \ \mathrm{yr}^{-1}$ to $4 \ \mathrm{yr}^{-1}$ (Lugo and Scatena, 1996; Phillips, 1996; Phillips et al., 2004). (Phillips, 1996) estimates 60 mortality rates collected across 40 pan-tropical sites for tree sizes greater than $10 - 25 \ \mathrm{cm} \ \mathrm{dbh}$, Later work by Phillips in 2004 used the demographic data from the RAINFOR dataset of trees $\geq 10\mathrm{cm} \ \mathrm{dbh}$. Using these site assessments,

10   we can make a comparison to BET-Tr equilibrium mortality rates by looking at the values of γ in areas where we would expect to see old growth forests. Within the top 25% of coverages, we assume represent areas of undisturbed forest. BET-Tr captures the baseline mortality rates seen post 2000 in the Amazon from site data (Figure 10).

Table 4. The area weighted median values of driving coverage and assimilate against μ_0 and γ for the upper quartile of grid-boxes of greater than 0.01 for each PFT.

PFT	Area weighted median			
	ν_{obs}	P (kgC m^{-2} yr^{-1})	μ_0	γ (yr^{-1})
BET-Tr	$0.793^{+0.019}_{-0.023}$	$0.731^{+0.054}_{-0.041}$	$0.232^{+0.008}_{-0.007}$	$0.032^{+0.002}_{-0.001}$
BET-Te	$0.402^{+0.020}_{-0.030}$	$0.349^{+0.022}_{-0.028}$	$0.340^{+0.006}_{-0.004}$	$0.059^{+0.003}_{-0.003}$
BDT	$0.238^{+0.011}_{-0.011}$	$0.143^{+0.018}_{-0.014}$	$0.377^{+0.013}_{-0.011}$	$0.052^{+0.003}_{-0.003}$
NET	$0.471^{+0.009}_{-0.011}$	$0.281^{+0.005}_{-0.013}$	$0.328^{+0.008}_{-0.009}$	$0.036^{+0.002}_{-0.002}$
NDT	$0.597^{+0.010}_{-0.015}$	$0.112^{+0.009}_{-0.008}$	$0.298^{+0.008}_{-0.007}$	$0.011^{+0.001}_{-0.001}$
C3	$0.566^{+0.011}_{-0.007}$	$0.124^{+0.008}_{-0.006}$	$0.163^{+0.017}_{-0.013}$	$0.023^{+0.002}_{-0.003}$
C4	$0.545^{+0.043}_{-0.053}$	$0.123^{+0.084}_{-0.040}$	$0.189^{+0.044}_{-0.027}$	$0.029^{+0.006}_{-0.010}$
ESh	$0.142^{+0.009}_{-0.007}$	$0.028^{+0.002}_{-0.001}$	$0.744^{+0.019}_{-0.021}$	$0.094^{+0.010}_{-0.004}$
DSh	$0.116^{+0.010}_{-0.015}$	$0.024^{+0.006}_{-0.004}$	$0.713^{+0.046}_{-0.027}$	$0.036^{+0.005}_{-0.007}$
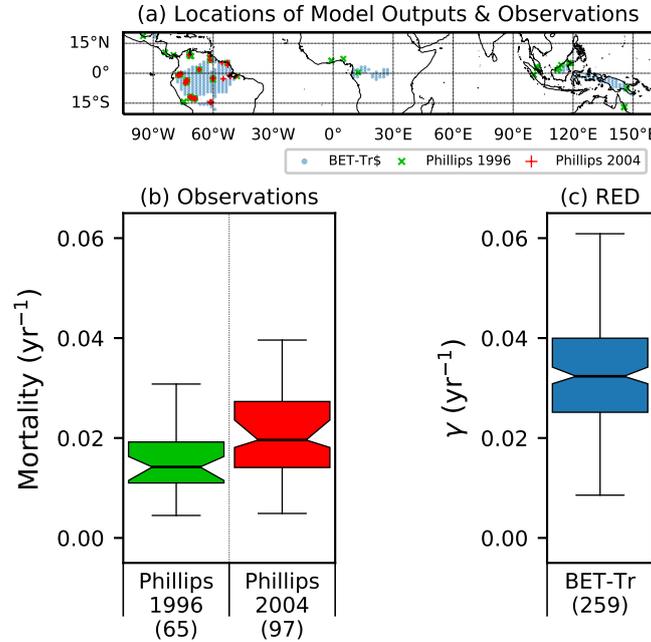
Figure 11. Comparision of observed tropical tree mortality with γ. Comparing datasets presented in (Phillips, 1996; Phillips et al., 2004) for an adjusted estimate of observed stand mortality with the equilibrium mortality rates for BET-Tr within the largest 25% of fractions. (a) shows the observations locations of sites (blue and green crosses) versus the chosen RED grid-points (red circles). (b) shows the distribution of mortality across grid-boxes, while (c) shows the mortality distribution across the BET-Tr grid-points. Red diamonds represent the mean, while the bracketed numbers represent the number of measurements from the datasets.

Next we investigate if the equilibrium mortality rates implicitly capture areas of disturbances. We compare the mean woody equilibrium mortality rate to fire and land-use surveys. The woody mortality is defined as the sum of γ weighted against the coverage of Trees and Shrubs to their collective coverage. Areas with large rates of disturbances area generally not expected to conform to the equilibrium assumptions, such as DET, used to initialise RED. Generally, biomes, that are humid and dry such
5 as savannah or grassland, wildfires play a natural part in maintaining the balance of vegetation (Bond et al., 2005), therefore for woody PFTs we expect to see a raised mortality/higher μ_0. Using the ESA FIRE_CCI dataset (Chuvieco et al., 2019) we can estimate the rate of burnt fraction per year:

$$\mathrm{BF} = \frac{\mathrm{Burnt\ Area}}{\mathrm{Burnable\ Fraction} \times \mathrm{Area}} \times \frac{1}{\Delta t}, \tag{34}$$

where Burnt Area and Burnable Fraction are given from the dataset. Area is inferred from the longitude, latitude quadrant.
10 The resolution of FIRE_CCI is $0.25° \times 0.25°$ is bi-linearly interpolated onto the simulated grid-boxes. Taking the averages of the burnt fraction rate between the months between 2000 and 2010, and converting into an annual burn rate.

We also carry-out a comparison with agriculture, we expect that in area of land-use we will see raised mortality for woody PFTs as the industry will use the space for crops and pasture. There is difficulty in getting an explicit rate of clearance from

land-use, however, by comparing with the fraction of cropland we achieve a non-direct geographic comparison to the to the dataset (Figure 12). For the crop fraction we use the 2000 ESA LC_CCI inferred PFT from (Li et al., 2019) of half a degree resolution - again interpolating onto the RED grids.
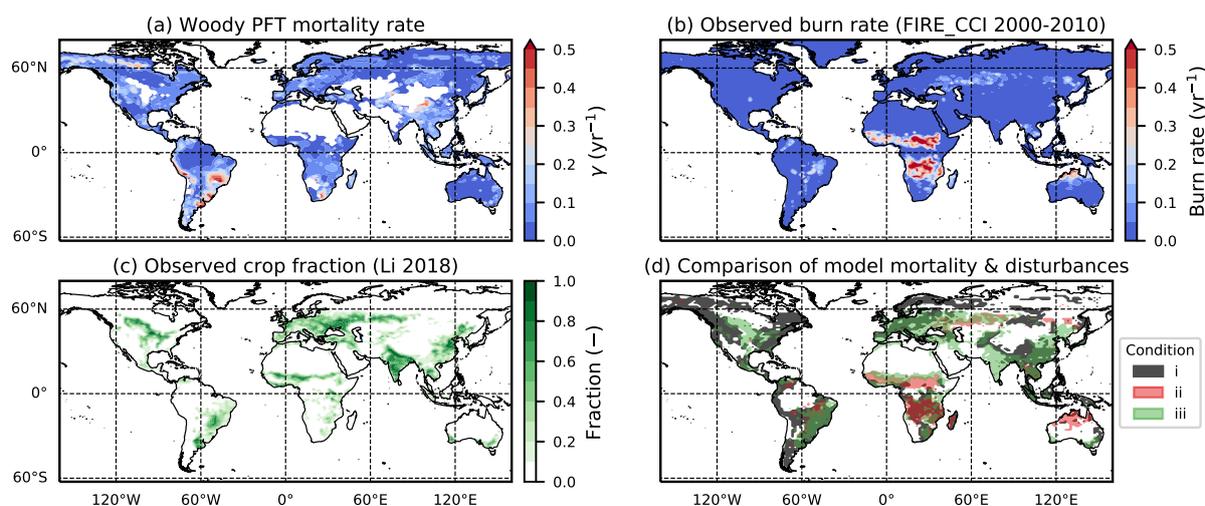
Figure 12. Comparison of the mean grid-box baseline mortality rate for 'Woody' PFTs (BET-Tr, BET-Te, BDT, ESh, DSh) (a) to disturbances - With fractional rate of burnt area (area burnt divided by burnable area) from ESA FIRE_CCI (b) and with a crop coverage from the PFT classification of the ESA LC_CCI 2000 Dataset (c). (d) describes where we see higher rates of morality (condition i, $\gamma > 0.05$ yr^{-1}) overlapped with areas of fire (condition ii, BF > 0.05 yr^{-1}) and land-use (condition iii, $\geq 10\%$).
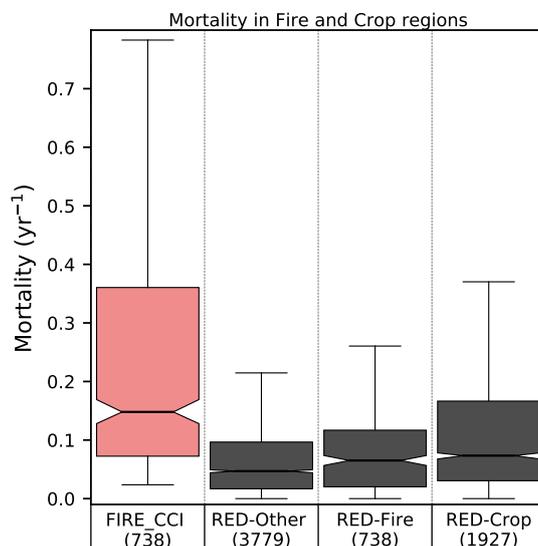
Figure 13. Comparison of the mean woody baseline mortality rate within areas of remotely sensed fire fractions from the FIRE_CCI and crop-land areas.

There are areas of spatial overlap of disturbance with that larger mortality rates (defined as $\gamma \geq 0.05\mathrm{yr}^{-1}$ and a similar burn-rate) with Figure 12.d. This would be expected, as areas of increased disturbances typically have PFTs with faster demographic turnovers with grasses and shrubs, dominating the environment, with the local areas either being more fire prone, or under agriculture. We see this clearly in areas such as South-Eastern Brazil and around Central Africa, there are areas of high

5   mortality overlapping with areas of high land-use and/or fire. Furthermore, we see in Europe and parts of China there are large mortality rates overlapping with high fractions of pastures.

If the carbon assimilate is similar to that of forests, to compensate for the lower observed coverage of woody PFTs the trees will posses higher mortality rates. Such compensation will be a non-direct approach to estimating mortality of disturbance prone areas. We see that in comparison to the distribution of mortalities increases in areas of observed land-use and fire (Figure 13),

10   however it does underestimate the annual burn rate.

4   Discussion

The response of land surface to climate change is a key uncertainty in climate projections. Ambitious climate targets also rely on land management practices such as reforestation and afforestation to increase the storage of carbon on land. First-generation Dynamic Global Vegetation Models (DGVMs) attempted to model the the land surface in terms of bulk properties such as mean

15   vegetation cover, vegetation carbon and leaf area index. These models lack information about the plant size-distribution, which compromised their ability to represent recovery from disturbance and the impact of land-management. Providing useful guidance on these issues requires improved DGVMs which can represent changes in tree size distributions within forests (so called 'demography').

A number of much more sophisticated second-generation DGVMs are now under development. These models often explicitly simulate the number of plants within different size or mass classes, and on different patches of land, which are defined by the time since a disturbance event. Such second generation models are therefore in principle able to simulate variations in plant number density as both a function of patch age and plant size. However, this completeness is at the expense of much

5   computational complexity. In principle, the number of age-defined patches grows indefinitely, and this can only be managed by arbitrarily merging patches of different ages after a certain age.

In this paper, we represent an intermediate complexity second generation DGVM ('RED'), which is designed to capture important features of plant demography, and yet avoid unwieldy computation. Our guiding principles in the development of RED have been that the model should: (i) simulate forest tree-size distributions; (ii) be globally applicable for ESM applications ;

10   (iii) be parameter sparse to minimise parameter uncertainties; (iv) be analytically soluble for steady-states to aid model initialisation. These design criteria, along with evaluation exercises against observed tree-size distributions in North America (Moore et al., 2018) and South America (Moore et al., 2019), has led us to make a number of simplifications.

Firstly, the number density for each PFT is treated as function of plant mass alone. This immediately eliminates the need to explicitly represent patches, and therefore removes age as an independent dimension. Secondly, we assume that plant growth-

15   rates vary as a power of plant mass. By default we assume a power of $\phi_g = 3/4$, which is consistent with Metabolic Scaling Theory (Enquist et al., 1998) and the empirically determined allometric relationships of (Niklas and Spatz, 2004). Finally, we assume that light-competition is only significant for the lowest 'seedling' mass class. This enables us to capture the impacts of light competition on seedling emergence through a simple 'gap' boundary condition.

These simplifications allow RED to be solved analytically for the steady state vegetation cover given information on the mor-

20   tality and growth rates per unit area for each PFT. Such analytical steady-state solutions mean that RED can be easily initialised in drift-free pre-industrial states, which is vital to avoid spurious sources and sinks in climate-carbon cycle projections. The analytical solutions also enable RED to be calibrated to the observed vegetation cover, via a single parameter (μ_0) which represents the ratio of mortality to growth for a tree of an arbitrary reference mass.

Aside from the existence of analytical stead-state solutions, RED is attractive for large-scale applications because it is both

25   parameter sparse ('parsimonious') and requires very few driving variables. The main driving variable is the time-varying net plant growth-rate for each PFT, which is defined as net primary production minus the local litter-fall. These driving data can be provided by a land-surface scheme, as we do in this study, or from observations. The only other driving variable for RED is the mortality rate, which we treat in this study as a PFT-specific constant that is independent of mass, but which could be dependent on plant mass and time to represent individual disturbance events (e.g. forest fires, disease outbreaks). Despite its

30   simplicity, the RED modelling is able to fit the global distribution of vegetation types (Figure 7), and simulates realistic succession including changes in forest demography (Figure 4).

There are inevitably weaknesses with any particular modelling approach. For RED, a current limitation is for competition to lead to a single PFT at each location within each co-competing vegetation class (i.e. tree, shrub, grass). The PFT with the highest equilibrium fraction will end-up excluding sub-dominant PFTs within the same vegetation class. It was necessary for

35   us to account for this eventual competitive exclusion to derive zero-drift steady states for the global runs presented in Section

3.1. Such competitive exclusion is a common problem in DGVMs (Fisher et al., 2018). Adapting the 'gap' boundary condition (Equation (12)) appears to be a promising way to deal with this issue in RED, without unduly increasing model complexity. We see this as a key priority for future research.

The existence of analytical steady-state solutions for RED also opens-up other promising research avenues. For example, these

5   solutions imply relationships between the fractional coverage of each PFT, total plant biomass, mean canopy height, and the ratio of mortality-to-growth. This in turn allows RED to be calibrated using observations of any two of these quantities (Figure 3). The analytical solutions also allow optimality hypotheses to be explored (e.g. the hypothesis that the fraction of net assimilate allocated to seed production maximises stand-density and/or biomass)

5   Conclusions

10   In this paper we have presented a new intermediate complexity second generation Dynamic Global Vegetation Model (DGVM), which captures important changes in forest demography. The *Robust Ecosystem Demography (RED)* model makes a number of important simplifications to achieve this. These simplifications are based-on theoretical ideas (e.g. metabolic scaling theory to estimate how plant growth-rate varies with plant mass, and perfect crown plasticity to minimise light competition) and also comparison to observed forest demography (Moore et al., 2018, 2019). As a result, RED is parameter sparse, and can be

15   driven with time-series of net plant growth rate (and optionally disturbance rates) for each Plant Functional Type (PFT). We have demonstrated that RED can be calibrated effectively to observed global vegetation maps, using a single fitting parameter (representing the ratio of mortality to growth for a plant of an arbitrary reference mass). The next stage will be to use RED in coupled climate-carbon cycle projections so to assess how changes in vegetation demography impact future CO_2 and climate. We have made the prototype RED code publically available, and we hope that Earth System and land-surface modellers will

20   make good use of this framework to further their own research.

Code availability. The RED model Python Code is archived at https://doi.org/10.5281/zenodo.3548678. Furthermore, RED is currently being coupled into JULES, where a basic integration currently exists as branch (vn5.4_veg3_ctrl) - this requires registration for the JULES repository (https://code.metoffice.gov.uk/trac).

Appendix A: Definitions

Table A1. List of Model Variables and Parameters

Symbol	Definitions	Units
Dimensions		
t	Time.	year
m	Carbon mass of an individual within a PFT.	kgC
ESM Inputs		
P	Total assimilate of Net Primary Productivity minus Local (Leaves, Wood and roots) Litterfall	kgC m^{-2} yr^{-1}
γ_d	Disturbance Mortality rate, the fraction of population dying over a year due explicitly modelled reason.	yr^{-1}
Individual		
m_0	Lowest/Sapling mass boundary.	kgC
g_0	(Sapling) Structural growth of an individual at the lowest mass boundary at a specific time.	kgC yr^{-1}
g	Structural growth of an individual at a given mass and time.	kgC yr^{-1}
h_0	(Sapling) Height of an individual at the lowest mass boundary.	m
h	(Height of an individual at a given mass.	m
a_0	(Sapling) Crown area of an individual at the lowest mass boundary.	m^2
a	Crown area of an individual at a given mass.	m^2
ϕ_g	Constant describing the power law scaling of structural growth across mass.	–
ϕ_h	Constant describing the power law scaling of height across mass.	–
ϕ_g	Constant describing the power law scaling of crown area across mass.	–
α	The fraction of total growth going into seedling recruitment.	–
Cohort		
n	Number density across mass space, the differential of N with respect to mass.	(kgC)$^{-1}$m^{-2}
N	Number density.	m^{-2}
G	Growth density.	kgC m^{-2} yr^{-1}
ν	The fractional coverage.	–
γ	Mortality rate, the summation of the baseline and additional mortalities across mass.	yr^{-1}
γ_b	Baseline Mortality rate, the fraction of population dying over a year due non-explicit reasons.	yr^{-1}
s	The fraction of avaliable space open for seedlings.	–
F	The flux of population density over time.	m^{-2}yr^{-1}
Λ_d	Demographic litter, the loss of carbon due to competition and mortality.	kgC m^{-2} yr^{-1}
M	Biomass density.	kgC m^{-2}
$c_{k,l}$	Competition coefficient, the fraction a PFT, k, is shaded by the canopy of PFT l.	–
Equilibrium		
μ_0	The boundary turnover parameter - the ratio of mass lost to growth gained in the boundary mass class.	–
λ_i	The proportional population of the ith class to the $i^\mathrm{th}-1$ class at equilibrium.	–
eq	Subscript denoting a variable in equilibrium.	–
Numerical		
k,l	Indices representing the PFT number.	–
i,j	Indices representing mass class number.	–
I	The largest mass class.	–
(k)	The current time-step.	–
ξ	The size scaling coefficient, mass classes are defined as $m_j = \xi\, m_{j-1}$, with $\xi > 1$.	–

Appendix B: RED and Demographic Equilibrium Theory

Equation (1), can be solved for the steady state if we assume metabolic scaling of growth using Eq. (2) and a size-independent mortality:

$$n = n_0 \left(\frac{m}{m_0} \right)^{-\phi_g} \exp \left\{ \frac{\mu_0}{(1 - \phi_g)} \left[1 - \left(\frac{m}{m_0} \right)^{1 - \phi_g} \right] \right\},$$

$$\mu_0 = \frac{\gamma m_0}{g_0}. \tag{B.1}$$

5   The scaling variable does not necessarily have to be biomass and can be any size-dimension with a power-scaling relationship with growth (height, basal diameter, etc.). The variable μ_0 can be thought of a parameter tied to the rate of biomass lost to biomass gained. The larger μ_0 is the greater the associated cost of replacing lost biomass - the smaller the total population density. Where n_0 is a boundary condition that describes the number density at the mass m_0.

Integrating Eq. (B.1) from m_0 to ∞ gives us estimates for the total number density:

10   $$N_{\text{eq}} = \frac{n_0 g_0}{\gamma} = \frac{n_0 m_0}{\mu_0}. \tag{B.2}$$

We can also gain estimates of the total growth and biomass values by integrating with the allometric relationships:

$$G_{\text{eq}} = g_0 N_{\text{eq}} \left(\frac{\mu_0}{1 - \phi_g} \right)^{\frac{1}{\phi_g - 1}}$$

$$\times \exp \left\{ \frac{\mu_0}{1 - \phi_g} \right\} \Gamma \left(\frac{1}{1 - \phi_g} + 1, \frac{\mu_0}{1 - \phi_g} \right), \tag{B.3}$$

the total biomass:

$$M_{\text{eq}} = m_0 N_{\text{eq}} \left(\frac{\mu_0}{1 - \phi_g} \right)^{\frac{1}{\phi_g - 1}}$$

$$\times \exp \left\{ \frac{\mu_0}{1 - \phi_g} \right\} \Gamma \left(\frac{1}{1 - \phi_g} + 1, \frac{\mu_0}{1 - \phi_g} \right), \tag{B.4}$$

15   and the total vegetation cover:

$$\nu_{\text{eq}} = a_0 N_{\text{eq}} \left(\frac{\mu_0}{1 - \phi_g} \right)^{\frac{1}{\phi_g - 1}}$$

$$\times \exp \left\{ \frac{\mu_0}{1 - \phi_g} \right\} \Gamma \left(\frac{\phi_a}{1 - \phi_g} + 1, \frac{\mu_0}{1 - \phi_g} \right). \tag{B.5}$$

Where $\Gamma(a, b)$ is the incomplete upper gamma function. When we assume the constants presented in (Niklas and Spatz, 2004) - $\phi_g = \frac{3}{4}, \phi_a = \frac{1}{3}$ simplifies towards:

$$G_{\text{eq}} = g_0 N_{\text{eq}} \left(1 + \frac{3}{4\mu_0} + \frac{3}{8\mu_0^2} + \frac{3}{32\mu_0^3} \right), \tag{B.6}$$

$$M_{\text{eq}} = m_0 N_{\text{eq}} \left(1 + \frac{1}{\mu_0} + \frac{3}{4\mu_0^2} + \frac{3}{8\mu_0^3} + \frac{3}{32\mu_0^4} \right), \tag{B.7}$$

$$\nu_{\text{eq}} = a_0 N_{\text{eq}} \left(1 + \frac{1}{2\mu_0} + \frac{1}{8\mu_0^2}\right). \tag{B.8}$$

For converting a μ_0 found using total dry mass ($\mu_{0,\text{tdm}}$) of 1 kg to that of dry carbon mass:

$$\mu_0 = 2^{1-\phi_g}\mu_{0,\text{tdm}}. \tag{B.9}$$

B1 Closed Continous Form

Using Eq. (B.2) and Eq. (B.6) with the competitive constraint, we find that the equilibrium fraction is given by:

$$\nu_{\text{eq}} = 1 - \left(\frac{1-\alpha}{\alpha}\right) \frac{\mu_0}{1 + \frac{3}{4\mu_0} + \frac{3}{8\mu_0^2} + \frac{3}{32\mu_0^3}}. \tag{B.10}$$

We can rearrange Eq. (B.10) into Eq. (B.5) allowing for the substitution of ν_{eq} into the equilibrium solutions (Equation (B.6) and Eq.(B.7)). For instance, the exact solution for the total number density is given as:

$$N_{\text{eq}} = \frac{\nu_{\text{eq}}}{a_0} \left(\frac{1}{1 + \frac{1}{2\mu_0} + \frac{1}{8\mu_0^2}}\right), \tag{B.11}$$

enables the calculation of total growth to Eq.(B.6)) as:

$$G_{\text{eq}} = \frac{\nu_{\text{eq}}g_0}{a_0(1-\alpha)} \left(\frac{1 + \frac{3}{4\mu_0} + \frac{3}{8\mu_0^2} + \frac{3}{32\mu_0^3}}{1 + \frac{1}{2\mu_0} + \frac{1}{8\mu_0^2}}\right), \tag{B.12}$$

and the biomass:

$$M_{\text{eq}} = \frac{\nu_{\text{eq}}m_0}{a_0} \left(\frac{1 + \frac{1}{\mu_0} + \frac{3}{4\mu_0^2} + \frac{3}{8\mu_0^3} + \frac{3}{32\mu_0^4}}{1 + \frac{1}{2\mu_0} + \frac{1}{8\mu_0^2}}\right). \tag{B.13}$$

B2 Continuous-Discrete Convergence

There are a few differences between the numerical steady state and the continuous form of RED. Firstly, the truncation point of the top mass class results in a underestimation of the total coverage, biomass and number density for an identical μ_0. The second source of difference arises from the binning of the mass classes. Discretising results in the continuous scaling of the growth and mass is not fully captured. In the current scheme a number density between the masses m_j of m_{j+1} will have its physiological characteristics represented at the m_j mass, this can lead to underestimating the total growth/biomass/coverage within the class. This is demonstrated in figure B1.a, where the total coverage of the biomass is lower in the discrete model than the continuous solution. There is also convergence when the bins between the methods when the number of classes goes towards infinity ($I \rightarrow \infty$) and class widths goes towards zero ($\xi \rightarrow 1$).

From Figure B1.c, we see that there is a clear optimum amount of scaling for a given number of classes which minimise the difference between continuous and discrete. This can be found by taking the difference of the continuous and discrete
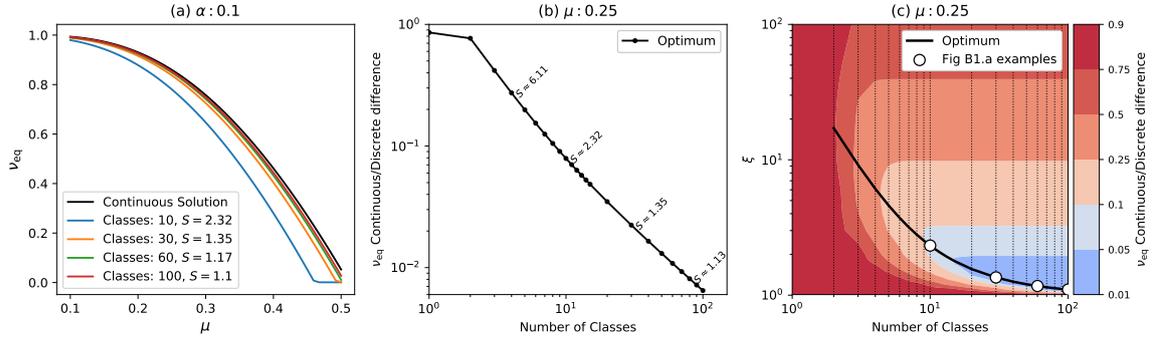
Figure B1. Showing the convergence of the numerical model towards the analytical equilibrium for Broadleaf Evergreen Tropical (BET-Tr) for coverage. With $\phi_g = 0.75, \alpha = 0.1$. (a) demonstrates over given μ_0 values the convergence. (b-c) shows that the diffences is minimised when the bin scaling, $\xi \to 1$, while the number of mass classes goes towards infinity. The black lines on (b,c) denote the optimum from Eq. (B.16). The white dots on (c) correspond with the discrete lines on (a).

coverages and differentiating with respect to ξ to find the minima. It should be noted that the as the continuous form is not dependent on ξ, we get:

$$\frac{\partial}{\partial \xi} \left[\nu_{\text{eq,continuous}} - \nu_{\text{eq}} \right] = -\frac{\partial}{\partial \xi} \left[\nu_{\text{eq}} \right]. \tag{B.14}$$

Where ν_{eq} corresponds with discrete equilibrium (Equation (31), with $\nu_{\text{eq}} = (1 - s)$). Setting Eq. (B.14) equal to zero we

5 reduce the relationship to just a dependence of X_N and X_G:

$$0 = \frac{\partial}{\partial \xi} \left[\frac{X_N}{X_G} \right] = X_G X_N' - X_G' X_N. \tag{B.15}$$

Finding the partial derivative of X_N, we get:

$$X_N' = \sum_{j=1}^{I} \left[\left(\prod_{i=1}^{j} \lambda_i \right) \left(\sum_{i=1}^{j} \frac{\lambda_i'}{\lambda_i} \right) \right], \tag{B.16}$$

and for X_G:

10 $$X_G' = \sum_{j=1}^{I} \left[\xi^{j\phi_g} \left(\prod_{i=1}^{j} \lambda_i \right) \left(j\phi_g \xi^{-1} + \sum_{i=1}^{j} \frac{\lambda_i'}{\lambda_i} \right) \right] \quad . \tag{B.17}$$

Finding λ_i' we get:

$$\begin{aligned} \lambda_i' =& \lambda_i \Big[(1 - i)(\phi_g - 1)\xi^{-1} \\ &- \lambda_i \left(i(\phi_g - 1)\xi^{\phi_g - 2} + \mu_0 \xi^{(i-1)(1-\phi_g)} \right) \Big], \end{aligned} \tag{B.18}$$

and for λ_i':

$$\lambda_i' = \left(\frac{(1 - \xi^{-1})(I - 1)(\phi_g - 1) - 1}{\xi - 1} \right) \lambda_i. \tag{B.19}$$

To numerically solve for the minimum, we must differentiate Eq. (B.15), with respect to ξ. Through the product rule we get:

$$\frac{\partial^2}{\partial \xi^2}\left[\frac{X_N}{X_G}\right] = X_G X_N'' - X_G'' X_N, \tag{B.20}$$

Eq. (B.16) differentiated simplifies towards:

$$X_N'' = \sum_{j=1}^{I}\left[\left(\prod_{i=1}^{j}\lambda_i\right)\left(\sum_{i=1}^{j}\frac{\lambda_i''}{\lambda_i}\right)\right], \tag{B.21}$$

5 and Eq. (B.17):

$$X_G'' = \sum_{j=1}^{I}\left[\xi^{j\phi_g}\left(\prod_{i=1}^{j}\lambda_i\right)\left(j\phi_g\xi^{-2}(j\phi_g-1)\right.\right.$$
$$\left.\left. + \sum_{i=1}^{j}\frac{2j\phi_g\xi^{-1}\lambda_i'-\lambda_i''}{\lambda_i}\right)\right]\qquad . \tag{B.22}$$

λ_i'' is given as:

$$\lambda_i'' = \lambda_i\left[-\frac{\lambda_i'}{\lambda_i}\left((i-1)(\phi_g-1)\xi^{-1}\right)\right.$$
$$-(i-1)(\phi_g-1)\xi^{-2}$$
$$-\lambda_i(\phi_g-1)\xi^{-1}\left(i(\phi_g-1)\xi^{\phi_g-2}\right.$$
$$\left.\left.-\mu_0(i-1)\xi^{(i-1)(1-\phi_g)}\right)\right]. \tag{B.23}$$

For the double differential of λ_i we get:

$$\lambda_i'' = \frac{\lambda_i''^2}{\lambda_i} + \frac{\lambda_i}{\xi-1}$$
$$\times\left(\frac{(I-1)(\phi-1)}{\xi^2} - \frac{\lambda_i'}{\lambda_i}\right). \tag{B.24}$$

10 We now possess the identities needed to perform a numerical root finding algorithm for the optimum bin scaling for a given class. Using a Newton root finding method for Eq. (B.15) with it's differential; Eq. (B.20), we find the optimum. On figure B1 the optimum line is shown as the bright dashed black line.

Author contributions. Originally the model framework in JRM's thesis (Moore, 2016) under the supervision of PMC and CH. The description

15 of PFT competition, the numerical model and the equilibrium solutions has been further developed by APKA, JRM and PMC. Currently RED is being integrated into JULES with the supervision of AJW and CJ. AJW also provided and processed the UKESM growth rates needed to drive RED globally within this paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work has been funded as part of the Newton Fund with the Met Office Climate Science for Service Partnership Brazil (CSSP-Brazil) and by the European Research Council (ERC) ECCLES project. A.A was funded through the CASE studentship from the University of Exeter and the Met Office. We are grateful to the Met Office for considering implementation in JULES, via a ticket 902 within a branch of the code repository.

References

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, Geoscientific Model Development, 4, 677–699, https://doi.org/10.5194/gmd-4-677-2011, https://www.geosci-model-dev.net/4/677/2011/, 2011.

Bond, W. J., Woodward, F. I., and Midgley, G. F.: The global distribution of ecosystems in a world without fire, New phytologist, 165, 525–538, 2005.

Brovkin, V., Boysen, L., Arora, V. K., Boisier, J. P., Cadule, P., Chini, L., Claussen, M., Friedlingstein, P., Gayler, V., Van den hurk, B. J., Hurtt, G. C., Jones, C. D., Kato, E., De noblet ducoudre, N., Pacifico, F., Pongratz, J., and Weiss, M.: Effect of anthropogenic land-use and land-cover changes on climate and land carbon storage in CMIP5 projections for the twenty-first century, Journal of Climate, 26, 6859–6881, https://doi.org/10.1175/JCLI-D-12-00623.1, 2013.

Burton, C., Betts, R., Cardoso, M., Feldpausch, T. R., Harper, A., Jones, C. D., Kelley, D. I., Robertson, E., and Wiltshire, A.: Representation of fire, land-use change and vegetation dynamics in the Joint UK Land Environment Simulator vn4. 9 (JULES), Geoscientific Model Development, 12, 179–193, 2019.

Chuvieco, E., Pettinari, M., Lizundia Loiola, J., Storm, T., and Padilla Parellada, M.: ESA Fire Climate Change Initiative (Fire_cci): MODIS Fire_cci Burned Area Grid product, version 5.1., doi:10.5285/3628cb2fdba443588155e15dee8e5352, 2019.

Cox, P. M.: Description of the" TRIFFID" dynamic global vegetation model, 2001.

Davies-Barnard, T., Valdes, P., Singarayer, J. S., Wiltshire, A., and Jones, C.: Quantifying the relative importance of land cover change from climate and land use in the representative concentration pathways, Global Biogeochemical Cycles, 29, 842–853, 2015.

Enquist, B. J., Brown, J. H., and West, G. B.: Allometric scaling of plant energetics and population density, Nature, 395, 163, 1998.

Fisher, R., McDowell, N., Purves, D., Moorcroft, P., Sitch, S., Cox, P., Huntingford, C., Meir, P., and Woodward, F. I.: Assessing uncertainties in a second-generation dynamic vegetation model caused by ecological scale limitations, New Phytologist, 187, 666–681, 2010.

Fisher, R. A., Koven, C. D., Anderegg, W. R., Christoffersen, B. O., Dietze, M. C., Farrior, C. E., Holm, J. A., Hurtt, G. C., Knox, R. G., Lawrence, P. J., Lichstein, J. W., Longo, M., Matheny, A. M., Medvigy, D., Muller-Landau, H. C., Powell, T. L., Serbin, S. P., Sato, H., Shuman, J. K., Smith, B., Trugman, A. T., Viskari, T., Verbeeck, H., Weng, E., Xu, C., Xu, X., Zhang, T., and Moorcroft, P. R.: Vegetation demographics in Earth System Models: A review of progress and priorities, Global Change Biology, 24, 35–54, https://doi.org/10.1111/gcb.13910, 2018.

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., and Knutti, R.: Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks, Journal of Climate, 27, 511–526, https://doi.org/10.1175/JCLI-D-12-00579.1, 2014.

Harper, A. B., Cox, P. M., Friedlingstein, P., Wiltshire, A. J., Jones, C. D., Sitch, S., Mercado, L. M., Groenendijk, M., Robertson, E., Kattge, J., Bönisch, G., Atkin, O. K., Bahn, M., Cornelissen, J., Niinemets, Ü., Onipchenko, V., Peñuelas, J., Poorter, L., Reich, P. B., Soudzilovskaia, N. A., and Van Bodegom, P.: Improved representation of plant functional types and physiology in the Joint UK Land Environment Simulator (JULES v4.2) using plant trait information, Geoscientific Model Development, 9, 2415–2440, https://doi.org/10.5194/gmd-9-2415-2016, 2016.

Harper, A. B., Wiltshire, A. J., Cox, P. M., Friedlingstein, P., Jones, C. D., Mercado, L. M., Sitch, S., Williams, K., and Duran-Rojas, C.: Vegetation distribution and terrestrial carbon cycle in a carbon cycle configuration of JULES4.6 with new plant functional types, Geoscientific Model Development, 11, 2857–2873, https://doi.org/10.5194/gmd-11-2857-2018, 2018.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., et al.: Global carbon budget 2018, Earth System Science Data (Online), 10, 2018.

Li, W., MacBean, N., Ciais, P., Defourny, P., Lamarche, C., Bontemps, S., and Peng, S.: Derivation of plant functional type (PFT) maps from the ESA CCI Land Cover product [Data set]., http://doi.org/10.5281/zenodo.1048163, 2019.

5 Lugo, A. E. and Scatena, F. N.: Background and catastrophic tree mortality in tropical moist, wet, and rain forests, Biotropica, pp. 585–599, 1996.

Medvigy, D., Wofsy, S. C., Munger, J. W., Hollinger, D. Y., and Moorcroft, P. R.: Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2, Journal of Geophysical Research: Biogeosciences, 114, 1–21, https://doi.org/10.1029/2008JG000812, 2009.

10 Moorcroft, P. R., Hurtt, G., and Pacala, S. W.: a Method for Scaling Vegetation Dynamics: the Ecosystem Demography Model (Ed), Ecological Monographs, 71, 557–586, https://doi.org/doi:10.1890/0012-9615(2001)071[0557:AMFSVD]2.0.CO;2, https://doi.org/10.1890/0012-9615(2001)071[0557:AMFSVD]2.0.CO{%}0Ahttp://0.0.0.2, 2001.

Moore, J. R.: Aspects of Land Surface Modelling: Role of Biodiversity in Ecosystem Resilience to Environmental Change and a Robust Ecosystem Demography Model, Ph.D. thesis, 2016.

15 Moore, J. R., Zhu, K., Huntingford, C., and Cox, P. M.: Equilibrium forest demography explains the distribution of tree sizes across North America, Environmental Research Letters, 13, https://doi.org/10.1088/1748-9326/aad6d1, 2018.

Moore, J. R., Argles, A. P. K., Zhu, K., Huntingford, C., and Cox, P. M.: Validation of demographic equilibrium theory against tree-size distributions and biomass density in Amazonia, Biogeosciences Discussions, 2019, 1–41, https://doi.org/10.5194/bg-2019-262, https://www.biogeosciences-discuss.net/bg-2019-262/, 2019.

20 Niklas, K. J. and Spatz, H.-C.: From The Cover: Growth and hydraulic (not mechanical) constraints govern the scaling of tree height and mass, Proceedings of the National Academy of Sciences, 101, 15 661–15 663, https://doi.org/10.1073/pnas.0405857101, http://www.pnas.org/cgi/doi/10.1073/pnas.0405857101, 2004.

Phillips, O. L.: Long-term environmental change in tropical forests: increasing tree turnover, Environmental Conservation, 23, 235–248, https://doi.org/10.1017/s0376892900038856, 1996.

25 Phillips, O. L., Baker, T. R., Arroyo, L., Higuchi, N., Killeen, T. J., Laurance, W. F., Lewis, S. L., Lloyd, J., Malhi, Y., Monteagudo, A., Neill, D. A., Núñez Vargas, P., Silva, J. N., Terborgh, J., Vásquez Martínez, R., Alexiades, M., Almeida, S., Brown, S., Chave, J., Comiskey, J. A., Czimczik, C. I., Di Fiore, A., Erwin, T., Kuebler, C., Laurance, S. G., Nascimento, H. E., Olivier, J., Palacios, W., Patiño, S., Pitman, N. C., Quesada, C. A., Saldias, M., Torres Lezama, A., and Vinceti, B.: Pattern and process in Amazon tree turnover, 1976-2001, Philosophical Transactions of the Royal Society B: Biological Sciences, 359, 381–407, https://doi.org/10.1098/rstb.2003.1438, 2004.

30 Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., Hagemann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Ottlé, C., Peters, M., and Peylin, P.: Plant functional type classification for earth system models: Results from the European Space Agency's Land Cover Climate Change Initiative, Geoscientific Model Development, 8, 2315–2328, https://doi.org/10.5194/gmd-8-2315-2015, 2015.

Pugh, T., Jones, C., Huntingford, C., Burton, C., Arneth, A., Brovkin, V., Ciais, P., Lomas, M., Robertson, E., Piao, S., et al.: A Large

35 Committed Long-Term Sink of Carbon due to Vegetation Dynamics, Earth's Future, 6, 1413–1432, 2018.

Sato, H., Itoh, A., and Kohyama, T.: SEIB-DGVM: A new Dynamic Global Vegetation Model using a spatially explicit individual-based approach, Ecological Modelling, 200, 279–307, https://doi.org/10.1016/j.ecolmodel.2006.09.006, 2007.

Scheiter, S., Langan, L., and Higgins, S. I.: Next-generation dynamic global vegetation models: Learning from community ecology, New Phytologist, 198, 957–969, https://doi.org/10.1111/nph.12210, 2013.

Sellar, A., Jones, C., Mulcahy, J., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., , and Others: UKESM1: Description and evaluation of the UK Earth system model, Journal of Advances in Modeling Earth Systems, in prep.

5 Strigul, N., Pristinski, D., Purves, D., Dushoff, J., and Pacala, S.: Scaling from trees to forests: tractable macroscopic equations for forest dynamics, Ecological Monographs, 78, 523–545, 2008.

Thomas, S. C. and Martin, A. R.: Carbon content of tree tissues: a synthesis, Forests, 3, 332–352, 2012.