

CLM(ED) model: Technical Documentation

ROSIE FISHER¹, STEFAN MUSZALA¹, ALLAN SPESSA², MARIANA VERTEINSTEIN¹, CHONGGANG XU³, ERIK KLUZEK¹, CHARLES KOVEN⁴, RYAN KNOX⁴, JENNIFER HOLM⁴, BRENDAN M. ROGERS⁵, DAVID LAWRENCE¹ & GORDON BONAN¹

¹*The National Center for Atmospheric Research, Boulder, Colorado*

²*Department of Meteorology, The Open University, Milton Keynes, UK.*

³*Los Alamos National Laboratory, Los Alamos, New Mexico.*

⁴*Lawrence Berkeley National Laboratory, Berkeley, California.*

⁵*Woods Hole Research Center, Falmouth, Massachusetts.*

Contents

1	Acknowledgements	2
2	Introduction	2
3	The representation of ecosystem heterogeneity in the CLM(ED)	3
4	Initialization of vegetation from bare ground	9
5	Canopy Structure and the Perfect Plasticity Approximation	11
6	Radiation Transfer	15
7	Photosynthesis	19
8	Plant respiration	23
9	Stomatal Conductance	25
10	Allocation	25

11 Control of Leaf Area Index	29
12 Phenology	30
13 Seed Dynamics and Recruitment	31
14 Fragmenting Litter Pool Production and Turnover	32
15 Plant Mortality	34
16 Fire	35

1 Acknowledgements

The development of the ideas in this document has benefitted from conversation with many people, including WILLIAM BOND, PETER COX, MATTHEW FOREST, ANDREW FOX, WILLIAM HOFFMAN, PETER LAWRENCE, JEREMY LICHSTEIN, HEIKE LISCHKE, MARK LOMAS, NATHAN McDOWELL, PAUL MOORCROFT, COLIN PRENTICE, DREW PURVES, NAOMI TAGUE, BENJAMIN SANDERSON, STEPHEN SITCH, ABIGAIL SWANN, ANTHONY WALKER, MATHEW WILLIAMS, F.IAN WOODWARD & SÖNKE ZAEHLE.

The code included in the CLM(ED) was initially developed as part of the QUEST program, funded by the UK Natural Environment Research Council. Subsequently, it's development continued as an activity of the CLM community, at Los Alamos National Laboratory, and at NCAR, funded by the National Science Foundation and the Department of Energy. The Ecosystem Demography concept was developed as a continuation of the original ED1.0 code structure [19] and was translated from the original 'C' code into Fortran90 for incorporation into land surface models. The coupling of the ED and SPITFIRE models was also funded under the QUEST program, in collaboration with Allan Spessa. Significant software design, engineering, and testing was completed on the CLM(ED) code in 2014 by Stefan Muszala and Mariana Vertenstein at NCAR, funded by the National Science Foundation.

2 Introduction

The CLM(ED) model is presented here as an option within the structure of the Community Land Model (CLM). Ecosystem Demography ('ED'), is a concept derived from the work of [19] and is a cohort model of vegetation competition and co-existence, allowing a representation of the biosphere which accounts for the division of the land surface into successional stages, and the competition for light between height structured cohorts of representative trees of various plant functional types. This implementation of the Ecosystem Demography concept links the surface flux and canopy physiology concepts in the CLM with numerous additional developments necessary to accommodate the new model also documented here. These include a version of the SPITFIRE (Spread and InTensity of Fire) model of [32], and an adoption of the concept of 'perfect plasticity' approach of [24, 13, 35] in accounting for the spatial arrangement of crowns. Novel algorithms accounting for the fragmentation

of coarse woody debris into chemical litter streams, for the physiological optimisation of canopy thickness, for the accumulation of seeds in the seed bank, for multi-layer multi-PFT radiation transfer, for drought-deciduous and cold-deciduous phenology, for carbon storage allocation, and for tree mortality under carbon stress, are also included and presented here.

2.1 This Document

This document describes the implementation of the Ecosystem Demography concept within the Community Land Model. It is intended as a supplementary document to the main CLM technical note, which is at the time of writing, the CLM4.5 technical note [21]. The material covered describes how the Ecosystem Demography code is organized, how it interacts with the pre-existing CLM routines, and the new aspects of plant physiology and vegetation dynamics that are introduced further to those already existing within the model. Numerous other implementations of the Ecosystem Demography concept are in existence, in particular the ED2 model, which is a fully operational land surface scheme [18] which contains both the basic ED concept and an array of physiological and ecological innovations surrounding the basic concept. Therefore, to avoid confusion between the concept of ‘Ecosystem Demography’ and the implementation of this concept in different models, we refer our model as the ‘CLM(ED)’ throughout.

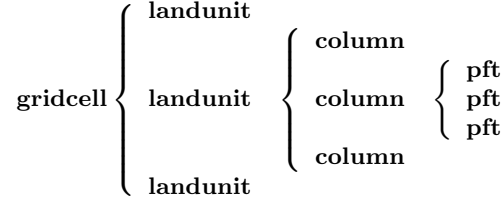
3 The representation of ecosystem heterogeneity in the CLM(ED)

The land surface of the Earth is heterogeneous for many reasons, driven by variations in climate, edaphic history, ecological variability, geological forcing and human interventions. Land surface models represent this variability first by introducing a grid structure to the land surface, allowing different atmospheric forcings to operate in each grid cell, and subsequently by representing ‘sub-grid’ variability in the surface properties. In the CLM, the land surface is divided into numerous ‘landunits’ corresponding to the underlying condition of the surface (e.g. soils, ice, lakes, bare ground) and then ‘columns’ referring to elements of the surface that share below ground resources (water & nutrients). Within the ‘soil’ landunit, for example, there are separate columns for crops, and for natural vegetation, as these are assumed to use separate resource pools. The CLM(ED) model at present only operates on the naturally vegetated column. The ‘soil’ column is sub-divided into numerous tiles, that correspond to statistical fractions of the potentially vegetated land area. In the CLM 4.5 (and all previous versions of the model), sub-grid tiling operates on the basis of plant functional types (PFTs). That is, each piece of land is assumed to be occupied by only one plant functional type, with multiple PFT-specific tiles sharing a common soil water and nutrient pool. This PFT-based tiling structure is the standard method used by most land surface models deployed in climate prediction.

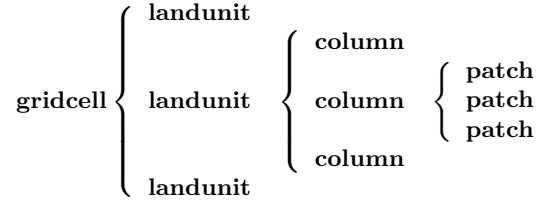
The introduction of the Ecosystem Demography concept introduces significant alterations to the representation of the land surface in the CLM. In the CLM(ED), the tiling structure represents the disturbance history of the ecosystem. Thus, some fraction of the land surface is characterized as ‘recently disturbed’, some fraction has escaped disturbance for a long time, and other areas will have intermediate disturbances. Thus the ED concept essentially discretizes the trajectory of succession from disturbed ground to ‘mature’ ecosystems. Within the ED code, each ‘disturbance history class’ is referred to as a ‘patch’. The word ‘patch’ has many possible interpretations, so it is important to note that: **there is no spatial location associated with the concept of a ‘patch’.** It refers to a fraction of the potential vegetated area consisting of all parts of the ecosystem with similar disturbance history.

The ‘patch’ organizational structure in CLM thus replaces the previous ‘PFT’ structure in the organization hierarchy. The original hierarchical land surface organizational structure of CLM as described in [21] may be

depicted as:



and the new structure is altered to the following:



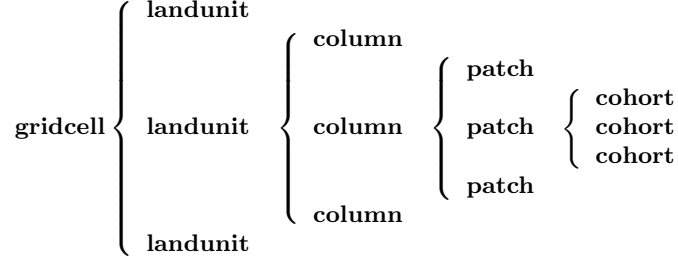
Thus, each gridcell becomes a matrix of ‘patches’ that are conceptualized by their ‘age since disturbance’ in years. This is the equivalent of grouping together all those areas of a gridcell that are ‘canopy gaps’, into a single entity, and all those areas that are ‘mature forest’ into a single entity.

3.1 Cohortized representation of tree populations

Each common-disturbance-history patch is a notional ecosystem that might in reality contain numerous individual plants which vary in their physiological attributes, in height and in spatial position. One way of addressing this heterogeneity is to simulate a forest of specific individuals, and to monitor their behavior through time. This is the approach taken by ‘gap’ and individual-based [31, 26, 33, 9]. The depiction of individuals typically implies that the outcome of the model is stochastic. This is because we lack the necessary detailed knowledge to simulate the individual plant’s fates. Thus gap models imply both stochastic locations and mortality of plants. Thus, (with a genuinely random seed) each model outcome is different, and an ensemble of model runs is required to generate an average representative solution. Because the random death of large individual trees can cause significant deviations from the mean trajectory for a small plot (a typical simulated plot size is 30m x 30 m) the number of runs required to minimize these deviations is large and computationally expensive. For this reason, models that resolve individual trees typically use a physiological timestep of one day or longer (e.g. [31, 36, 26]).

The approach introduced by the Ecosystem Demography model [19] is to group the hypothetical population of plants into ‘cohorts’. In the notional ecosystem, after the land-surface is divided into common-disturbance-history patches, the population in each patch is divided first into plant functional types (the standard approach to representing plant diversity in large scale vegetation models), and then each plant type is represented as numerous height classes. Importantly, **for each PFT/height class bin, we model *one* representative individual plant, which tracks the average properties of this ‘cohort’ of individual plants.** Thus, each common-disturbance-history patch is typically occupied by a set of cohorts of different plant functional types, and different height classes within those plant functional types. Each cohort is associated with a number of identical trees, n_{coh} (where *coh* denotes the identification or index number for a given cohort)..

The complete hierarchy of elements in the CLM(ED) is therefore now described as follows:



3.2 Discretization of cohorts and patches

Newly disturbed land and newly recruited seedlings can in theory be generated at each new model timestep as the result of germination and disturbance processes. If the new patches and cohorts established at *every* timestep were tracked by the model structure, the computational load would of course be extremely high (and thus equivalent to an individual-based approach). A signature feature of the ED model is the system by which ‘functionally equivalent’ patches and cohorts are fused into single model entities to save memory and computational time.

¹ This functionality requires that criteria are established for the meaning of ‘functional equivalence’, which are by necessity slightly subjective, as they represent ways of abstracting reality into a more tractable mathematical representation. As an example of this, for height-structured cohorts, we calculate the relativized differences in height (h_{coh} , m) between two cohorts of the same pft, p and q as

$$d_{\text{hite},p,q} = \frac{\text{abs.}(h_p - h_q)}{\frac{1}{2}(h_p + h_q)} \quad (1)$$

If $d_{\text{hite},p,q}$ is smaller than some threshold t_{ch} , and they are of the same plant functional type, the two cohorts are considered equivalent and merged to form a third cohort r , with the properties of cohort p and q averaged such that they conserve mass. The model parameter t_{ch} can be adjusted to adjust the trade-off between simulation accuracy and computational load. There is no theoretical optimal value for this threshold but it may be altered to have finer or coarser model resolutions as needed.

² Similarly, for common-disturbance-history patches, we again assign a threshold criteria, which is then compared to the difference between patches m and n , and if the difference is less than some threshold value (t_p) then patches are merged together, otherwise they are kept separate. However, in contrast with height-structured cohorts, where the meaning of the difference criteria is relatively clear, how the landscape should be divided into common-disturbance-history units is less clear. Several alternative criteria are possible, including Leaf Area Index, total biomass and total stem basal area.

In this implementation of the CLM(ED) we assess the amount of above-ground biomass in each PFT/plant diameter bin. Biomass is first grouped into fixed diameter bins for each PFT (ft) and a significant difference in any bin will cause patches to remain separated. This means that if two patches have similar total biomass, but differ in the distribution of that biomass between diameter classes or plant types, they remain as separate entities. Thus

$$B_{\text{profile},m,dc,ft} = \sum_{d_{c,\min}}^{d_{c,\max}} (B_{\text{ag},\text{coh}} n_{\text{coh}}) \quad (2)$$

$B_{\text{profile},m,dc,ft}$ is the binned above-ground biomass profile for patch m , d_c is the diameter class. $d_{c,\min}$ and $d_{c,\max}$ are the lower and upper boundaries for the d_c diameter class. $B_{\text{ag},\text{coh}}$ and n_{coh} depict the biomass (KgC m^{-2})

¹This description covers algorithms in the ‘fuse_cohorts’ subroutine.

²This description covers algorithms in the ‘fuse_patches’ subroutine.

and the number of individuals of each cohort respectively. A difference matrix between patches m and n is thus calculated as

$$d_{\text{biomass},mn,dc,ft} = \frac{\text{abs}(B_{\text{profile},m,hc,ft} - B_{\text{profile},n,hc,ft})}{\frac{1}{2}(B_{\text{profile},m,hc,ft} + B_{\text{profile},n,hc,ft})} \quad (3)$$

If all the values of $d_{\text{biomass},mn,hc,ft}$ are smaller than the threshold, t_p , then the patches m and n are fused together to form a new patch o .

To increase computational efficiency and to simplify the coding structure of the model, the maximum number of patches is capped at $P_{\text{no},\text{max}}$. To force the fusion of patches down to this number, the simulation begins with a relatively sensitive discretization of patches ($t_p = 0.2$) but if the patch number exceeds the maximum, the fusion routine is repeated iteratively until the two most similar patches reach their fusion threshold. This approach maintains an even discretization along the biomass gradient, in contrast to, for example, simply fusing the oldest or youngest patches together.

³ The area of the new patch ($A_{\text{patch},o}$, m^2) is the sum of the area of the two existing patches,

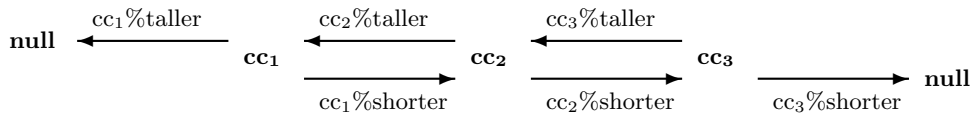
$$A_{\text{patch},o} = A_{\text{patch},n} + A_{\text{patch},m} \quad (4)$$

and the cohorts ‘belonging’ to patches m and n now co-occupy patch o . The state properties of m and n (litter, seed pools, etc.) are also averaged in accordance with mass conservation .

3.3 Linked Lists: the general code structure of CLM(ED)

⁴ The number of patches in each natural vegetation column and the number of cohorts in any given patch are variable through time because they are re-calculated for each daily timestep of the model. The more complex an ecosystem, the larger the number of patches and cohorts. For a slowly growing ecosystem, where maximum cohort size achieved between disturbance intervals is low, the number of cohorts is also low. For fast-growing ecosystems where many plant types are viable and maximum heights are large, more cohorts are required to represent the ecosystem with adequate complexity.

In terms of variable structure, the creation of an array whose size could accommodate every possible cohort would mean defining the maximum potential number of cohorts for every potential patch, which would result in very large amounts of wasted allocated memory, on account of the heterogeneity in the number of cohorts between complex and simple ecosystems (n.b. this does still happen for some variables at restart timesteps). To resolve this, the cohort structure in the CLM(ED) model does not use an array system for internal calculations. Instead it uses a system of *linked lists* where each cohort structure is linked to the cohorts larger than and smaller than itself using a system of pointers. The shortest cohort in each patch has a ‘shorter’ pointer that points to the *null* value, and the tallest cohort has a ‘taller’ pointer that points to the null value. The cohort structure which is the ‘currentCohort’ is typically the object name used, and the data values associated with that structure are accessed using the ‘%’ symbol. In this example, currentCohort is abbreviated to ‘cc’. cc_1 is the tallest cohort and cc_3 is the shortest. Thus:



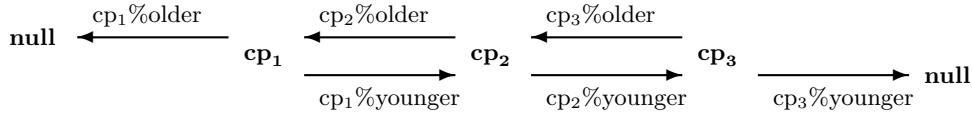
³This description covers algorithms in the ‘fuse_2_patches’ subroutine.

⁴This description covers the structure of code in all modules in clm4.5 that are located in ‘ED’ subdirectories

Instead of iterating along a vector indexed by *coh*, the code structures typically begin at the tallest cohort in a given patch, and iterate until a null pointer is encountered. An example of this, for a linked-list loop calculating total cohort biomass (*currentCohort%b*) from the structural and live biomass pools (*currentCohort%balive*, *currentCohort%bstruc*) is shown below

```
currentCohort => currentPatch%tallest
do while(associated(currentCohort))
  currentCohort%b = currentCohort%balive + cc%bstruc
  currentCohort => currentCohort%shorter
end do
```

Using this structure, it is therefore possible to have an unbounded upper limit on cohort number, and also to easily alter the ordering of cohorts if, for example, a cohort of one functional type begins to grow faster than a competitor of another functional type, and the cohort list can easily be re-ordered by altering the pointer structure. Each cohort has pointers indicating to which patch and gridcell it belongs. The patch system is analogous to the cohort system, except that patches (here denoted using the ‘currentPatch’ (cp) identifier) are ordered in terms of their relative age, with pointers to older and younger patches where *cp₁* is the oldest:



thus a loop around the patch structures typically takes this form

```
currentPatch => currentSite%oldestpatch
do while(associated(currentPatch))
  currentPatch%age = currentPatch%age +  $\delta_t$ 
  currentPatch => currentPatch%younger
end do
```

For this example, we increase the age stamp on each patch by δ_t , which is the increase in patch age with each timestep (days).

3.4 Indices used in CLM(ED)

Some of the indices used in the CLM(ED) are similar to those used in the standard CLM4.5 model; column (*c*), land unit (*l*), grid cell (*g*) and soil layer (*j*). On account of the additional complexity of the new representation of plant function, several additional indices are introduced that describe the discretization of plant type, fuel type, litter type, plant height, canopy identity, leaf vertical structure and fuel moisture characteristics. To provide a reference with which to interpret the equations that follow, they are listed here.

Table 1: Table of subscripts used in this document

Parameter Symbol	Parameter Name
<i>ft</i>	Plant Functional Type
<i>fc</i>	Fuel Class
<i>lsc</i>	Litter Size Class
<i>coh</i>	Cohort Index
<i>patch</i>	Patch Index
<i>Cl</i>	Canopy Layer
<i>z</i>	Leaf Layer
<i>mc</i>	Moisture Class

3.5 Cohort State Variables

The unit of allometry in the ED model is the cohort. Each cohort represents a group of plants with similar functional types and heights that occupy portions of column with similar disturbance histories. The state variables of each cohort therefore consist of several pieces of information that fully describe the growth status of the plant and its position in the ecosystem structure, and from which the model can be restarted. The state variables of a cohort are as follows:

Table 2: State Variables of ‘cohort’ structure

Quantity	Variable name	Units	Notes
Plant Functional Type	ft_{coh}	integer	
Number of Individuals	n_{coh}	n per 10000m ⁻²	
Height	h_{coh}	m	
Diameter	dbh_{coh}	cm	
Structural Biomass	$b_{struc,coh}$	KgC plant ⁻¹	Stem wood (above and below ground)
Alive Biomass	$b_{alive,coh}$	KgC plant ⁻¹	Leaf, fine root and sapwood
Stored Biomass	$b_{store,coh}$	KgC plant ⁻¹	Labile carbon reserve
Leaf memory	$l_{memory,coh}$	KgC plant ⁻¹	Leaf mass when leaves are dropped
Canopy Layer	$C_{l,coh}$	integer	1 = top layer
Phenological Status	$S_{phen,coh}$	integer	1=leaves off. 2=leaves on
Canopy trimming	$C_{trim,coh}$	fraction	1.0=max leaf area
Patch Index	p_{coh}	integer	To which patch does this cohort belong?

3.6 Patch State Variables

A patch, as discuss earlier, is a fraction of the landscape which contains ecosystems with similar structure and disturbance history. A patch has no spatial location. The state variables, which are ‘ecosystem’ rather than ‘tree’ scale properties, from which the model can be restarted, are as follows

Table 3: State variables of ‘patch’ structure

Quantity	Variable name	Units	Indexed By
Area	A_{patch}	m ²	-
Age	age_{patch}	years	-
Seed	$Seed_{patch}$	KgC m ⁻²	<i>ft</i>
Leaf Litter	$l_{litter,patch}$	KgC m ⁻²	<i>ft</i>
Root Litter	$r_{litter,patch}$	KgC m ⁻²	<i>ft</i>
AG Coarse Woody Debris	$CWD_{AG,patch}$	KgC m ⁻²	Size Class (<i>lsc</i>)
BG Coarse Woody Debris	$CWD_{BG,patch}$	KgC m ⁻²	Size Class (<i>lsc</i>)
Canopy Spread	$S_{c,patch}$	-	Canopy Layer
Column Index	l_{patch}	integer	-

3.7 Model Structure

Code concerned with the Ecosystem Demography model interfaces with the CLM model in four ways: i) During initialization, ii) During the calculation of surface processes (albedo, radiation absorption, canopy fluxes) each model time step (typically half-hourly), iii) During the main invocation of the ED model code at the end of each day. Daily cohort-level NPP is used to grow plants and alter the cohort structures, disturbance processes (fire and mortality) operate to alter the patch structures, and all fragmenting carbon pool dynamics are calculated. iv) during restart reading and writing. The net assimilation (NPP) fluxes attributed to each cohort are accumulated throughout each daily cycle and passed into the ED code as the major driver of vegetation dynamics.

4 Initialization of vegetation from bare ground

⁵ If the model is restarted from a bare ground state (as opposed to a pre-existing vegetation state), the state variables above are initialized as follows. First, the number of plants per PFT is allocated according to the initial seeding density (S_{init} , individuals per m^2) and the area of the patch A_{patch} , which in the first timestep is the same as the area of the notional ecosystem A_{tot} . The model has no meaningful spatial dimension, but we assign a notional area such that the values of ‘ n_{coh} ’ can be attributed. The default value of A_{tot} is one hectare ($10,000 \text{ m}^2$), but the model will behave identically irrespective of the value of this parameter.

$$n_{\text{coh},0} = S_{\text{init}} A_{\text{patch}} \quad (5)$$

Each cohort is initialized at the minimum canopy height $h_{\text{min},\text{ft}}$, which is specified as a parameter for each plant functional type and denotes the smallest size of plant which is tracked by the model. Smaller plants are not considered, and their emergence from the recruitment processes is unresolved and therefore implicitly parameterized in the seedling establishment model.. The diameter of each cohort is then specified using the log-linear allometry between stem diameter and canopy height

$$dbh_{\text{coh}} = 10^{\frac{\log_{10}(h_{\text{coh}}) - c_{\text{allom}}}{m_{\text{allom}}}} \quad (6)$$

where the slope of the log-log relationship, m_{allom} is 0.64 and the intercept c_{allom} is 0.37. The structural biomass associated with a plant of this diameter and height is given (as a function of wood density, ρ , g cm^{-3})

$$b_{\text{struc},\text{coh}} = c_{\text{str}} h_{\text{coh}}^{\text{e}_{\text{str},\text{hite}}} dbh_{\text{coh}}^{\text{e}_{\text{str},\text{dbh}}} \rho_{\text{ft}}^{\text{e}_{\text{str},\text{dens}}} \quad (7)$$

taken from the original ED1.0 allometry [19] (values of the allometric constants in Table 5. The maximum amount of leaf biomass associated with this diameter of tree is calculated according to the following allometry

$$b_{\text{max},\text{leaf},\text{coh}} = c_{\text{leaf}} dbh_{\text{coh}}^{\text{e}_{\text{leaf},\text{dbh}}} \rho_{\text{ft}}^{\text{e}_{\text{leaf},\text{dens}}} \quad (8)$$

from this quantity, we calculate the active/fine root biomass $b_{\text{root},\text{coh}}$ as

$$b_{\text{root},\text{coh}} = b_{\text{max},\text{leaf},\text{coh}} \cdot f_{\text{frla}} \quad (9)$$

where f_{frla} is the fraction of fine root biomass to leaf biomass, assigned per PFT

Table 4: Parameters needed for model initialization.

Parameter Symbol	Parameter Name	Units	Default Value
h_{min}	Minimum plant height	m	1.5
S_{init}	Initial Planting density	Individuals m^{-2}	
A_{tot}	Model area	m^2	10,000

⁵This description covers algorithms in the ‘init_cohorts’ subroutine

4.1 Allocation of live biomass to leaves roots and sapwood

⁶ Total live biomass b_{alive} is the state variable of the model that describes the sum of the three live biomass pools leaf b_{leaf} , root b_{root} and sapwood b_{sw} (all in $\text{kgC individual}^{-1}$). The quantities are constrained by the following

$$b_{\text{alive}} = b_{\text{leaf}} + b_{\text{root}} + b_{\text{sw}} \quad (10)$$

Sapwood volume is a function of tree height and leaf biomass

$$b_{\text{sw}} = b_{\text{leaf}} \cdot h_{\text{coh}} \cdot f_{\text{swh}} \quad (11)$$

where f_{swh} is the ratio of sapwood mass (kgC) to leaf mass per unit tree height (m). Also, root mass is a function of leaf mass

$$b_{\text{root}} = b_{\text{leaf}} \cdot f_{\text{frla}} \quad (12)$$

Thus

$$b_{\text{alive}} = b_{\text{leaf}} + b_{\text{leaf}} \cdot f_{\text{frla}} + b_{\text{leaf}} \cdot h_{\text{coh}} \cdot f_{\text{swh}} \quad (13)$$

Rearranging gives the fraction of biomass in the leaf pool f_{leaf} as

$$f_{\text{leaf}} = \frac{1}{1 + h_{\text{coh}} \cdot f_{\text{swh}} + f_{\text{frla}}} \quad (14)$$

Thus, we can determine the leaf fraction from the height at the tissue ratios, and the phenological status of the cohort $S_{\text{phen,coh}}$.

$$b_{\text{leaf}} = b_{\text{alive}} \cdot l_{\text{frac}} \quad (15)$$

To divide the live biomass pool at restart, or whenever it is recalculated, into its constituent parts, we first

$$b_{\text{leaf}} = \begin{cases} b_{\text{alive}} \cdot l_{\text{frac}} & \text{for } S_{\text{phen,coh}} = 1 \\ 0 & \text{for } S_{\text{phen,coh}} = 0 \end{cases} \quad (16)$$

Because sometimes the leaves are dropped, using leaf biomass as a predictor of root and sapwood would produce zero live biomass in the winter. To account for this, we add the LAI memory variable l_{memory} to the live biomass pool to account for the need to maintain root biomass when leaf biomass is zero. Thus, to calculate the root biomass, we use

$$b_{\text{root}} = (b_{\text{alive}} + l_{\text{memory}}) \cdot l_{\text{frac}} \cdot f_{\text{frla}} \quad (17)$$

To calculate the sapwood biomass, we use

$$b_{\text{sw}} = (b_{\text{alive}} + l_{\text{memory}}) \cdot l_{\text{frac}} \cdot f_{\text{swh}} \cdot h_{\text{coh}} \quad (18)$$

Table 5: Allometric Constants

Parameter Symbol	Parameter Name	Units	Default Value
c_{allom}	Allometry intercept		0.37
m_{allom}	Allometry slope		0.64
c_{str}	Structural biomass multiplier		0.06896
$e_{\text{str,dbh}}$	Structural Biomass dbh exponent		1.94
$e_{\text{str,hite}}$	Structural Biomass height exponent		0.572
$e_{\text{str,dens}}$	Structural Biomass density exponent		0.931
c_{leaf}	Leaf biomass multiplier		0.0419
$e_{\text{leaf,dbh}}$	Leaf biomass dbh exponent		1.56
$e_{\text{leaf,dens}}$	Leaf biomass density exponent		0.55
f_{swh}	Ratio of sapwood mass to height	m^{-1}	
f_{frla}	Ratio of fine root mass to leaf mass	-	1.0

⁶This description relates to algorithms in the `allocate_live_biomass` subroutine

5 Canopy Structure and the Perfect Plasticity Approximation

⁷ During initialization and every subsequent daily ED timestep, the canopy structure model is called to determine how the leaf area of the different cohorts is arranged relative to the incoming radiation, which will then be used to drive the radiation and photosynthesis calculations. This task requires that some assumptions are made about 1) the shape and depth of the canopy within which the plant leaves are arranged and 2) how the leaves of different cohorts are arranged relative to each other. This set of assumptions are critical to model performance in ED-like cohort based models, since they determine how light resources are partitioned between competing plants of varying heights, which has a very significant impact on how vegetation distribution emerges from competition [7].

The standard ED1.0 model makes a simple ‘flat disk’ assumption, that the leaf area of each cohort is spread in an homogenous layer at one exact height across entire the ground area represented by each patch. The CLM(ED) model has diverged from this representation due to (at least) two problematic emergent properties that we identified as generating unrealistic behaviours especially for large-area patches.

1. Over-estimation of light competition . The vertical stacking of cohorts which have all their leaf area at the same nominal height means that when one cohort is only very slightly taller than it’s competitor, it is completely shaded by it. This means that any small advantage in terms of height growth translates into a large advantage in terms of light competition, even at the seedling stage. This property of the model artificially exaggerates the process of light competition. In reality, trees do not compete for light until their canopies begin to overlap and canopy closure is approached.

2. Unrealistic over-crowding. The ‘flat-disk’ assumption has no consideration of the spatial extent of tree crowns. Therefore it has no control on the packing density of plants in the model. Given a mismatch between production and mortality, entirely unrealistic tree densities are thus possible for some combinations of recruitment, growth and mortality rates.

To account for the filling of space in three dimensions using the one-dimensional representation of the canopy employed by CLM, we implement a new scheme derived from that of [24]. Their argument follows the development of an individual-based variant of the SORTIE model, called SHELL, which allows the location of individual plant crowns to be highly flexible in space. Ultimately, the solutions of this model possess an emergent property whereby the crowns of the plants simply fill all of the available space in the canopy before forming a distinct understorey.

Purves et al. developed a model that uses this feature, called the ‘perfect plasticity approximation’, which assumes the plants are able to perfectly fill all of the available canopy space. That is, at canopy closure, all of the available horizontal space is filled, with negligible gaps, owing to lateral tree growth and the ability of tree canopies to grow into the available gaps (this is of course, an over-simplified but potential useful ecosystem property). The ‘perfect plasticity approximation’ (PPA) implies that the community of trees is subdivided into discrete canopy layers, and by extension, each cohort represented by the CLM(ED) model is assigned a canopy layer status flag, C_L . In this version, we set the maximum number of canopy layers at 2 for simplicity, although is possible to have a larger number of layers in theory. $C_{L,coh} = 1$ means that all the trees of cohort *coh* are in the upper canopy (overstorey), and $C_{L,coh} = 2$ means that all the trees of cohort *coh* are in the understorey.

In this model, all the trees in the canopy experience full light on their uppermost leaf layer, and all trees in the understorey experience the same light (full sunlight attenuated by the average LAI of the upper canopy) on their uppermost leaves, as described in the radiation transfer section (more nuanced versions of this approach may be investigated in future model versions). The canopy is assumed to be cylindrical, the lower layers of which experience self-shading by the upper layers.

To determine whether a second canopy layer is required, the model needs to know the spatial extent of tree crowns. Crown area, A_{crown} , m^2 , is defined as

$$A_{crown,coh} = \pi(dbh_{coh}S_{c,patch,C_L})^{1.56} \quad (19)$$

where A_{crown} is the crown area of a single tree canopy (m^2) and $S_{c,patch,C_L}$ is the ‘canopy spread’ parameter ($m\ cm^{-1}$) of this canopy layer, which is assigned as a function of canopy space filling, discussed below. In contrast to [24], we use an exponent, identical to that for leaf biomass, of 1.56, not 2.0, such that tree leaf area index does not change as a function of diameter.

⁷This description relates to algorithms in the EDCanopyStructure subroutine

To determine whether the canopy is closed, we calculate the total canopy area as:

$$A_{\text{canopy}} = \sum_{\text{coh}=1}^{\text{nc,patch}} A_{\text{crown,coh}} \cdot n_{\text{coh}} \quad (20)$$

where n_{cpatch} is the number of cohorts in a given patch. If the area of all crowns A_{canopy} (m^2) is larger than the total ground area of a patch (A_{patch}), then some fraction of each cohort is demoted to the understorey.

Under these circumstances, the ‘extra’ crown area A_{loss} (i.e., $A_{\text{canopy}} - A_p$) is moved into the understorey. For each cohort already in the canopy, we determine a fraction of trees that are moved from the canopy (L_c) to the understorey. L_c is calculated as [7]

$$L_c = \frac{A_{\text{loss,patch}} w_{\text{coh}}}{\sum_{\text{coh}=1}^{\text{nc,patch}} w_{\text{coh}}}, \quad (21)$$

where w_{coh} is a weighting of each cohort determined by basal diameter dbh (cm) and the competitive exclusion coefficient C_e

$$w_{\text{coh}} = dbh_{\text{coh}} C_e. \quad (22)$$

The higher the value of C_e the greater the impact of tree diameter on the probability of a given tree obtaining a position in the canopy layer. That is, for high C_e values, competition is highly deterministic. The smaller the value of C_e , the greater the influence of random factors on the competitive exclusion process, and the higher the probability that slower growing trees will get into the canopy. Appropriate values of C_e are poorly constrained but alter the outcome of competitive processes.

The process by which trees are moved between canopy layers is complex because 1) the crown area predicted for a cohort to lose may be larger than the total crown area of the cohort, which requires iterative solutions, and 2) on some occasions (e.g. after fire), the canopy may open up and require ‘promotion’ of cohorts from the understorey, and 3) canopy area may change due to the variations of canopy spread values ($S_{\text{c,patch,Cl}}$, see the section below for details) when fractions of cohorts are demoted or promoted. Further details can be found in the code references in the footnote.

5.1 Horizontal Canopy Spread

⁸ [24] estimated the ratio between canopy and stem diameter c_p as 0.1 m cm^{-1} for canopy trees in North American forests, but this estimate was made on trees in closed canopies, whose shape is subject to space competition from other individuals. Sapling trees have no constraints in their horizontal spatial structure, and as such, are more likely to display their leaves to full sunlight. Also, prior to canopy closure, light interception by leaves on the sides of the canopy is also higher than it would be in a closed canopy forest. If the ‘canopy spread’ parameter is constant for all trees, then we simulate high levels of self-shading for plants in unclosed canopies, which is arguably unrealistic and can lower the productivity of trees in areas of unclosed canopy (e.g. low productivity areas of boreal or semi-arid regions where LAI and canopy cover might naturally be low). We here interpret the degree of canopy spread, S_c as a function of how much tree crowns interfere with each other in space, or the total canopy area A_{canopy} . However A_{canopy} itself is a function of S_c , leading to a circularity. S_c is thus solved iteratively through time.

Each daily model step, A_{canopy} and the fraction of the gridcell occupied by tree canopies in each canopy layer ($A_{f,Cl} = A_{\text{canopy,Cl}}/A_{\text{patch}}$) is calculated based on S_c from the previous timestep. If A_f is greater than a threshold value A_t , S_c is increased by a small increment i . The threshold A_t is, hypothetically, the canopy fraction at which light competition begins to impact on tree growth. This is less than 1.0 owing to the non-perfect spatial spacing of tree canopies. If $A_{f,Cl}$ is greater than A_t , then S_c is reduced by an increment i , to reduce the spatial extent of the canopy, thus.

$$S_{\text{c,patch,Cl,t+1}} = \begin{cases} S_{\text{c,patch,Cl,t}} + i & \text{for } A_{f,Cl} < A_t \\ S_{\text{c,patch,Cl,t}} - i & \text{for } A_{f,Cl} > A_t \end{cases} \quad (23)$$

The values of S_c are bounded to upper and lower limits. The lower limit corresponds to the observed canopy spread parameter for canopy trees $S_{c,\text{min}}$ and the upper limit corresponds to the largest canopy extent $S_{c,\text{max}}$

$$S_{\text{c,patch,Cl}} = \begin{cases} S_{c,\text{min}} & \text{for } S_{\text{c,patch,Cl}} < S_{c,\text{min}} \\ S_{c,\text{max}} & \text{for } S_{\text{c,patch,Cl}} > S_{c,\text{max}} \end{cases} \quad (24)$$

⁸This description relates to algorithms in the canopy_spread subroutine

This iterative scheme requires two additional parameters (i and A_t). i affects the speed with which canopy spread (and hence leaf area index) increase as canopy closure is neared. However, the model is relatively insensitive to the choice of either i or A_t .

5.2 Definition of Leaf Area Profile

⁹ Within each patch, the model defines and tracks cohorts of multiple plant functional types that exist either in the canopy or understorey. Light on the top leaf surface of each cohort in the canopy is the same, and the rate of decay through the canopy is also the same for each PFT. Therefore, we accumulate all the cohorts of a given PFT together for the sake of the radiation and photosynthesis calculations (to avoid separate calculations for every cohort).

Therefore, the leaf area index for each patch is defined as a three-dimensional array $lai_{Cl,ft,z}$ where Cl is the canopy layer, ft is the functional type and z is the leaf layer within each canopy. This three-dimensional structure is the basis of the radiation and photosynthetic models. In addition to a leaf area profile matrix, we also define, for each patch, the area which is covered by leaves at each layer as $care_{Cl,ft,z}$.

Each plant cohort is already defined as a member of a single canopy layer and functional type. This means that to generate the $lai_{Cl,ft,z}$ matrix, it only remains to divide the leaf area of each cohort into leaf layers. First, we determine how many leaf layers are occupied by a single cohort, by calculating the ‘tree LAI’ as the total leaf area of each cohort divided by its crown area (both in m^2)

$$tree_{lai,coh} = \frac{b_{leaf,coh} \cdot sla_{ft}}{A_{crown,coh}} \quad (25)$$

where sla_{ft} is the specific leaf area in $m^2 KgC^{-1}$ and b_{leaf} is in kGC per plant.

5.2.1 Stem Area Index

Stem area index (SAI) is ratio of the total area of all woody stems on a plant to the area of ground covered by the plant. During winter in deciduous areas, the extra absorption by woody stems can have a significant impact on the surface energy budget. However, in previous ‘big leaf’ versions of the CLM, computing the circumstances under which stem area was visible in the absence of leaves was difficult and the algorithm was largely heuristic as a result. Given the multi-layer canopy introduced for CLM(ED), we can determine the leaves in the higher canopy layers will likely shade stem area in the lower layers when leaves are on, and therefore stem area index can be calculated as a function of woody biomass directly.

Literature on stem area index is particularly poor, as it’s estimation is complex and not particularly amenable to the use of, for example, assumptions of random distribution in space that are typically used to calculate leaf area from light interception. [10] estimated that SAI visible from an LAI2000 sensor was around $0.5 m^2 m^{-2}$. [11] estimate that the wood area index for Ponderosa Pine forest is 0.27-0.33. The existing CLM(CN) algorithm sets the minimum SAI at 0.25 to match MODIS observations, but then allows SAI to rise as a function of the LAI lost, meaning that in some places, predicted SAI can reach value of 8 or more. Clearly, greater scientific input on this quantity is badly needed. Here we determine that SAI is a linear function of woody biomass, to at very least provide a mechanistic link between the existence of wood and radiation absorbed by it. The non-linearity between how much woody area exists and how much radiation is absorbed is provided by the radiation absorption algorithm. Specifically, the SAI of an individual cohort ($tree_{sai,coh}$, $m^2 m^{-2}$) is calculated as follows,

$$tree_{sai,coh} = k_{sai} \cdot b_{struc,coh}, \quad (26)$$

where k_{sai} is the coefficient linking structural biomass to SAI. The number of occupied leaf layers for cohort coh ($n_{z,coh}$) is then equal to the rounded up integer value of the tree SAI ($tree_{sai,coh}$) and LAI ($tree_{lai,coh}$) divided by the layer thickness (i.e., the resolution of the canopy layer model, in units of vegetation index ($lai+sai$) with a default value of 1.0, δ_{vai}),

$$n_{z,coh} = \frac{tree_{lai,coh} + tree_{sai,coh}}{\delta_{vai}}. \quad (27)$$

The fraction of each layer that is leaf (as opposed to stem) can then be calculated as

$$f_{leaf,coh} = \frac{tree_{lai,coh}}{tree_{sai,coh} + tree_{lai,coh}}. \quad (28)$$

⁹This description relates to algorithms in the `canopy_leaf_area_profile` subroutine

Finally, the leaf area in each leaf layer pertaining to this cohort is thus

$$lai_{z,coh} = \begin{cases} \delta_{vai} \cdot f_{leaf,coh} \frac{A_{canopy,coh}}{A_{canopy,patch}} & \text{for } i = 1, \dots, i = n_{z,coh} - 1 \\ \delta_{vai} \cdot f_{leaf,coh} \frac{A_{canopy,coh}}{A_{canopy,patch}} \cdot r_{vai} & \text{for } i = n_{z,coh} \end{cases} \quad (29)$$

and the stem area index is

$$sai_{z,coh} = \begin{cases} \delta_{vai} \cdot (1 - f_{leaf,coh}) \frac{A_{canopy,coh}}{A_{canopy,patch}} & \text{for } i = 1, \dots, i = n_{z,coh} - 1 \\ \delta_{vai} \cdot (1 - f_{leaf,coh}) \frac{A_{canopy,coh}}{A_{canopy,patch}} \cdot r_{vai} & \text{for } i = n_{z,coh} \end{cases} \quad (30)$$

where r_{vai} is the remainder of the canopy that is below the last full leaf layer

$$r_{vai} = (\text{tree}_{lai,coh} + \text{tree}_{sai,coh}) - (\delta_{vai} \cdot (n_{z,coh} - 1)). \quad (31)$$

$A_{canopy,patch}$ is the total canopy area occupied by plants in a given patch (m^2) and is calculated as follows,

$$A_{canopy,patch} = \min \left(\sum_{coh=1}^{coh=n_{coh}} A_{canopy,coh}, A_{patch} \right). \quad (32)$$

The canopy is conceived as a cylinder, although this assumption could be altered given sufficient evidence that canopy shape was an important determinant of competitive outcomes, and the area of ground covered by each leaf layer is the same through the cohort canopy. With the calculated SAI and LAI, we are able to calculate the complete canopy profile. Specifically, the relative canopy area for the cohort coh is calculated as

$$area_{1:nz,coh} = \frac{A_{crown,coh}}{A_{canopy,patch}}. \quad (33)$$

The total occupied canopy area for each canopy layer (Cl), plant functional type (ft) and leaf layer (z) bin is thus

$$c_{area,Cl,ft,z} = \sum_{coh=1}^{coh=n_{coh}} area_{1:nz,coh} \quad \text{where } ft_{coh} = ft \text{ and } Cl_{coh} = Cl. \quad (34)$$

All of these quantities are summed across cohorts to give the complete leaf and stem area profiles,

$$\begin{aligned} lai_{Cl,ft,z} &= \sum_{coh=1}^{coh=n_{coh}} lai_{z,coh} \quad \text{where } ft_{coh} = ft \text{ and } Cl_{coh} = Cl, \\ sai_{Cl,ft,z} &= \sum_{coh=1}^{coh=n_{coh}} sai_{z,coh} \quad \text{where } ft_{coh} = ft \text{ and } Cl_{coh} = Cl. \end{aligned} \quad (35)$$

5.2.2 Burial of leaf area by snow

The calculations above all pertain to the total leaf and stem area indices which characterize the vegetation structure. In addition, the model must know when the vegetation is covered by snow, and by how much, so that the albedo and energy balance calculations can be adjusted accordingly. Therefore, we calculated a ‘total’ and ‘exposed’ lai and sai profile using a representation of the bottom and top canopy heights, and the depth of the average snow pack. For each leaf layer z of each cohort, we calculate an ‘exposed fraction’ $f_{exp,z}$ via consideration of the top and bottom heights of that layer $h_{top,z}$ and $h_{bot,z}$ (m),

$$\begin{aligned} h_{top,z} &= h_{coh} - h_{coh} \cdot f_{crown,ft} \cdot \frac{z}{n_{z,coh}} \\ h_{bot,z} &= h_{coh} - h_{coh} \cdot f_{crown,ft} \cdot \frac{z+1}{n_{z,coh}} \end{aligned} \quad (36)$$

where $f_{crown,ft}$ is the plant functional type (ft) specific fraction of the cohort height that is occupied by the crown. Specifically, the ‘exposed fraction’ $f_{exp,z}$ is calculated as follows,

$$f_{exp,z} = \begin{cases} = 1.0 & \text{where } h_{bot,z} > d_{snow} \\ = \frac{d_{snow} - h_{bot,z}}{h_{top,z} - h_{bot,z}} & \text{where } h_{top,z} > d_{snow} \text{ and } h_{bot,z} < d_{snow} \\ = 0.0 & \text{where } h_{top,z} < d_{snow} \end{cases} \quad (37)$$

The resulting exposed ($elai, esai$) and total ($tlai, tsai$) leaf and stem area indices are calculated as

$$\begin{aligned} elai_{C1,ft,z} &= lai_{C1,ft,z} \cdot f_{exp,z} \\ esai_{C1,ft,z} &= sai_{C1,ft,z} \cdot f_{exp,z} \\ tlai_{C1,ft,z} &= lai_{C1,ft,z} \\ tsai_{C1,ft,z} &= sai_{C1,ft,z} \end{aligned}, \quad (38)$$

and are used in the radiation interception and photosynthesis algorithms described later.

Table 6: Parameters needed for canopy structure model.

Parameter Symbol	Parameter Name	Units	Notes	Indexed by
δ_{vai}	Thickness of single canopy layer	$m^{-2}m^{-2}$		
C_e	Competitive Exclusion Parameter	none		
$c_{p,min}$	Minimum canopy spread	$m^2 cm^{-1}$		
$c_{p,max}$	Competitive Exclusion Parameter	$m^2 cm^{-1}$		
i	Incremental change in c_p	$m^2 cm^{-1} y^{-1}$		
A_t	Threshold canopy closure	none		
$f_{crown,ft}$	Crown fraction	none		ft
k_{sai}	Stem area per unit woody biomass	$m^2 KgC^{-1}$		

6 Radiation Transfer

6.1 Fundamental Radiation Transfer Theory

¹⁰ The first interaction of the land surface with the properties of vegetation concerns the partitioning of energy into that which is absorbed by vegetation, reflected back into the atmosphere, and absorbed by the ground surface. Older versions of the CLM have utilized a ‘two-stream’ approximation [27, 28] that provided an empirical solution for the radiation partitioning of a multi-layer canopy for two streams, of diffuse and direct light. However, implementation of the Ecosystem Demography model requires a) the adoption of an explicit multiple layer canopy b) the implementation of a multiple plant type canopy and c) the distinction of canopy and under-storey layers, in-between which the radiation streams are fully mixed. The radiation mixing between canopy layers is necessary as the position of different plants in the under-storey is not defined spatially or relative to the canopy trees above. In this new scheme, we thus implemented a one-dimensional scheme that traces the absorption, transmittance and reflectance of each canopy layer and the soil, iterating the upwards and downwards passes of radiation through the canopy until a pre-defined accuracy tolerance is reached. This approach is based on the work of [20].

Here we describe the basic theory of the radiation transfer model for the case of a single homogenous canopy, and in the next section we discuss how this is applied to the multi layer multi PFT canopy in the CLM(ED) implementation. The code considers the fractions of a single unit of incoming direct and a single unit of incoming diffuse light, that are absorbed at each layer of the canopy for a given solar angle (α_s , radians). Direct radiation is extinguished through the canopy according to the coefficient k_{dir} that is calculated from the incoming solar angle and the dimensionless leaf angle distribution parameter (χ) as

$$k_{dir} = g_{dir} / \sin(\alpha_s) \quad (39)$$

where

$$g_{dir} = \phi_1 + \phi_2 \cdot \sin(\alpha_s) \quad (40)$$

and

$$\begin{aligned} \phi_1 &= 0.5 - 0.633\chi_1 - 0.33\chi_1^2 \\ \phi_2 &= 0.877(1 - 2\phi_1) \end{aligned} \quad (41)$$

The leaf angle distribution is a descriptor of how leaf surfaces are arranged in space. Values approaching 1.0 indicate that (on average) the majority of leaves are horizontally arranged with respect to the ground. Values approaching -1.0 indicate that leaves are mostly vertically arranged, and a value of 0.0 denotes a canopy where leaf angle is random (a ‘spherical’ distribution).

¹⁰This description relates to algorithms in the ED_norman_radiation subroutine

According to Beer's Law, the fraction of light that is transferred through a single layer of vegetation (leaves or stems) of thickness δ_{vai} , without being intercepted by any surface, is

$$tr_{\text{dir}} = e^{-k_{\text{dir}}\delta_{\text{vai}}} \quad (42)$$

and the incident direct radiation transmitted to each layer of the canopy ($dir_{\text{tr},z}$) is thus calculated from the cumulative leaf area (L_{above}) shading each layer (z):

$$dir_{\text{tr},z} = e^{-k_{\text{dir}}L_{\text{above},z}} \quad (43)$$

The fraction of the leaves f_{sun} that are exposed to direct light is also calculated from the decay coefficient k_{dir} .

$$\begin{aligned} f_{\text{sun},z} &= e^{-k_{\text{dir}}L_{\text{above},z}} \\ \text{and} \\ f_{\text{shade},z} &= 1 - f_{\text{sun},z} \end{aligned} \quad (44)$$

where $f_{\text{shade},z}$ is the fraction of leaves that are shaded from direct radiation and only receive diffuse light. Diffuse radiation, by definition, enters the canopy from a spectrum of potential incident directions, therefore the un-intercepted transfer (tr_{dif}) through a leaf layer of thickness δ_l is calculated as the mean of the transfer rate from each of 9 different incident light directions (α_s) between 0 and 180 degrees to the horizontal.

$$tr_{\text{dif}} = \frac{1}{9} \sum_{\alpha_s=5\pi/180}^{\alpha_s=85\pi/180} e^{-k_{\text{dir},l}\delta_{\text{vai}}} \quad (45)$$

$$tr_{\text{dif}} = \frac{1}{9} \pi \sum_{\alpha_s=0}^{\pi/2} \frac{e^{-g_{\text{dir}}\alpha_s}}{\delta_{\text{vai}} \cdot \sin(\alpha_s) \sin(\alpha_s) \cos(\alpha_s)} \quad (46)$$

The fraction ($1-tr_{\text{dif}}$) of the diffuse radiation is intercepted by leaves as it passes through each leaf layer. Of this, some fraction is reflected by the leaf surfaces and some is transmitted through. The fractions of diffuse radiation reflected from ($refl_{\text{dif}}$) and transmitted through ($tran_{\text{dif}}$) each layer of leaves are thus, respectively

$$\begin{aligned} refl_{\text{dif}} &= (1 - tr_{\text{dif}})\rho_{l,\text{ft}} \\ tran_{\text{dif}} &= (1 - tr_{\text{dif}})\tau_{l,\text{ft}} + tr_{\text{dif}} \end{aligned} \quad (47)$$

where $\rho_{l,\text{ft}}$ and $\tau_{l,\text{ft}}$ are the fractions of incident light reflected and transmitted by individual leaf surfaces. Once we know the fractions of light that are transmitted and reflected by each leaf layer, we begin the process of distributing light through the canopy. Starting with the first leaf layer ($z=1$), where the incident downwards diffuse radiation (dif_{down}) is 1.0, we work downwards for n_z layers, calculating the radiation in the next layer down ($z+1$) as:

$$dif_{\text{down},z+1} = \frac{dif_{\text{down},z} tran_{\text{dif}}}{1 - r_{z+1} refl_{\text{dif}}} \quad (48)$$

Here, $dif_{\text{down},z} tran_{\text{dif}}$ calculates the fraction of incoming energy transmitted downwards onto layer $z+1$. This flux is then increased by the additional radiation r_z that is reflected upwards from further down in the canopy to layer z , and then is reflected back downwards according to the reflected fraction $refl_{\text{dif}}$. The more radiation in $r_{z+1} refl_{\text{dif}}$, the smaller the denominator and the larger the downwards flux. r is also calculated sequentially, starting this time at the soil surface layer (where $z = n_z + 1$)

$$r_{n_z+1} = alb_s \quad (49)$$

where alb_s is the soil albedo characteristic. The upwards reflected fraction r_z for each leaf layer, moving upwards, is then [20]

$$r_z = \frac{r_{z+1} \times tran_{\text{dif}}^2}{(1 - r_{z+1} refl_{\text{dif}}) + refl_{\text{dif}}} \quad (50)$$

The corresponding upwards diffuse radiation flux is therefore the fraction of downwards radiation that is incident on a particular layer, multiplied by the fraction that is reflected from all the lower layers:

$$dif_{\text{up},z} = r_z dif_{\text{down},z+1} \quad (51)$$

Now we have initial conditions for the upwards and downwards diffuse fluxes, these must be modified to account for the fact that, on interception with leaves, direct radiation is transformed into diffuse radiation. In addition, the initial solutions to the upwards and downwards radiation only allow a single ‘bounce’ of radiation through the canopy, so some radiation which might be intercepted by leaves higher up is potentially lost. Therefore, the solution to this model is iterative. The iterative solution has upwards and a downwards components that calculate the upwards and downwards fluxes of total radiation at each leaf layer ($rad_{dn, z}$ and $rad_{up, z}$). The downwards component begins at the top canopy layer ($z = 1$). Here we define the incoming solar diffuse and direct radiation ($solar_{dif}$ and $solar_{dir}$ respectively).

$$\begin{aligned} dif_{dn,1} &= solar_{dif} \\ rad_{dn, z+1} &= dif_{dn,z} \cdot tran_{dif} + dif_{up,z+1} \cdot refl_{dif} + solar_{dir} \cdot dir_{tr,z}(1 - tr_{dir})\tau_l. \end{aligned} \quad (52)$$

The first term of the right-hand side deals with the diffuse radiation transmitted downwards, the second with the diffuse radiation travelling upwards, and the third with the direct radiation incoming at each layer ($dir_{tr,z}$) that is intercepted by leaves ($1 - tr_{dir}$) and then transmitted through the leaf matrix as diffuse radiation (τ_l). At the bottom of the canopy, the light reflected off the soil surface is calculated as

$$rad_{up, nz} = dif_{down,nz} \cdot salb_{dif} + solar_{dir} \cdot dir_{tr,z} salb_{dir}. \quad (53)$$

The upwards propagation of the reflected radiation is then

$$rad_{up, z} = dif_{up,z+1} \cdot tran_{dif} + dif_{dn,z} \cdot refl_{dif} + solar_{dir} \cdot dir_{tr,z}(1 - tr_{dir})\rho_l. \quad (54)$$

Here the first two terms deal with the diffuse downwards and upwards fluxes, as before, and the third deals direct beam light that is intercepted by leaves and reflected upwards. These upwards and downwards fluxes are computed for multiple iterations, and at each iteration, $rad_{up, z}$ and $rad_{down, z}$ are compared to their values in the previous iteration. The iteration scheme stops once the differences between iterations for all layers is below a predefined tolerance factor, (set here at 10^{-4}). Subsequently, the fractions of absorbed direct ($abs_{dir,z}$) and diffuse ($abs_{dif,z}$) radiation for each leaf layer then

$$abs_{dir,z} = solar_{dir} \cdot dir_{tr,z} \cdot (1 - tr_{dir}) \cdot (1 - \rho_l - \tau_l) \quad (55)$$

$$abs_{dif,z} = (dif_{dn,z} + dif_{up,z+1}) \cdot (1 - tr_{dif}) \cdot (1 - \rho_l - \tau_l). \quad (56)$$

and, the radiation energy absorbed by the soil for the diffuse and direct streams is calculated as

$$abs_{soil} = dif_{down,nz+1} \cdot (1 - salb_{dif}) + solar_{dir} \cdot dir_{tr,nz+1} \cdot (1 - salb_{dir}). \quad (57)$$

Canopy level albedo is denoted as the upwards flux from the top leaf layer

$$alb_{canopy} = \frac{dif_{up,z+1}}{solar_{dir} + solar_{dif}} \quad (58)$$

and the division of absorbed energy into sunlit and shaded leaf fractions, (required by the photosynthesis calculations), is

$$abs_{sha,z} = abs_{dif,z} \cdot f_{sha} \quad (59)$$

$$abs_{sun,z} = abs_{dif,z} \cdot f_{sun} + abs_{dir,z} \quad (60)$$

6.2 Resolution of radiation transfer theory within the CLM(ED) canopy structure

The radiation transfer theory above, was described with reference to a single canopy of one plant functional type, for the sake of clarity of explanation. The CLM(ED) model, however, calculates radiative and photosynthetic fluxes for a more complex hierarchical structure within each patch/time-since-disturbance class, as described in the leaf area profile section. Firstly, we denote two or more canopy layers (denoted C_l). The concept of a ‘canopy layer’ refers to the idea that plants are organized into discrete over and under-stories, as predicted by the Perfect Plasticity Approximation [24, 7]. Within each canopy layer there potentially exist multiple cohorts of different plant functional types and heights. Within each canopy layer, C_l , and functional type, ft , the model resolves numerous leaf layers ‘ z ’, and, for some processes, notably photosynthesis, each leaf layer is split into a fraction of sun and shade leaves, f_{sun} and f_{sha} , respectively.

The radiation scheme described in Section is solved explicitly for this structure, for both the visible and near-infrared wavebands, according to the following assumptions.

- A ‘canopy layer’ (C_L) refers to either the over or understorey
- A ‘leaf layer’ (z) refers to the discretization of the LAI within the canopy of a given plant functional type.
- All PFT’s in the same canopy layer have the same solar radiation incident on the top layer of the canopy
- Light is transmitted through the canopy of each plant functional type independently
- Between canopy layers, the light streams from different plant functional types are mixed, such that the (undefined) spatial location of plants in lower canopy layers does not impact the amount of light received.
- Where understorey layers fill less area than the overstorey layers, radiation is directly transferred to the soil surface.
- All these calculations pertain to a single patch, so we omit the ‘patch’ subscript for simplicity in the following discussion.

Within this framework, the majority of the terms in the radiative transfer scheme are calculated with indices of C_L , ft and z . In the following text, we revisit the simplified version of the radiation model described above, and explain how it is modified to account for the more complex canopy structure used by the CLM(ED).

Firstly, the light penetration functions, k_{dir} and g_{dir} are described as functions of ft , because the leaf angle distribution, χ_l , is a pft-specific parameter. Thus, the diffuse irradiance transfer rate, tr_{dif} is also ft specific because g_{dir} , on which it depends, is a function of χ_l .

The amount of direct light reaching each leaf layer is a function of the leaves existing above the layer in question. If a leaf layer ‘ z ’ is in the top canopy layer (the over-storey), it is only shaded by leaves of the same PFT so k_{dir} is unchanged from equation. If there is more than one canopy layer ($C_{l,max} > 1$), then the amount of direct light reaching the top leaf surfaces of the second/lower layer is the weighted average of the light attenuated by all the parallel tree canopies in the canopy layer above, thus.

$$dir_{tr,C_l,z,1} = \sum_{ft=1}^{npft} (dir_{tr,C_l,ft,z_{max}} \cdot c_{area,C_l-1,ft,z_{max}}) \quad (61)$$

where pft_{wt} is the areal fraction of each canopy layer occupied by each functional type and z_{max} is the index of the bottom canopy layer of each pft in each canopy layer (the subscripts C_l and ft are implied but omitted from all z_{max} references to avoid additional complications)

Similarly, the sunlit fraction for a leaf layer ‘ z ’ in the second canopy layer (where $C_l > 1$) is

$$f_{sun,C_l,ft,z} = W_{sun,C_l} \cdot e^{k_{dir,ft,laic,z}} \quad (62)$$

where W_{sun,C_l} is the weighted average sunlit fraction in the bottom layer of a given canopy layer.

$$W_{sun,C_l} = \sum_{ft=1}^{npft} (f_{sun,C_l-1,ft,z_{max}} \cdot c_{area,C_l-1,ft,z_{max}}) \quad (63)$$

Following through the sequence of equations for the simple single pft and canopy layer approach above, the $refl_{dif}$ and $tran_{dif}$ fluxes are also indexed by C_l , ft , and z . The diffuse radiation reflectance ratio r_z is also calculated in a manner that homogenizes fluxes between canopy layers. For the canopy layer nearest the soil ($C_l = C_{l,max}$). For the top canopy layer ($C_l=1$), a weighted average reflectance from the lower layers is used as the baseline, in lieu of the soil albedo. Thus:

$$r_{z,C_l,z,1} = \sum_{ft=1}^{npft} (r_{z,C_l-1,ft,1} pft_{wt,C_l-1,ft,1}) \quad (64)$$

For the iterative flux resolution, the upwards and downwards fluxes are also averaged between canopy layers, thus where $C_l > 1$

$$rad_{dn,C_l,ft,1} = \sum_{ft=1}^{npft} (rad_{dn,C_l-1,ft,z_{max}} \cdot pft_{wt,C_l-1,ft,z_{max}}) \quad (65)$$

and where $C_l=1$, and $C_{l,max} > 1$

$$rad_{up,C_l,ft,z_{max}} = \sum_{ft=1}^{npft} (rad_{up,C_l+1,ft,1} \cdot pft_{wt,C_l+1,ft,1}) \quad (66)$$

The remaining terms in the radiation calculations are all also indexed by C_l , ft and z so that the fraction of absorbed radiation outputs are termed $abs_{dir,C_l,ft,z}$ and $abs_{dif,C_l,ft,z}$. The sunlit and shaded absorption rates are therefore

$$abs_{sha,C_l,ft,z} = abs_{dif,C_l,ft,z} \cdot f_{sha,C_l,ft,z} \quad (67)$$

and

$$abs_{sun,C_l,ft,z} = abs_{dif,C_l,ft,z} \cdot f_{sun,C_l,ft,z} + abs_{dir,C_l,ft,z} \quad (68)$$

The albedo of the mixed pft canopy is calculated as the weighted average of the upwards radiation from the top leaf layer of each pft where $C_l=1$:

$$alb_{canopy} = \sum_{ft=1}^{npft} \frac{dif_{up,1,ft,1} pft_{wt,1,ft,1}}{solar_{dir} + solar_{dif}} \quad (69)$$

The radiation absorbed by the soil after passing through through under-storey vegetation is:

$$abs_{soil} = \sum_{ft=1}^{npft} pft_{wt,1,ft,1} (dif_{down,nz+1} (1 - salb_{dif}) + solar_{dir} dir_{tr,nz+1} (1 - salb_{dir})) \quad (70)$$

to which is added the diffuse flux coming directly from the upper canopy and hitting no understorey vegetation.

$$abs_{soil} = abs_{soil} + dif_{dn,2,1} (1 - \sum_{ft=1}^{npft} pft_{wt,1,ft,1}) (1 - salb_{dif}) \quad (71)$$

and the direct flux coming directly from the upper canopy and hitting no understorey vegetation.

$$abs_{soil} = abs_{soil} + solar_{dir} dir_{tr,2,1} (1 - \sum_{ft=1}^{npft} pft_{wt,1,ft,1}) (1 - salb_{dir}) \quad (72)$$

These changes to the radiation code are designed to be structurally flexible, and the scheme may be collapsed down to only include on canopy layer, functional type and pft for testing if necessary.

Table 7: Parameters needed for radiation transfer model.

Parameter Symbol	Parameter Name	Units	indexed by
χ	Leaf angle distribution parameter	none	ft
ρ_l	Fraction of light reflected by leaf surface	none	ft
τ_l	Fraction of light transmitted by leaf surface	none	ft
alb_s	Fraction of light reflected by soil	none	direct vs diffuse

7 Photosynthesis

7.1 Fundamental photosynthetic physiology theory

¹¹ In this section we describe the physiological basis of the photosynthesis model before describing its application to the CLM(ED) canopy structure. This description in this section is largely repeated from the Oleson et al. CLM4.5 technical note but included here for comparison with its implementation in CLM(ED). Photosynthesis in C3 plants is based on the model of [6] as modified by [4]. Photosynthetic assimilation in C4 plants is based on the model of [5]. In both models, leaf photosynthesis, gpp ($\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$) is calculated as the minimum of three potentially limiting fluxes, described below:

$$gpp = \min(w_j, w_c, w_p). \quad (73)$$

The RuBP carboxylase (Rubisco) limited rate of carboxylation w_c ($\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$) is determined as

$$w_c = \begin{cases} \frac{V_{c,max}(c_i - \Gamma_*)}{c_i + K_c(1 + o_i/K_o)} & \text{for } C_3 \text{ plants} \\ V_{c,max} & \text{for } C_4 \text{ plants} \end{cases} \quad c_i - \Gamma_* \geq 0 \quad (74)$$

¹¹This description relates to algorithms in the ED_photosynthesis subroutine

where c_i is the internal leaf CO_2 partial pressure (Pa) and $o_i(0.209P_{\text{atm}})$ is the O_2 partial pressure (Pa). K_c and K_o are the Michaelis-Menten constants (Pa) for CO_2 and O_2 . These vary with vegetation temperature T_v ($^\circ\text{C}$) according to an Arrhenious function described in [21]. $V_{c,\text{max}}$ is the leaf layer photosynthetic capacity ($\mu \text{ mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$).

The maximum rate of carboxylation allowed by the capacity to regenerate RuBP (i.e., the light-limited rate) w_j ($\mu \text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$) is

$$w_j = \begin{cases} \frac{J(c_i - \Gamma_*)}{4c_i + 8\Gamma_*} & \text{for C}_3 \text{ plants} \\ 4.6\phi\alpha & \text{for C}_4 \text{ plants} \end{cases} \quad c_i - \Gamma_* \geq 0 \quad (75)$$

To find J , the electron transport rate ($\mu \text{ mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$), we solve the following quadratic term and take its smaller root,

$$\Theta_{\text{psII}} J^2 - (I_{\text{psII}} + J_{\text{max}})J + I_{\text{psII}} J_{\text{max}} = 0 \quad (76)$$

where J_{max} is the maximum potential rate of electron transport ($\mu \text{mol m}^{-2} \text{ s}^{-1}$), I_{psII} is the is the light utilized in electron transport by photosystem II ($\mu \text{mol m}^{-2} \text{ s}^{-1}$) and Θ_{psII} is is curvature parameter. I_{psII} is determined as

$$I_{\text{psII}} = 0.5\Phi_{\text{psII}}(4.6\phi) \quad (77)$$

where ϕ is the absorbed photosynthetically active radiation (Wm^{-2}) for either sunlit or shaded leaves (abs_{sun} and abs_{sha}). ϕ is converted to photosynthetic photon flux assuming $4.6 \mu \text{mol}$ photons per joule. Parameter values are $\Phi_{\text{psII}} = 0.7$ for C3 and $\Phi_{\text{psII}} = 0.85$ for C4 plants.

The export limited rate of carboxylation for C3 plants and the PEP carboxylase limited rate of carboxylation for C4 plants w_e (also in $\mu \text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$) is

$$w_e = \begin{cases} 3T_{p,0} & \text{for C}_3 \text{ plants} \\ k_p \frac{c_i}{P_{\text{atm}}} & \text{for C}_4 \text{ plants.} \end{cases} \quad (78)$$

T_p is the triose-phosphate limited rate of photosynthesis, which is equal to $0.167V_{c,\text{max}0}$. k_p is the initial slope of C4 CO_2 response curve. The Michaelis-Menten constants K_c and K_o are modeled as follows,

$$K_c = K_{c,25}(a_{kc})^{\frac{T_v - 25}{10}}, \quad (79)$$

$$K_o = K_{o,25}(a_{ko})^{\frac{T_v - 25}{10}}, \quad (80)$$

where $K_{c,25} = 30.0$ and $K_{o,25} = 30000.0$ are values (Pa) at 25°C , and $a_{kc} = 2.1$ and $a_{ko} = 1.2$ are the relative changes in $K_{c,25}$ and $K_{o,25}$ respectively, for a 10°C change in temperature. The CO_2 compensation point Γ_* (Pa) is

$$\Gamma_* = \frac{1}{2} \frac{K_c}{K_o} 0.21o_i, \quad (81)$$

where the term 0.21 represents the ratio of maximum rates of oxygenation to carboxylation, which is virtually constant with temperature [6].

7.2 Resolution of the photosynthesis theory within the CLM(ED) canopy structure.

The photosynthesis scheme is modified from the CLM4.5 model to give estimates of photosynthesis, respiration and stomatal conductance for a three dimensional matrix indexed by canopy level (C_l), plant functional type (ft) and leaf layer (z). We conduct the photosynthesis calculations at each layer for both sunlit and shaded leaves. Thus, the model also generates estimates of w_c , w_j and w_e indexed in the same three dimensional matrix. In this implementation, some properties (stomatal conductance parameters, top-of-canopy photosynthetic capacity) vary with plant functional type, and some vary with both functional type and canopy depth (absorbed photosynthetically active radiation, nitrogen-based variation in photosynthetic properties). The remaining drivers of photosynthesis (P_{atm} , K_c , o_i , K_o , temperature, atmospheric CO_2) remain the same throughout the canopy.

The rate of gross photosynthesis ($\text{gpp}_{\text{Cl,ft,z}}$) is the smoothed minimum of the three potentially limiting processes (carboxylation, electron transport, export limitation), but calculated independently for each leaf layer:

$$\text{gpp}_{\text{Cl,ft,z}} = \min(w_{\text{c,Cl,ft,z}}, w_{\text{j,Cl,ft,z}}, w_{\text{e,Cl,ft,z}}). \quad (82)$$

For $w_{\text{c,Cl,ft,z}}$, we use

$$w_{\text{c,Cl,ft,z}} = \begin{cases} \frac{V_{\text{c,max,Cl,ft,z}}(c_{\text{i,Cl,ft,z}} - \Gamma_*)}{c_{\text{i,Cl,ft,z}} + K_c(1 + o_i/K_o)} & \text{for } C_3 \text{ plants} \\ V_{\text{c,max,Cl,ft,z}} & \text{for } C_4 \text{ plants} \end{cases} \quad c_{\text{i,Cl,ft,z}} - \Gamma_* \geq 0 \quad (83)$$

where $V_{\text{c,max}}$ now varies with PFT, canopy depth and layer (see below). Internal leaf CO_2 ($c_{\text{i,Cl,ft,z}}$) is tracked separately for each leaf layer. For the light limited rate w_{j} , we use

$$w_{\text{j}} = \begin{cases} \frac{J(c_{\text{i}} - \Gamma_*)4.6\phi\alpha}{4c_{\text{i}} + 8\Gamma_*} & \text{for } C_3 \text{ plants} \\ 4.6\phi\alpha & \text{for } C_4 \text{ plants} \end{cases} \quad (84)$$

where J is calculated as above but based on the absorbed photosynthetically active radiation ($\phi_{\text{Cl,ft,z}}$) for either sunlit or shaded leaves in Wm^{-2} . Specifically,

$$\phi_{\text{Cl,ft,z}} = \begin{cases} \text{abs}_{\text{sun,Cl,ft,z}} & \text{for sunlit leaves} \\ \text{abs}_{\text{sha,Cl,ft,z}} & \text{for shaded leaves} \end{cases} \quad (85)$$

The export limited rate of carboxylation for C_3 plants and the PEP carboxylase limited rate of carboxylation for C_4 plants w_{e} (also in $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$) is calculated in a similar fashion,

$$w_{\text{e,Cl,ft,z}} = \begin{cases} 0.5V_{\text{c,max,Cl,ft,z}} & \text{for } C_3 \text{ plants} \\ 4000V_{\text{c,max,Cl,ft,z}} \frac{c_{\text{i,Cl,ft,z}}}{P_{\text{atm}}} & \text{for } C_4 \text{ plants.} \end{cases} \quad (86)$$

7.2.1 Variation in plant physiology with canopy depth

Both $V_{\text{c,max}}$ and J_{max} vary with vertical depth in the canopy on account of the well-documented reduction in canopy nitrogen through the leaf profile, see [1] for details). Thus, both $V_{\text{c,max}}$ and J_{max} are indexed by C_l , ft and z according to the nitrogen decay coefficient K_n and the amount of vegetation area shading each leaf layer V_{above} ,

$$\begin{aligned} V_{\text{c,max,Cl,ft,z}} &= V_{\text{c,max0,ft}} e^{-K_n \text{ft}} V_{\text{above,Cl,ft,z}}, \\ J_{\text{max,Cl,ft,z}} &= J_{\text{max0,ft}} e^{-K_n \text{ft}} V_{\text{above,Cl,ft,z}}, \end{aligned} \quad (87)$$

where $V_{\text{c,max,0}}$ and $J_{\text{max,0}}$ are the top-of-canopy photosynthetic rates. V_{above} is the sum of exposed leaf area index ($\text{elai}_{\text{Cl,ft,z}}$) and the exposed stem area index ($\text{esai}_{\text{Cl,ft,z}}$) ($\text{m}^2 \text{ m}^{-2}$). Namely,

$$V_{\text{Cl,ft,z}} = \text{elai}_{\text{Cl,ft,z}} + \text{esai}_{\text{Cl,ft,z}}. \quad (88)$$

The vegetation index shading a particular leaf layer in the top canopy layer is equal to

$$V_{\text{above,Cl,ft,z}} = \sum_1^z V_{\text{Cl,ft,z}} \quad \text{for } Cl = 1. \quad (89)$$

For lower canopy layers, the weighted average vegetation index of the canopy layer above (V_{canopy}) is added to this within-canopy shading. Thus,

$$V_{\text{above,Cl,ft,z}} = \sum_1^z V_{\text{Cl,ft,z}} + V_{\text{canopy,Cl-1}} \quad \text{for } Cl > 1, \quad (90)$$

where V_{canopy} is calculated as

$$V_{\text{canopy,Cl}} = \sum_{\text{ft}=1}^{npft} \sum_{z=1}^{nz(\text{ft})} (V_{\text{Cl,ft,z}} \cdot pft_{\text{wt,Cl,ft,1}}). \quad (91)$$

K_n is the coefficient of nitrogen decay with canopy depth. The value of this parameter is taken from the work of [15] who determined, from 204 vertical profiles of leaf traits, that the decay rate of N through canopies of tropical rainforests was a function of the $V_{\text{c,max}}$ at the top of the canopy. They obtain the following term to predict K_n ,

$$K_{n,\text{ft}} = e^{0.00963 V_{\text{c,max0,ft}} - 2.43}, \quad (92)$$

where $V_{\text{c,max}}$ is again in $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$.

7.2.2 Water Stress Function

¹² The top of canopy leaf photosynthetic capacity, $V_{c,\max0}$, is also adjusted for the availability of water to plants as

$$V_{c,\max0,25} = V_{c,\max0,25}\beta_{sw}, \quad (93)$$

where the adjusting factor β_{sw} ranges from one when the soil is wet to zero when the soil is dry. It depends on the soil water potential of each soil layer, the root distribution of the plant functional type, and a plant-dependent response to soil water stress,

$$\beta_{sw} = \sum_{j=1}^{nj} w_j r_j, \quad (94)$$

where w_j is a plant wilting factor for layer j and r_j is the fraction of roots in layer j . The plant wilting factor w_j is

$$w_j = \begin{cases} \frac{\psi_c - \psi_j}{\psi_c - \psi_o} \left(\frac{\theta_{sat,j} - \theta_{ice,j}}{\theta_{sat,j}} \right) & \text{for } T_i > -2^\circ\text{C} \\ 0 & \text{for } T_j \geq -2^\circ\text{C} \end{cases} \quad (95)$$

where ψ_i is the soil water matric potential (mm) and ψ_c and ψ_o are the soil water potential (mm) when stomata are fully closed or fully open, respectively. The term in brackets scales w_i the ratio of the effective porosity (after accounting for the ice fraction) relative to the total porosity. $w_i = 0$ when the temperature of the soil layer (T_i) is below some threshold (-2°C) or when there is no liquid water in the soil layer ($\theta_{liq,i} \leq 0$). For more details on the calculation of soil matric potential, see the CLM4.5 technical note.

7.2.3 Variation of water stress and water uptake within tiles

The remaining drivers of the photosynthesis model remain constant (atmospheric CO_2 and O_2 and canopy temperature) throughout the canopy, except for the water stress index β_{sw} . β_{sw} must be indexed by ft , because plants of differing functional types have the capacity to have varying root depth, and thus access different soil moisture profile and experience differing stress functions. Thus, the water stress function applied to gas exchange calculation is now calculated as

$$\beta_{sw,ft} = \sum_{j=1}^{nj} w_{j,ft} r_{j,ft}, \quad (96)$$

where w_j is the water stress at each soil layer j and $r_{j,ft}$ is the root fraction of each PFT's root mass in layer j . Note that this alteration of the β_{sw} parameter also necessitates recalculation of the vertical water extraction profiles. In the original model, the fraction of extraction from each layer ($r_{e,j,\text{patch}}$) is the product of a single root distribution, because each patch only has one plant functional type. In the CLM(ED), we need to calculate a new weighted patch effective rooting depth profile $r_{e,j,\text{patch}}$ as the weighted average of the functional-type level stress functions and their relative contributions to canopy conductance. Thus for each layer j , the extraction fraction is summed over all PFTs as

$$r_{e,j,\text{patch}} = \sum_{ft=1}^{ft=npft} \frac{w_{j,ft}}{\sum_{j=1}^{nj} w_{j,ft}} \frac{G_{s,ft}}{G_{s,\text{canopy}}}, \quad (97)$$

where nj is the number of soil layers, $G_{s,\text{canopy}}$ is the total canopy (see section 9 for details) and $G_{s,ft}$ is the canopy conductance for plant functional type ft ,

$$G_{s,ft} = \sum_1 w_{ncoh,ft} g_{s,\text{can},coh} n_{coh}. \quad (98)$$

7.2.4 Aggregation of assimilated carbon into cohorts

The derivation of photosynthetic rates per leaf layer, as above, give us the estimated rate of assimilation for a unit area of leaf at a given point in the canopy in $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$. To allow the integration of these rates into fluxes per individual tree, or cohort of trees ($\text{gCO}_2 \text{ tree}^{-1} \text{ s}^{-1}$), they must be multiplied by the amount of leaf

¹²This description relates to algorithms in the ED_btran subroutine

area placed in each layer by each cohort. Each cohort is described by a single functional type, ft and canopy layer C_l flag, so the problem is constrained to integrating these fluxes through the vertical profile (z).

We first make a weighted average of photosynthesis rates from sun (gpp_{sun} , $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$) and shade leaves ($\text{gpp}_{\text{shade}}$, $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$) as

$$\text{gpp}_{\text{Cl,ft,z}} = \text{gpp}_{\text{sun,Cl,ft,z}} f_{\text{sun,Cl,ft,z}} + \text{gpp}_{\text{sha,Cl,ft,z}} (1 - f_{\text{sun,Cl,ft,z}}). \quad (99)$$

The assimilation per leaf layer is then accumulated across all the leaf layers in a given cohort (coh) to give the cohort-specific gross primary productivity (GPP_{coh}),

$$GPP_{\text{coh}} = 12 \times 10^{-9} \sum_{z=1}^{\text{nz}(coh)} \text{gpp}_{\text{Cl,ft,z}} A_{\text{crown,coh}} \text{elai}_{\text{Cl,ft,z}} \quad (100)$$

The $\text{elai}_{\text{Cl,ft,z}}$ is the exposed leaf area which is present in each leaf layer in $\text{m}^2 \text{ m}^{-2}$. (For all the leaf layers that are completely occupied by a cohort, this is the same as the leaf fraction of δ_{vai}). The fluxes are converted from μmol into mol and then multiplied by 12 (the molecular weight of carbon) to give units for GPP_{coh} of $\text{KgC cohort}^{-1} \text{ s}^{-1}$. These are integrated for each timestep to give $\text{KgC cohort}^{-1} \text{ day}^{-1}$.

Table 8: Parameters needed for photosynthesis model.

Parameter Symbol	Parameter Name	Units	indexed by
$V_{\text{c,max0}}$	Maximum carboxylation capacity	$\mu \text{ mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$	ft
r_b	Base Rate of Respiration	$\text{gC gN}^{-1} \text{ s}^{-1}$	
q_{10}	Temp. Response of stem and root respiration		
$R_{\text{cn,leaf,ft}}$	CN ratio of leaf matter	gC/gN	ft
$R_{\text{cn,root,ft}}$	CN ratio of root matter	gC/gN	ft
f_{gr}	Growth Respiration Fraction	none	
ψ_c	Water content when stomata close	Pa	ft
ψ_o	Water content above which stomata are open	Pa	ft

8 Plant respiration

¹³ Plant respiration per individual $R_{\text{plant,coh}}$ ($\text{KgC individual}^{-1} \text{ s}^{-1}$) is the sum of two terms, growth and maintenance respiration $R_{\text{g,coh}}$ and $R_{\text{m,coh}}$

$$R_{\text{plant}} = R_{\text{g,coh}} + R_{\text{m,coh}} \quad (101)$$

Maintenance respiration is the sum of the respiration terms from four different plant tissues, leaf, $R_{\text{m,leaf,coh}}$, fine root $R_{\text{m,froot,coh}}$, coarse root $R_{\text{m,croot,coh}}$ and stem $R_{\text{m,stem,coh}}$, all also in ($\text{KgC individual}^{-1} \text{ s}^{-1}$).

$$R_{\text{m,coh}} = R_{\text{m,leaf,coh}} + R_{\text{m,froot,coh}} + R_{\text{m,croot,coh}} + R_{\text{m,stem,coh}} \quad (102)$$

To calculate canopy leaf respiration, which varies through we canopy, we first determine the top-of-canopy leaf respiration rate ($r_{\text{m,leaf,ft,0}}$, $\text{gC s}^{-1} \text{ m}^{-2}$) is calculated from a base rate of respiration per unit leaf nitrogen derived from [25]. The base rate for leaf respiration (r_b) is $2.525 \text{ gC/gN s}^{-1}$,

$$r_{\text{m,leaf,ft,0}} = r_b N_{\text{a,ft}} (1.5^{(25-20)/10}) \quad (103)$$

where r_b is the base rate of metabolism ($2.525 \times 10^6 \text{ gC/gN s}^{-1}$). This base rate is adjusted assuming a Q_{10} of 1.5 to scale from the baseline of 20C to the CLM default base rate temperature of 25C. For use in the calculations of net photosynthesis and stomatal conductance, leaf respiration is converted from $\text{gC s}^{-1} \text{ m}^{-2}$, into $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$ ($/12 \cdot 10^{-6}$).

This top-of-canopy flux is scaled to account for variation in N_a through the vertical canopy, in the same manner as the $V_{\text{c,max}}$ values are scaled using V_{above} .

$$r_{\text{leaf,Cl,ft,z}} = r_{\text{m,leaf,ft,0}} e^{-K_{\text{n,ft}} V_{\text{above,Cl,ft,z}}} \beta_{\text{ft}} f(t) \quad (104)$$

¹³This description relates to algorithms in the ED_photosynthesis subroutine

Leaf respiration is also adjusted such that it is reduced by drought stress, β_{ft} , and canopy temperature, $f(t_{veg})$. For details of the temperature functions affecting leaf respiration see the CLM4 technical note, Section 8, Equations 8.13 and 8.14. The adjusted leaf level fluxes are scaled to individual-level ($\text{gC individual}^{-1} \text{s}^{-1}$) in the same fashion as the GPP_{coh} calculations

$$R_{m,leaf,coh} = 12 \times 10^{-9} \sum_{z=1}^{nz(coh)} r_{leaf,C1,ft,z} A_{crown} elai_{C1,ft,z} \quad (105)$$

The stem and the coarse-root respiration terms are derived using the same base rate of respiration per unit of tissue Nitrogen.

$$R_{m,croot,coh} = 10^{-3} r_b t_c \beta_{ft} N_{livecroot,coh} \quad (106)$$

$$R_{m,stem,coh} = 10^{-3} r_b t_c \beta_{ft} N_{stem,coh} \quad (107)$$

Here, t_c is a temperature relationship based on a q_{10} value of 1.5, where t_v is the vegetation temperature. We use a base rate of 20 here as, again, this is the baseline temperature used by [25]. The 10^{-3} converts from $\text{gC individual}^{-1} \text{s}^{-1}$ to $\text{KgC individual}^{-1} \text{s}^{-1}$

$$t_c = q_{10}^{(t_v - 20)/10} \quad (108)$$

The tissue N contents for live sapwood are derived from the leaf CN ratios, and for fine roots from the root CN ratio as:

$$N_{stem,coh} = \frac{B_{sapwood,coh}}{R_{cn,leaf,ft}} \quad (109)$$

and

$$N_{livecroot,coh} = \frac{B_{root,coh} w_{frac,ft}}{R_{cn,root,ft}} \quad (110)$$

where $B_{sapwood,coh}$ and $B_{root,coh}$ are the biomass pools of sapwood and live root biomass respectively (KgC individual) and $w_{frac,ft}$ is the fraction of coarse root tissue in the root pool (0.5 for woody plants, 0.0 for grasses and crops). We assume here that stem CN ratio is the same as the leaf C:N ratio, for simplicity. The final maintenance respiration term is derived from the fine root respiration, which accounts for gradients of temperature in the soil profile and thus calculated for each soil layer j as follows:

$$R_{m,root,j} = \frac{(1 - w_{frac,ft}) B_{root,coh} b_r \beta_{ft}}{10^3 R_{cn,leaf,ft}} \sum_{j=1}^{nj} t_{c,soi,j} r_{i,ft,j} \quad (111)$$

$t_{c,soi}$ is a function of soil temperature in layer j that has the same form as that for stem respiration, but uses vertically resolved soil temperature instead of canopy temperature. In the CLM4.5, only coarse and not fine root respiration varies as a function of soil depth, and we maintain this assumption here, although it may be altered in later versions. The growth respiration, $R_{g,coh}$ is a fixed fraction f_{gr} of the carbon remaining after maintenance respiration has occurred.

$$R_{g,coh} = \max(0, \text{GPP}_{g,coh} - R_{m,coh}) f_{gr} \quad (112)$$

Table 9: Parameters needed for plant respiration model.

Parameter Symbol	Parameter Name	Units	indexed by
$-K_{n,ft}$	Rate of reduction of N through the canopy	none	-
r_b	Base Rate of Respiration	$\text{gC gN}^{-1} \text{s}^{-1}$	-
q_{10}	Temp. Response of stem and root respiration	-	-
$R_{cn,leaf,ft}$	CN ratio of leaf matter	gC/gN	ft
$R_{cn,root,ft}$	CN ratio of root matter	gC/gN	ft
f_{gr}	Growth Respiration Fraction	none	-

9 Stomatal Conductance

9.0.5 Fundamental stomatal conductance theory

¹⁴ Stomatal conductance is unchanged in concept from the CLM4.5 approach. Leaf stomatal resistance is calculated from the Ball-Berry conductance model as described by [4] and implemented in a global climate model by [28]. The model relates stomatal conductance (i.e., the inverse of resistance) to net leaf photosynthesis, scaled by the relative humidity at the leaf surface and the CO₂ concentration at the leaf surface. The primary difference between the CLM implementation and that used by [4] and [28] is that they used net photosynthesis (i.e., leaf photosynthesis minus leaf respiration) instead of gross photosynthesis. As implemented here, stomatal conductance equals the minimum conductance (b) when gross photosynthesis (A) is zero. Leaf stomatal resistance is

$$\frac{1}{r_s} = m_{ft} \frac{A}{c_s} \frac{e_s}{e_i} P_{atm} + b_{ft} \beta_{sw} \quad (113)$$

where r_s is leaf stomatal resistance ($s\ m^2\ \mu\text{mol}^{-1}$), b_{ft} is a plant functional type dependent parameter equivalent to g_0 in the Ball-Berry model literature. This parameter is also scaled by the water stress index β_{sw} . Similarly, m_{ft} is the slope of the relationship between the assimilation, c_s and humidity dependant term and the stomatal conductance, and so is equivalent to the g_1 term in the stomatal literature. A is leaf photosynthesis ($\mu\text{mol}\ \text{CO}_2\ \text{m}^{-2}\ \text{s}^{-1}$), c_s is the CO₂ partial pressure at the leaf surface (Pa), e_s is the vapor pressure at the leaf surface (Pa), e_i is the saturation vapor pressure (Pa) inside the leaf at the vegetation temperature conductance ($\mu\text{mol}\ \text{m}^{-2}\ \text{s}^{-1}$) when $A = 0$. Typical values are $m_{ft} = 9$ for C₃ plants and $m_{ft} = 4$ for C₄ plants [4, 5, 28]. [28] used $b = 10000$ for C₃ plants and $b = 40000$ for C₄ plants. Here, b was chosen to give a maximum stomatal resistance of $20000\ \text{s}\ \text{m}^{-1}$. These terms are nevertheless plant strategy dependent, and have been found to vary widely with plant type [17].

Resistance is converted from units of $s\ \text{m}^2\ \mu\text{mol}^{-1}$ to $s\ \text{m}^{-1}$ as: $1\ \text{s}\ \text{m}^{-1} = 1 \times 10^{-9} R_{gas} \theta_{atm} P_{atm} (\mu\text{mol}^{-1}\ \text{m}^2\ \text{s})$, where R_{gas} is the universal gas constant ($\text{J}\ \text{K}^{-1}\ \text{kmol}^{-1}$) and θ_{atm} is the atmospheric potential temperature (K).

9.0.6 Resolution of the stomatal conductance theory in the CLM(ED) canopy structure

The stomatal conductance is calculated, as with photosynthesis, for each canopy, PFT and leaf layer. The CLM code requires a single canopy conductance estimate to be generated from the multi-layer multi-PFT array. In previous iterations of the CLM, sun and shade-leaf specific values have been reported and then averaged by their respective leaf areas. In this version, the total canopy conductance $G_{s,canopy}$, is calculated as the sum of the cohort-level conductance values.

$$G_{s,canopy} = \sum \frac{gs_{can,coh} n_{coh}}{A_{patch}} \quad (114)$$

Cohort conductance is the sum of the inverse of the leaf resistances at each canopy layer ($r_{s,z}$) multiplied by the area of each cohort.

$$gs_{can,coh} = \sum_{z=1}^{z=nv,coh} \frac{A_{crown,coh}}{r_{s,cl,ft,z} + r_b} \quad (115)$$

Table 10: Parameters needed for stomatal conductance model.

Parameter Symbol	Parameter Name	Units	indexed by
b_{ft}	Slope of Ball-Berry term	none	<i>ft</i>
m_{ft}	Slope of Ball-Berry term	none	<i>ft</i>

10 Allocation

¹⁵ Total assimilation carbon enters the ED model each day as a cohort-specific Net Primary Productivity NPP_{coh} , which is calculated as

$$NPP_{coh} = GPP_{coh} - R_{plant,coh} \quad (116)$$

¹⁴This description relates to algorithms in the ED_photosynthesis subroutine

¹⁵This description relates to algorithms in the Growth_Derivatives subroutine

This flux of carbon is allocated between the demands of tissue turnover, of carbohydrate storage and of growth (increase in size of one or many plant organs). Priority is explicitly given to maintenance respiration, followed by tissue maintenance and storage, then allocation to live biomass and then to the expansion of structural and live biomass pools. All fluxes here are first converted into in KgC individual⁻¹ year⁻¹ and ultimately integrated using a timesteps of 1/365 years for each day.

10.1 Tissue maintenance demand

We calculate a ‘tissue maintenance’ flux. The magnitude of this flux is such that the quantity of biomass in each pool will remain constant, given background turnover rates. For roots, this maintenace demand is simply

$$r_{\text{md,coh}} = b_{\text{root}} \cdot \alpha_{\text{root,ft}} \quad (117)$$

Where $\alpha_{\text{root,ft}}$ is the root turnover rate in y⁻¹. Given that, for deciduous trees, loss of leaves is assumed to happen only one per growing season, the algorithm is dependent on phenological habit (whether or not this PFT is evergreen), thus

$$l_{\text{md,coh}} = \begin{cases} b_{\text{leaf}} \cdot \alpha_{\text{leaf,ft}} & \text{for } P_{\text{evergreen}} = 1 \\ 0 & \text{for } P_{\text{evergreen}} = 0 \end{cases} \quad (118)$$

Leaf litter resulting from deciduous senescence is handled in the phenology section. The total quantity of maintenance demand ($t_{\text{md,coh}}$ KgC individual y⁻¹) is therefore

$$t_{\text{md,coh}} = l_{\text{md,coh}} + r_{\text{md,coh}} \quad (119)$$

10.2 Allocation to storage and turnover

The model must now determine whether the NPP input is sufficient to meet the maintenance demand and keep tissue levels constant. To determine this, we introduce the idea of ‘carbon balance’ $C_{\text{bal,coh}}$ (KgC individual⁻¹) where

$$C_{\text{bal,coh}} = \text{NPP}_{\text{coh}} - t_{\text{md,coh}} \cdot f_{\text{md,min,ft}} \quad (120)$$

where $f_{\text{md,min,ft}}$ is the minimum fraction of the maintenance demand that the plant must meet each timestep, which is indexed by ft and represents a life-history-strategy decision concerning whether leaves should remain on in the case of low carbon uptake (a risky strategy) or not be replaced (a conservative strategy). Subsequently, we determine a flux to the storage pool, where the flux into the pool, as a fraction of $C_{\text{bal,coh}}$, is proportional to the discrepancy between the target pool size and the actual pool size f_{tstore} where

$$f_{\text{tstore}} = \max \left(0, \frac{b_{\text{store}}}{b_{\text{leaf}} \cdot S_{\text{cushion}}} \right) \quad (121)$$

The allocation to storage is a fourth power function of f_{tstore} to mimic the qualitative behaviour found for carbon allocation in arabidopsis by [30].

$$\frac{\delta b_{\text{store}}}{\delta t} = \begin{cases} C_{\text{bal,coh}} \cdot e^{-f_{\text{tstore}}^4} & \text{for } C_{\text{bal,coh}} > 0 \\ C_{\text{bal,coh}} & \text{for } C_{\text{bal,coh}} \leq 0 \end{cases} \quad (122)$$

If the carbon remaining after the storage and minimum turnover fluxes have been met, the next priority is the remaining flux to leaves $t_{\text{md}} \cdot (1 - f_{\text{md,min}})$. If the quantity of carbon left ($C_{\text{bal,coh}} - \frac{\delta b_{\text{store}}}{\delta t}$) is insufficient to supply this amount of carbon, then the store of alive carbon is depleted (to represent those leaves that have fallen off and not been replaced)

$$\frac{\delta b_{\text{alive}}}{\delta t} = \begin{cases} 0 & \text{for } (C_{\text{bal,coh}} - \frac{\delta b_{\text{store}}}{\delta t}) > t_{\text{md}} \cdot (1 - f_{\text{md,min}}) \\ t_{\text{md}} \cdot (1 - f_{\text{md,min}}) - (C_{\text{bal,coh}} - \frac{\delta b_{\text{store}}}{\delta t}) & \text{for } (C_{\text{bal,coh}} - \frac{\delta b_{\text{store}}}{\delta t}) \leq t_{\text{md}} \cdot (1 - f_{\text{md,min}}) \end{cases} \quad (123)$$

correspondingly, the carbon left over for growth (C_{growth} (KgC individual⁻¹ year⁻¹) is therefore

$$C_{\text{growth}} = \begin{cases} C_{\text{bal,coh}} - \frac{\delta b_{\text{store}}}{\delta t} & \text{for } (C_{\text{bal,coh}} - \frac{\delta b_{\text{store}}}{\delta t}) > 0 \\ 0 & \text{for } (C_{\text{bal,coh}} - \frac{\delta b_{\text{store}}}{\delta t}) \leq 0 \end{cases} \quad (124)$$

to allocate the remaining carbon (if there is any), we first ascertain whether the live biomass pool is at its target, or whether it has been depleted by previous low carbon timesteps. Thus

$$\begin{aligned} b_{\text{alive,target}} &= b_{\text{leaf,target}}(1 + f_{\text{firla}} + f_{\text{sw}}h_{\text{coh}}) & \text{for } S_{\text{phen,coh}} = 2 \\ b_{\text{alive,target}} &= b_{\text{leaf,target}}(f_{\text{firla}} + f_{\text{sw}}h_{\text{coh}}) & \text{for } S_{\text{phen,coh}} = 1 \end{aligned} \quad (125)$$

where the target leaf biomass $b_{\text{leaf,target}}$ ((Kg C individual⁻¹)) is the allometric relationship between dbh and leaf biomass, ameliorated by the leaf trimming fraction (see ‘control of leaf area’ below)

$$b_{\text{leaf,target}} = c_{\text{leaf}} \cdot dbh_{\text{coh}}^{e_{\text{leaf,dbh}}} \rho_{\text{ft}}^{e_{\text{leaf,dens}}} \cdot C_{\text{trim,coh}} \quad (126)$$

ρ_{ft} is the wood density, in g cm³.

10.2.1 Allocation to Seeds

The fraction remaining for growth (expansion of live and structural tissues) f_{growth} is 1 minus that allocated to seeds.

$$f_{\text{growth,coh}} = 1 - f_{\text{seed,coh}} \quad (127)$$

Allocation to seeds only occurs if the alive biomass is not below its target, and then is a predefined fixed fraction of the carbon remaining for growth. Allocation to clonal reproduction (primarily for grasses) occurs when \max_{dbh} is achieved.

$$f_{\text{seed,coh}} = \begin{cases} R_{\text{frac,ft}} & \text{for } \max_{\text{dbh}} < dbh_{\text{coh}} \\ (R_{\text{frac,ft}} + C_{\text{frac,ft}}) & \text{for } \max_{\text{dbh}} \geq dbh_{\text{coh}} \end{cases} \quad (128)$$

the total amount allocated to seed production ($p_{\text{seed,coh}}$ in KgC individual⁻¹ y⁻¹) is thus

$$p_{\text{seed,coh}} = C_{\text{growth}} \cdot f_{\text{seed,coh}} \quad (129)$$

10.2.2 Allocation to growing pools

¹⁶ The carbon is then partitioned into carbon available to grow the b_{alive} and b_{struc} pools. A fraction v_a is available to live biomass pools, and a fraction v_s is available to structural pools.

$$\frac{\delta b_{\text{alive}}}{\delta t} = C_{\text{growth}} \cdot f_{\text{growth}} v_a \quad (130)$$

$$\frac{\delta b_{\text{struc}}}{\delta t} = C_{\text{growth}} \cdot f_{\text{growth}} v_s \quad (131)$$

If the alive biomass is lower than its ideal target, all of the available carbon is directed into that pool. Thus:

$$v_a = \begin{cases} \frac{1}{1+u} & \text{for } b_{\text{alive}} \geq b_{\text{alive,target}} \\ 1.0 & \text{for } b_{\text{alive}} < b_{\text{alive,target}} \end{cases} \quad (132)$$

$$v_s = \begin{cases} \frac{u}{1+u} & \text{for } b_{\text{alive}} \geq b_{\text{alive,target}} \\ 0.0 & \text{for } b_{\text{alive}} < b_{\text{alive,target}} \end{cases} \quad (133)$$

In this case, the division of carbon between the live and structural pools u is derived as the inverse of the sum of the rates of change in live biomass with respect to structural:

$$u = \frac{1}{\frac{\delta b_{\text{leaf}}}{\delta b_{\text{struc}}} + \frac{\delta b_{\text{root}}}{\delta b_{\text{struc}}} + \frac{\delta b_{\text{sw}}}{\delta b_{\text{struc}}}} \quad (134)$$

To calculate all these differentials, we first start with $\delta b_{\text{leaf}}/\delta b_{\text{struc}}$, where

$$\frac{\delta b_{\text{leaf}}}{\delta b_{\text{struc}}} = \frac{\frac{\delta dbh}{\delta b_{\text{struc}}}}{\frac{\delta dbh}{\delta b_{\text{leaf}}}} \quad (135)$$

¹⁶This description relates to algorithms in the ED_GrowthFunctions subroutine

The rates of change of dbh with respect to leaf and structural biomass are the differentials of the allometric equations linking these terms to each other. Hence,

$$\frac{\delta \text{dbh}}{\delta b_{\text{leaf}}} = \frac{1}{b_{\text{trim,coh}}} \cdot (e_{\text{leaf,dbh}} - 1) \exp \left(c_{\text{leaf}} \text{dbh}^{(e_{\text{leaf,dbh}} - 1)} \rho_{\text{ft}}^{e_{\text{leaf,dens}}} \right) \quad (136)$$

and where $\text{dbh}_{\text{coh}} > \text{dbh}_{\text{max}}$

$$\frac{\delta b_{\text{struc}}}{\delta \text{dbh}} = e_{\text{str,dbh}} \cdot c_{\text{str}} \cdot e_{\text{str,hite}} h_{\text{coh}}^{e_{\text{str,dbh}} - 1} \text{dbh}_{\text{coh}}^{e_{\text{str,dbh}}} \rho_{\text{ft}}^{e_{\text{str,dens}}} \quad (137)$$

If $\text{dbh}_{\text{coh}} \leq \text{dbh}_{\text{max}}$ then we must also account for allocation for growing taller as:

$$\frac{\delta b_{\text{struc}}}{\delta \text{dbh}} = \frac{\delta b_{\text{struc}}}{\delta \text{dbh}} + \frac{\delta h}{\delta \text{dbh}} \cdot \frac{\delta b_{\text{struc}}}{\delta \text{dbh}} \quad (138)$$

where

$$\frac{\delta h}{\delta \text{dbh}} = 1.4976 \text{dbh}_{\text{coh}}^{\text{m}_{\text{allom}} - 1} \quad (139)$$

$$\frac{\delta \text{dbh}}{\delta b_{\text{struc}}} = \frac{1}{\frac{\delta b_{\text{struc}}}{\delta \text{dbh}}} \quad (140)$$

Once we have the $\delta b_{\text{leaf}}/\delta b_{\text{struc}}$, we calculate $\delta b_{\text{root}}/\delta b_{\text{struc}}$ as

$$\frac{\delta b_{\text{root}}}{\delta b_{\text{struc}}} = \frac{\delta b_{\text{leaf}}}{\delta b_{\text{struc}}} \cdot f_{\text{frla}} \quad (141)$$

and the sapwood differential as

$$\frac{\delta b_{\text{sw}}}{\delta b_{\text{struc}}} = f_{\text{swh}} \left(h_{\text{coh}} \frac{\delta b_{\text{leaf}}}{\delta b_{\text{struc}}} + b_{\text{leaf,coh}} \frac{\delta h}{\delta b_{\text{struc}}} \right) \quad (142)$$

where

$$\frac{\delta h}{\delta b_{\text{struc}}} = \frac{1}{c_{\text{str}} \times e_{\text{str,hite}} h_{\text{coh}}^{e_{\text{str,dbh}} - 1} \text{dbh}_{\text{coh}}^{e_{\text{str,dbh}}} \rho_{\text{ft}}^{e_{\text{str,dens}}}} \quad (143)$$

In all of the above terms, height in m, dbh is in cm, and all biomass pools are in KgCm^{-2} . The allometric terms for the growth trajectory are all taken from the ED1.0 model, but could in theory be altered to accomodate alternative allometric relationships. Critically, the non-linear relationships between live and structural biomass pools are maintained in this algorithm, which diverges from the methodology currently deployed in the CLM4.5.

10.3 Integration of allocated fluxes

All of the flux calculations generate differential of the biomass state variables against time (in years). To integrate these differential rates into changes in the state variables, we use a simple forward Euler integration. Other methods exist (e.g. ODEINT solvers, Runge Kutta methods etc.), but they are more prone to errors that become difficult to diagnose, and the typically slow rates of change of carbon pools mean that these are less important than they might be in strongly non-linear systems (soil drainage, energy balance, etc.)

$$b_{\text{alive},t+1} = \min \left(0, b_{\text{alive},t} + \frac{\delta b_{\text{alive}}}{\delta t} \delta t \right) \quad (144)$$

$$b_{\text{struc},t+1} = \min \left(0, b_{\text{struc},t} + \frac{\delta b_{\text{struc}}}{\delta t} \delta t \right) \quad (145)$$

$$b_{\text{store},t+1} = \min \left(0, b_{\text{store},t} + \frac{\delta b_{\text{store}}}{\delta t} \delta t \right) \quad (146)$$

In this case, δt is set to be one day ($\frac{1}{365}$ years).

Table 11: Parameters needed for allocation model.

Parameter Symbol	Parameter Name	Units	indexed by
S_{cushion}	Target stored biomass as fraction of b_{leaf}	none	<i>ft</i>
$f_{\text{md,min}}$	Minimum fraction of turnover that must be met	none	<i>ft</i>
$R_{\text{frac,ft}}$	Fraction allocated to seeds	none	<i>ft</i>
$C_{\text{frac,ft}}$	Fraction allocated to clonal reproduction	none	<i>ft</i>
max_{dbh}	Diameter at which maximum height is achieved	m	<i>ft</i>
$P_{\text{evergreen,ft}}$	Does this cohort have an evergreen phenological habit?	1=yes, 0=no	<i>ft</i>

11 Control of Leaf Area Index

¹⁷ The leaf area A_{leaf} (m^2) of each cohort is calculated from leaf biomass $b_{\text{leaf,coh}}$ ($\text{kgC individual}^{-1}$) and specific leaf area (SLA, $\text{m}^2 \text{ kg C}^{-1}$)

$$A_{\text{leaf,coh}} = b_{\text{leaf,coh}} \cdot \text{SLA}_{\text{ft}} \quad (147)$$

For a given tree allometry, leaf biomass is determined from basal area using the function used by [19] where d_w is wood density in g cm^{-3} .

$$b_{\text{leaf,coh}} = c_{\text{leaf}} \cdot dbh_{\text{coh}}^{e_{\text{leaf,dbh}}} \rho_{\text{ft}}^{e_{\text{leaf,dens}}} \quad (148)$$

However, using this model, where leaf area and crown area are both functions of diameter, the leaf area index of each tree in a closed canopy forest is always the same (where $S_{\text{c,patch}} = S_{\text{c,min}}$, irrespective of the growth conditions. To allow greater plasticity in tree canopy structure, and for tree leaf area index to adapt to prevailing conditions, we implemented a methodology for removing those leaves in the canopy that exist in negative carbon balance. That is, their total annual assimilation rate is insufficient to pay for the turnover and maintenance costs associated with their supportive root and stem tissue, plus the costs of growing the leaf. The tissue turnover maintenance cost ($\text{KgC m}^{-2} \text{y}^{-1}$ of leaf is the total maintenance demand divided by the leaf area:

$$L_{\text{cost,coh}} = \frac{t_{\text{md,coh}}}{b_{\text{leaf,coh}} \cdot \text{SLA}} \quad (149)$$

The net uptake for each leaf layer $U_{\text{net,z}}$ in ($\text{KgC m}^{-2} \text{ year}^{-1}$) is

$$U_{\text{net,coh,z}} = g_{\text{coh,z}} - r_{\text{m,leaf,coh,z}} \quad (150)$$

where g_z is the GPP of each layer of leaves in each tree ($\text{KgC m}^{-2} \text{ year}^{-1}$), $r_{\text{m,leaf,z}}$ is the rate of leaf dark respiration (also $\text{KgC m}^{-2} \text{ year}^{-1}$). We use an iterative scheme to define the cohort specific canopy trimming fraction $C_{\text{trim,coh}}$, on an annual time-step, where

$$b_{\text{leaf,coh}} = C_{\text{trim}} \times 0.0419 dbh_{\text{coh}}^{1.56} d_w^{0.55} \quad (151)$$

If the annual maintenance cost of the bottom layer of leaves ($\text{KgC m}^{-2} \text{ year}^{-1}$) is less than then the canopy is trimmed by an increment ι_l (0.01), which is applied until the end of the next calander year. Because this is an optimality model, there is an issue of the timescale over which net assimilation is evaluated, the timescale of response, and the plasticity of plants to respond to these pressures. These properties should be investigated further in future efforts.

$$C_{\text{trim,y+1}} = \begin{cases} \max(C_{\text{trim,y}} - \iota_l, 1.0) & \text{for}(L_{\text{cost,coh}} > U_{\text{net,coh,nz}}) \\ \min(C_{\text{trim,y}} + \iota_l, L_{\text{trim,min}}) & \text{for}(L_{\text{cost,coh}} < U_{\text{net,coh,nz}}) \end{cases} \quad (152)$$

We impose an arbitrary minimum value on the scope of canopy trimming of $L_{\text{trim,min}}$ (0.5). If plants are able simply to drop all of their canopy in times of stress, with no consequences, then tree mortality from carbon starvation is much less likely to occur because of the greatly reduced maintenance and turnover requirements.

Table 12: Parameters needed for leaf area control model.

Parameter Symbol	Parameter Name	Units	indexed by
ι_l	Fraction by which leaf mass is reduced next year	none	-
$L_{\text{trim,min}}$	Minimum fraction to which leaf mass can be reduced	-	-

¹⁷This description relates to algorithms in the trim_canopy subroutine

12 Phenology

12.1 Cold Deciduous Phenology

12.1.1 Cold Leaf-out timing

¹⁸. The phenology model of [2] is used in CLM(ED) to determine the leaf-on timing. The Botta et al. model was verified against satellite data and is one of the only globally verified and published models of leaf-out phenology. This model differs from the phenology model in the CLM4.5. The model simulates leaf-on date as a function of the number of growing degree days (GDD), defined by the sum of mean daily temperatures (T_{day} °C) above a given threshold T_g (0 °C).

$$GDD = \sum \max(T_{\text{day}} - T_g, 0) \quad (153)$$

Budburst occurs when GDD exceeds a threshold (GDD_{crit}). The threshold is modulated by the number of chilling days experienced (NCD) where the mean daily temperature falls below a threshold determined by [2] as 5°C. A greater number of chilling days means that fewer growing degree days are required before budburst:

$$GDD_{\text{crit}} = a + be^{c \cdot \text{NCD}} \quad (154)$$

where $a = -68$, $b = 638$ and $c = -0.01$ [2]. In the Northern Hemisphere, counting of degree days begins on 1st January, and of chilling days on 1st November. The calendar opposite of these dates is used for points in the Southern Hemisphere.

If the growing degree days exceed the critical threshold, leaf-on is triggered by a change in the gridcell phenology status flag $S_{\text{phen,grid}}$ where ‘2’ indicates that leaves should come on and ‘1’ indicates that they should fall.

$$S_{\text{phen,grid}} = 2 \quad \text{if } S_{\text{phen,grid}} = 1 \text{ and } GDD_{\text{grid}} \geq GDD_{\text{crit}} \quad (155)$$

12.1.2 Cold Leaf-off timing

The leaf-off model is taken from the Sheffield Dynamic Vegetation Model (SDGVM) and is similar to that for LPJ [29] and IBIS [8] models. The average daily temperatures of the previous 10 day period are stored. Senescence is triggered when the number of days with an average temperature below 7.5° (n_{colddays}) rises above a threshold values $n_{\text{crit,cold}}$, set at 5 days.

$$S_{\text{phen,grid}} = 1 \quad \text{if } S_{\text{phen,grid}} = 2 \text{ and } n_{\text{colddays}} \geq n_{\text{crit,cold}} \quad (156)$$

12.1.3 Global implementation modifications

Because of the global implementation of the cold-deciduous phenology scheme, adjustments must be made to account for the possibility of cold-deciduous plants experiencing situations where no chilling period triggering leaf-off ever happens. If left unaccounted for, these leaves will last indefinitely, resulting in highly unrealistic behaviour. Therefore, we implement two additional rules. Firstly, if the number of days since the last senescence event was triggered is larger than 364, then leaf-off is triggered on that day. Secondly, if no chilling days have occurred during the winter accumulation period, then leaf-on is not triggered. This means that in effect, where there are no cold periods, leaves will fall off and not come back on, meaning that cold-deciduous plants can only grow in places where there is a cold season.

Further to this rule, we introduce a ‘buffer’ time periods after leaf-on of 30 days, so that cold-snap periods in the spring cannot trigger a leaf senescence. The 30 day limit is an arbitrary limit. In addition, we constrain growing degree day accumulation to the second half of the year (Jult onwards in the Northern hemisphere, or Jan-June in the Southern) and only allow GDD accumulation while the leaves are off.

12.2 Drought-deciduous Phenology

In the current version of the model, a drought deciduous algorithm exists, but is not yet operational, due to issue detected in the existing CN and soil moisture modules, which also affect the behaviour of the native ED drought deciduous model. This is a priority to address before the science tag is released.

¹⁸This description relates to algorithms in the phenology subroutine

12.3 Carbon Dynamics of deciduous plants

¹⁹ In the present version, leaf expansion and senescence happen over the course of a single day. This is clearly not an empirically robust representation of leaf behaviour, whereby leaf expansion occurs over a period of 10-14 days, and senescence over a similar period. This will be incorporated in later versions. When the cold or drought phenological status of the gridcell status changes ($S_{\text{phen,grid}}$) from ‘2’ to ‘1’, and the leaves are still on ($S_{\text{phen,coh}}=2$), the leaf biomass at this timestep is ‘remembered’ by the model state variable $l_{\text{memory,coh}}$. This provides a ‘target’ biomass for leaf onset at the beginning of the next growing season (it is a target, since depletion of stored carbon in the off season may render achieving the target impossible).

$$l_{\text{memory,coh}} = b_{\text{leaf,coh}} \quad (157)$$

Leaf carbon is then added to the leaf litter flux $l_{\text{leaf,coh}}$ (KgC individual⁻¹)

$$l_{\text{leaf,coh}} = b_{\text{leaf,coh}} \quad (158)$$

The alive biomass is depleted by the quantity of leaf mass lost, and the leaf biomass is set to zero

$$b_{\text{alive,coh}} = b_{\text{alive,coh}} - b_{\text{leaf,coh}} \quad (159)$$

$$b_{\text{leaf,coh}} = 0 \quad (160)$$

Finally, the status $S_{\text{phen,coh}}$ is set to 1, indicating that the leaves have fallen off.

For bud burst, or leaf-on, the same occurs in reverse. If the leaves are off ($S_{\text{phen,coh}}=1$) and the phenological status triggers budburst ($S_{\text{phen,grid}}=2$) then the leaf mass is set the maximum of the leaf memory and the available store

$$b_{\text{leaf,coh}} = \max(l_{\text{memory,coh}}, b_{\text{store,coh}}) \quad (161)$$

this amount of carbon is removed from the store

$$b_{\text{store,coh}} = b_{\text{store,coh}} - b_{\text{leaf,coh}} \quad (162)$$

and the new leaf biomass is added to the alive pool

$$b_{\text{alive,coh}} = b_{\text{alive,coh}} + b_{\text{leaf,coh}} \quad (163)$$

Lastly, the leaf memory variable is set to zero and the phenological status of the cohort back to ‘2’. No parameters are currently required for this carbon accounting scheme.

Table 13: Parameters needed for phenology model.

Parameter Symbol	Parameter Name	Units	indexed by
$n_{\text{crit,cold}}$	Threshold of cold days for senescence	none	-
T_g	Threshold for counting growing degree days	°C	

13 Seed Dynamics and Recruitment

²⁰ The production of seeds and their subsequent germination is a process that must be captured explicitly or implicitly in vegetation models. The CLM(ED) contains a seed bank model designed to allow the dynamics of seed production and germination to be simulated independently. In the ED1.0 model, seed recruitment occurs in the same timestep as allocation to seeds, which prohibits the survival of a viable seed bank through a period of disturbance or low productivity (winter, drought). In the CLM(ED) model, a plant functional type specific seed bank is tracked in each patch ($Seeds_{\text{patch}}$ KgC m⁻²), whose rate of change (KgC m⁻² y⁻¹) is the balance of inputs, germination and decay:

$$\frac{\delta Seeds_{\text{FT}}}{\delta t} = Seed_{\text{in,ft}} - Seed_{\text{germ,ft}} - Seed_{\text{decay,ft}} \quad (164)$$

¹⁹This description relates to algorithms in the phenology_leafoff subroutine

²⁰This description relates to algorithms in the seed_in, seed_decay and seed_germination subroutines

where $Seed_{in}$, $Seed_{germ}$ and $Seed_{decay}$ are the production, germination and decay (or onset of inviability) of seeds, all in $\text{KgC m}^{-2} \text{ year}^{-1}$.

Seeds are assumed to be distributed evenly across the site (in this version of the model), so the total input to the seed pool is therefore the sum of all of the reproductive output of all the cohorts in each patch of the correct PFT type.

$$Seed_{in,ft} = \frac{\sum_{p=1}^{n_{\text{patch}}} \sum_{i=1}^{n_{\text{coh}}} p_{\text{seed},i} \cdot n_{\text{coh}}}{area_{\text{site}}} \quad (165)$$

Seed decay is the sum of all the processes that reduce the number of seeds, taken from [14]. Firstly, the rate at which seeds become inviable is described as a constant rate ϕ (y^{-1}) which is set to 0.51, the mean of the parameters used by [14].

$$Seed_{decay,ft} = Seeds_{FT} \cdot \phi \quad (166)$$

The seed germination flux is also prescribed as a fraction of the existing pool (α_{sgerm}), but with a cap on maximum germination rate β_{sgerm} , to prevent excessive dominance of one plant functional type over the seed pool.

$$Seed_{germ,ft} = \max(Seeds_{FT} \cdot \alpha_{sgerm}, \beta_{sgerm}) \quad (167)$$

Table 14: Parameters needed for seed model.

Parameter Symbol	Parameter Name	Units	indexed by
K_s	Maximum seed mass	kgC m^{-2}	
α_{sgerm}	Proportional germination rate	-	
β_{sgerm}	Maximum germination rate	$\text{KgC m}^{-2} \text{ y}^{-1}$	
ϕ	Decay rate of viable seeds	none	FT
$R_{frac,ft}$	Fraction of C_{bal} devoted to reproduction	none	FT

14 Fragmenting Litter Pool Production and Turnover

The original CLM4.5 model contains streams of carbon pertaining to different chemical properties of litter (lignin, cellulose and labile streams, specifically). In the CLM(ED) model, the fire simulation scheme in the SPITFIRE model requires that the model tracks the pools of litter pools that differ with respect to their propensity to burn (surface area-volume ratio, bulk density etc.). Therefore, this model contains more complexity in the representation of coarse woody debris. We also introduce the concept of 'fragmenting' pools, which are pools that can be burned, but are not available for decomposition or respiration. In this way, we can both maintain above-ground pools that affect the rate of burning, and the lag between tree mortality and availability of woody material for decomposition.

The CLM(ED) recognizes four classes of litter. Above- and below-ground coarse woody debris (CWD_{AG} , CWD_{BG}) and leaf litter (l_{leaf} and fine root litter l_{root}). All pools are represented per patch, and with units of kGC m^{-2} . Further to this, CWD_{AG} , CWD_{BG} are split into four litter size classes (lsc) for the purposes of proscribing this to the SPITFIRE fire model (see 'Fuel Load' section for more detail. 1-hour (twigs), 10-hour (small branches), 100-hour (large branches) and 1000-hour (boles or trunks). 4.5 %, 7.5%, 21 % and 67% of the woody biomass ($b_{store,coh} + b_{sw,coh}$) is partitioned into each class, respectively.

l_{leaf} and l_{root} are indexed by plant functional type (ft). The rationale for indexing leaf and fine root by PFT is that leaf and fine root matter typically vary in their carbon:nitrogen ratio, whereas woody pools typically do not.

Rates of change of litter, all in $\text{kGC m}^{-2} \text{ year}^{-1}$, are calculated as

$$\frac{\delta CWD_{AG,out,lsc}}{\delta t} = CWD_{AG,in,lsc} - CWD_{AG,out,lsc} \quad (168)$$

$$\frac{\delta CWD_{BG,out,lsc}}{\delta t} = CWD_{BG,in,lsc} - CWD_{BG,out,lsc} \quad (169)$$

$$\frac{\delta l_{leaf,out,ft}}{\delta t} = l_{leaf,in,ft} - l_{leaf,out,ft} \quad (170)$$

$$\frac{\delta l_{\text{root,out,ft}}}{\delta t} = l_{\text{root,in,ft}} - l_{\text{root,out,ft}} \quad (171)$$

14.1 Litter Inputs

²¹ Inputs into the litter pools come from tissue turnover, mortality of canopy trees, mortality of understorey trees, mortality of seeds, and leaf senescence of deciduous plants.

$$l_{\text{leaf,in,ft}} = \left(\sum_{i=1}^{n_{\text{coh,ft}}} n_{\text{coh}} (l_{\text{md,coh}} + l_{\text{leaf,coh}}) + M_{\text{t,coh}} \cdot b_{\text{leaf,coh}} \right) / \sum_{p=1}^{n_{\text{pat}}} A_{\text{patch}} \quad (172)$$

where $l_{\text{md,coh}}$ is the leaf turnover rate for evergreen trees and $l_{\text{leaf,coh}}$ is the leaf loss from phenology in that timestep (KgC m^{-2}). $M_{\text{t,coh}}$ is the total mortality flux in that timestep (in individuals). For fine root input. $n_{\text{coh,ft}}$ is the number of cohorts of functional type 'FT' in the current patch.

$$l_{\text{root,in,ft}} = \left(\sum_{i=1}^{n_{\text{coh,ft}}} n_{\text{coh}} (r_{\text{md,coh}}) + M_{\text{t,coh}} \cdot b_{\text{root,coh}} \right) / \sum_{p=1}^{n_{\text{pat}}} A_p \quad (173)$$

where $r_{\text{md,coh}}$ is the root turnover rate. For coarse woody debris input ($CWD_{\text{AG,in,lsc}}$), we first calculate the sum of the mortality $M_{\text{t,coh}} \cdot (b_{\text{struc,coh}} + b_{\text{sw,coh}})$ and turnover $n_{\text{coh}} (w_{\text{md,coh}})$ fluxes, then separate these into size classes and above/below ground fractions using the fixed fractions assigned to each (f_{lsc} and f_{ag})

$$CWD_{\text{AG,in,lsc}} = \left(f_{\text{lsc}} \cdot f_{\text{ag}} \sum_{i=1}^{n_{\text{coh,ft}}} n_{\text{coh}} w_{\text{md,coh}} + M_{\text{t,coh}} \cdot (b_{\text{struc,coh}} + b_{\text{sw,coh}}) \right) / \sum_{p=1}^{n_{\text{pat}}} A_p \quad (174)$$

$$CWD_{\text{BG,in,lsc}} = \left(f_{\text{lsc}} \cdot (1 - f_{\text{ag}}) \sum_{i=1}^{n_{\text{coh,ft}}} n_{\text{coh}} w_{\text{md,coh}} + M_{\text{t,coh}} \cdot (b_{\text{struc,coh}} + b_{\text{sw,coh}}) \right) / \sum_{p=1}^{n_{\text{pat}}} A_p \quad (175)$$

14.2 Litter Outputs

²² The fragmenting litter pool is available for burning but not for respiration or decomposition. Fragmentation rates are calculated according to a maximum fragmentation rate ($\alpha_{\text{cwd,lsc}}$ or α_{litter}) which is ameliorated by a temperature and water dependent scalar S_{tw} . The form of the temperature scalar is taken from the existing CLM4.5BGC decomposition cascade calculations). The water scaler is equal to the water limitation on photosynthesis (since the CLM4.5BGC water scaler pertains to the water potential of individual soil layers, which it is difficult to meaningfully average, given the non-linearities in the impact of soil moisture). The scaler code is modular, and new functions may be implemented trivially. Rate constants for the decay of the litter pools are extremely uncertain in literature, as few studies either separate litter into size classes, nor examine its decomposition under non-limiting moisture and temperature conditions. Thus, these parameters should be considered as part of sensitivity analyses of the model outputs.

$$CWD_{\text{AG,out,lsc}} = CWD_{\text{AG,lsc}} \cdot \alpha_{\text{cwd,lsc}} \cdot S_{\text{tw}} \quad (176)$$

$$CWD_{\text{BG,out,lsc}} = CWD_{\text{BG,lsc}} \cdot \alpha_{\text{cwd,lsc}} \cdot S_{\text{tw}} \quad (177)$$

$$l_{\text{leaf,out,ft}} = l_{\text{leaf,ft}} \cdot \alpha_{\text{litter}} \cdot S_{\text{tw}} \quad (178)$$

$$l_{\text{root,out,ft}} = l_{\text{root,ft}} \cdot \alpha_{\text{root,ft}} \cdot S_{\text{tw}} \quad (179)$$

²¹This description relates to algorithms in the `CWD_input`, `mortality_litter_fluxes` and `fire_litter_fluxes` subroutines

²²This description relates to algorithms in the `CWD_out` subroutine

14.3 Flux into decomposition cascade

²³ Upon fragmentation and release from the litter pool, carbon is transferred into the labile, lignin and cellulose decomposition pools. These pools are vertically resolved in the biogeochemistry model. The movement of carbon into each vertical layer is obviously different for above- and below-ground fragmenting pools. For each layer z and chemical litter type i , we derive a flux from ED into the decomposition cascade as $ED_{lit,i,z}$ ($\text{kGC m}^{-2} \text{s}^{-1}$)

$$ED_{lit,lab,z} = t_c \sum_{i=1}^{n_{patch}} \left(l_{prof,z} \sum_{i=1}^{n_{FT}} l_{leaf,out,FT} f_{lab,l,FT} + \sum_{i=1}^{n_{FT}} r_{f,prof,z,FT} l_{root,out,FT} f_{lab,root,FT} \right) \quad (180)$$

$$ED_{lit,cel,z} = t_c \sum_{i=1}^{n_{patch}} \left(l_{prof,z} \left(\sum_{i=1}^{n_{lsc}} CWD_{AG,out,lsc} f_{cel,CWD} + \sum_{i=1}^{n_{FT}} l_{leaf,out,ft} f_{cel,l,FT} \right) + r_{c,prof,z} \sum_{i=1}^{n_{lsc}} CWD_{BG,out,lsc} f_{cel,CWD} + \sum_{i=1}^{n_{FT}} r_{f,prof,z,ft} l_{root,out,ft} f_{cel,fr,ft} \right) \quad (181)$$

$$ED_{lit,lig,z} = t_c \sum_{i=1}^{n_{patch}} \left(l_{prof,z} \left(\sum_{i=1}^{n_{lsc}} CWD_{AG,out,lsc} f_{lig,CWD} + \sum_{i=1}^{n_{FT}} l_{leaf,out,ft} f_{lig,l,FT} \right) + r_{c,prof,z} \sum_{i=1}^{n_{lsc}} CWD_{BG,out,lsc} f_{lig,CWD} + \sum_{i=1}^{n_{FT}} r_{f,prof,z,ft} l_{root,out,ft} f_{lig,fr,ft} \right) \quad (182)$$

where t_c is the time conversion factor from years to seconds, $f_{lab,l}$, $f_{cel,l}$ and $f_{lig,l}$ are the fractions of labile, cellulose and lignin in leaf litter, and $f_{lab,r}$, $f_{cel,r}$ and $f_{lig,r}$ are their counterparts for root matter. Similarly, l_{prof} , $r_{f,prof}$ and $r_{c,prof}$ are the fractions of leaf, coarse root and fine root matter that are passed into each vertical soil layer z , derived from the CLM(BGC) model.

Table 15: Parameters needed for litter model.

Parameter Symbol	Parameter Name	Units	indexed by
$\alpha_{cwd,lsc}$	Maximum fragmentation rate of CWD	y^{-1}	
α_{litter}	Maximum fragmentation rate of leaf litter	y^{-1}	
α_{root}	Maximum fragmentation rate of fine root litter	y^{-1}	
$f_{lab,l}$	Fraction of leaf mass in labile carbon pool	none	
$f_{cel,l}$	Fraction of leaf mass in cellulose carbon pool	none	
$f_{lig,l}$	Fraction of leaf mass in lignin carbon pool	none	
$f_{lab,r}$	Fraction of root mass in labile carbon pool	none	
$f_{cel,r}$	Fraction of root mass in cellulose carbon pool	none	
$f_{lig,r}$	Fraction of root mass in lignin carbon pool	none	
$l_{prof,z}$	Fraction of leaf matter directed to soil layer z	none	soil layer
$r_{c,prof,z}$	Fraction of coarse root matter directed to soil layer z	none	soil layer
$r_{f,prof,z}$	Fraction of fine root matter directed to soil layer z	none	soil layer

15 Plant Mortality

Total plant mortality per cohort $M_{t,coh}$, (fraction year^{-1}) is simulated as the sum of four additive terms,

$$M_{t,coh} = M_{b,coh} + M_{cs,coh} + M_{hf,coh} + M_{f,coh}, \quad (183)$$

²³This description relates to algorithms in the flux.into.litter.pools subroutine

where M_b is the background mortality that is unaccounted by any of the other mortality rates and is fixed at 0.014. M_{cs} is the carbon starvation derived mortality, which is a function of the non-structural carbon storage term $b_{store,coh}$ and the PFT-specific ‘target’ carbon storage, $l_{targ,ft}$, as follows:

$$M_{cs,coh} = \max \left(0.0, S_{m,ft} \left(0.5 - \frac{b_{store,coh}}{l_{targ,ft} b_{leaf}} \right) \right) \quad (184)$$

where $S_{m,ft}$ is the ‘stress mortality’ parameter, or the fraction of trees in a landscape that die when the mean condition of a given cohort triggers mortality. This parameter is needed to scale from individual-level mortality simulation to grid-cell average conditions.

Mechanistic simulation of hydraulic failure is not undertaken on account of it’s mechanistic complexity (see [16] for details). Instead, we use a proxy for hydraulic failure induced mortality ($M_{hf,coh}$) that uses a water potential threshold beyond mortality is triggered, such that the tolerance of low water potentials is a function of plant functional type (as expressed via the ψ_c parameter). For each day that the aggregate water potential falls below a threshold value, a set fraction of the trees are killed. The aggregation of soil moisture potential across the root zone is expressed using the β function. We thus determine plant mortality caused by extremely low water potentials as

$$M_{hf,coh} = \begin{cases} S_{m,ft} & \text{for } \beta_{ft} < 10^{-6} \\ 0.0 & \text{for } \beta_{ft} \geq 10^{-6}. \end{cases} \quad (185)$$

The threshold value of 10^{-6} represents a state where the average soil moisture potential is within 10^{-6} of the wilting point (a PFT specific parameter $\theta_{w,ft}$).

$M_{hf,coh}$ is the fire-induced mortality, as described in the fire modelling section.

Table 16: Parameters needed for mortality model.

Parameter Symbol	Parameter Name	Units	indexed by
$S_{m,ft}$	Stress Mortality Scaler	none	
$l_{targ,ft}$	Target carbon storage fraction	none	ft

16 Fire

The influence of fire on vegetation is estimated using the SPITFIRE model, which has been modified for use in ED following it’s original implementation in the LPJ-SPITFIRE model ([32, 23]). This model as described is substantially different from the existing CLM4.5 fire model [12], however, further developments are intended to increase the merging of SPITFIRE’s natural vegetation fire scheme with the fire suppression, forest-clearing and peat fire estimations in the existing model. The coupling to the ED model allows fires to interact with vegetation in a ‘size-structured’ manner, so small fires can burn only understorey vegetation. Also, the patch structure and representation of succession in the ED model allows the model to track the impacts of fire on different forest stands, therefore removing the problem of area-averaging implicit in area-based DGVMs. The SPITFIRE approach has also been coupled to the LPJ-GUESS individual-based model (Forrest et al. in prep) and so this is not the only implementation of this type of scheme in existence.

The SPITFIRE model operates at a daily timestep and at the patch level, meaning that different litter pools and vegetation characteristics of open and closed forests can be represented effectively (we omit the ‘patch’ subscript throughout for simplicity).

16.1 Properties of fuel load

Many fire processes are impacted by the properties of the litter pool in the SPITFIRE model. There are one live (live grasses) and five dead fuel categories (dead leaf litter and four pools of coarse woody debris). Coarse woody debris is classified into 1h, 10h, 100h, and 1000h fuels, defined by “*the order of magnitude of time required for fuel to lose (or gain) 63% of the difference between its current moisture content and the equilibrium moisture content under defined atmospheric conditions.*” [32]. For the purposes of describing the behaviour of fire, we introduce a new index ‘fuel class’ fc , the values of which correspond to each of the six possible fuel categories as follows.

Table 17: Fuel categories

<i>fc</i> index	Fuel type	Drying Time
1	dead grass	n/a
2	twigs	1h fuels
3	small branches	10h fuel
4	large branches	100h fuel
5	stems and trunks	1000h fuel
6	live grasses	n/a

16.1.1 Fuel load

Total fuel load $F_{\text{tot,patch}}$ for a given patch is the sum of the above ground coarse woody debris and the leaf litter, plus the alive grass leaf biomass $b_{\text{l,grass}}$ multiplied by the non-mineral fraction $(1-M_f)$.

$$F_{\text{tot,patch}} = \left(\sum_{fc=1}^{fc=5} CWD_{AG,fc} + l_{\text{litter}} + b_{\text{l,grass}} \right) (1 - M_f) \quad (186)$$

Many of the model behaviours are affected by the patch-level weighted average properties of the fuel load. Typically, these are calculated in the absence of 1000-h fuels because these do not contribute greatly to fire spread properties.

16.1.2 Nesterov Index

Dead fuel moisture ($moist_{\text{df,fc}}$), and several other properties of fire behaviour, are a function of the ‘Nesterov Index’ (N_I) which is an accumulation over time of a function of temperature and humidity (Eqn 5, [32]).

$$N_I = \sum \max(T_d(T_d - D), 0) \quad (187)$$

where T_d is the daily mean temperature in °C and D is the dew point calculated as .

$$v = \frac{17.27T_d}{237.70 + T_d} + \log(RH/100) \quad (188)$$

$$D = \frac{237.70v}{17.27 - v} \quad (189)$$

where RH is the relative humidity (%).

16.1.3 Dead Fuel Moisture Content

Dead fuel moisture is calculated as

$$moist_{\text{df,fc}} = e^{-\alpha_{\text{fmc,fc}} N_I} \quad (190)$$

where $\alpha_{\text{fmc,fc}}$ is a parameter defining how fuel moisture content varies between the first four dead fuel classes.

16.1.4 Live grass moisture Content

The live grass fractional moisture content ($moist_{\text{lg}}$) is a function of the soil moisture content. (Equation B2 in [32])

$$moist_{\text{lg}} = \max(0.0, \frac{10}{9}\theta_{30} - \frac{1}{9}) \quad (191)$$

where θ_{30} is the fractional moisture content of the top 30cm of soil.

16.1.5 Patch Fuel Moisture

The total patch fuel moisture is based on the weighted average of the different moisture contents associated with each of the different live grass and dead fuel types available (except 1000-h fuels).

$$F_{\text{m,patch}} = \sum_{fc=1}^{fc=4} \frac{F_{\text{fc}}}{F_{\text{tot}}} moist_{\text{df,fc}} + \frac{b_{\text{l,grass}}}{F_{\text{tot}}} moist_{\text{lg}} \quad (192)$$

16.1.6 Effective Fuel Moisture Content

Effective Fuel Moisture Content is used for calculations of fuel consumed, and is a function of the ratio of dead fuel moisture content $M_{df,fc}$ and the moisture of extinction factor, $m_{ef,fc}$

$$E_{moist,fc} = \frac{moist_{fc}}{m_{ef,fc}} \quad (193)$$

where the m_{ef} is a function of surface-area to volume ratio.

$$m_{ef,fc} = 0.524 - 0.066 \log_{10} \sigma_{fc} \quad (194)$$

16.1.7 Patch Fuel Moisture of Extinction

The patch ‘moisture of extinction’ factor (F_{mef}) is the weighted average of the m_{ef} of the different fuel classes

$$F_{mef,patch} = \sum_{fc=1}^{fc=5} \frac{F_{fc}}{F_{tot}} m_{ef,fc} + \frac{b_{l,grass}}{F_{tot}} m_{ef,grass} \quad (195)$$

16.1.8 Patch Fuel Bulk Density

The patch fuel bulk density is the weighted average of the bulk density of the different fuel classes (except 1000-h fuels).

$$F_{bd,patch} = \sum_{fc=1}^{fc=4} \frac{F_{fc}}{F_{tot}} \beta_{fuel,fc} + \frac{b_{l,grass}}{F_{tot}} \beta_{fuel,lgrass} \quad (196)$$

where $\beta_{fuel,fc}$ is the bulk density of each fuel size class (kG m^{-3})

16.1.9 Patch Fuel Surface Area to Volume

The patch surface area to volume ratio (F_{σ}) is the weighted average of the surface area to volume ratios (σ_{fuel}) of the different fuel classes (except 1000-h fuels).

$$F_{\sigma} = \sum_{fc=1}^{fc=4} \frac{F_{fc}}{F_{tot}} \sigma_{fuel,fc} + \frac{b_{l,grass}}{F_{tot}} \sigma_{fuel,grass} \quad (197)$$

16.2 Forward rate of spread

For each patch and each day, we calculate the rate of forward spread of the fire ros_f (nominally in the direction of the wind).

$$ros_f = \frac{i_r x_i (1 - \phi_w)}{F_{bd,patch} e_{ps} q_{ig}} \quad (198)$$

e_{ps} is the effective heating number ($e^{\frac{-4.528}{F_{\sigma,patch}}}$). q_{ig} is the heat of pre-ignition ($581 + 2594 F_m$). x_i is the propagating flux calculated as (see [32] Appendix A).

$$x_i = \begin{cases} 0.0 & \text{for } F_{\sigma,patch} < 0.00001 \\ \frac{e^{0.792 + 3.7597 F_{\sigma,patch}^{0.5} (\frac{F_{bd,patch}}{p_d} + 0.1)}}{192 + 7.9095 F_{\sigma,patch}} & \text{for } F_{\sigma,patch} \geq 0.00001 \end{cases} \quad (199)$$

ϕ_w is the influence of windspeed on rate of spread.

$$\phi_w = cb_w^b \beta^{-e} \quad (200)$$

Where b , c and e are all functions of surface-area-volume ratio $F_{\sigma,patch}$: $b = 0.15988 F_{\sigma,patch}^{0.54}$, $c = 7.47 e^{-0.8711 F_{\sigma,patch}^{0.55}}$, $e = 0.715 e^{-0.01094 F_{\sigma,patch}}$. $b_w = 196.86 W$ where W is the the windspeed in ms^{-1} , and $\beta = \frac{F_{bd}/p_d}{0.200395 F_{\sigma,patch}^{0.8189}}$ where p_d is the particle density (513).

i_r is the reaction intensity, calculated using the following set of expressions (from [32] Appendix A):

$$i_r = \Gamma_{\text{opt}} F_{\text{tot}} H d_{\text{moist}} d_{\text{miner}} \quad (201)$$

$$d_{\text{moist}} = \max\left(0.0, (1 - 2.59m_w + 5.11m_w^2 - 3.52m_w^3)\right) \quad (202)$$

$$m_w = \frac{F_{\text{m,patch}}}{F_{\text{mef,patch}}} \quad (203)$$

$$\Gamma_{\text{opt}} = \Gamma_{\text{max}} \beta^a \lambda \quad (204)$$

$$\Gamma_{\text{max}} = \frac{1}{0.0591 + 2.926 F_{\sigma, \text{patch}}^{-1.5}} \quad (205)$$

$$\lambda = e^{a(1-\beta)} \quad (206)$$

$$a = 8.9033 F_{\sigma, \text{patch}}^{-0.7913} \quad (207)$$

Γ_{opt} is the residence time of the fire, and d_{miner} is the mineral damping coefficient ($=0.174S_e^{-0.19}$, where S_e is 0.01 and so $=d_{\text{miner}} 0.41739$).

16.3 Fuel Consumption

The fuel consumption (fraction of biomass pools) of each dead biomass pool in the area affected by fire on a given day ($f_{\text{c,dead,fc}}$) is a function of effective fuel moisture $E_{\text{moist,fc}}$ and size class fc (Eqn B1, B4 and B5, [32]). The fraction of each fuel class that is consumed decreases as its moisture content relative to its moisture of extinction ($E_{\text{moist,fc}}$) increases.

$$f_{\text{cdead,fc}} = \max\left(0, \min(1, m_{\text{int,mc,fc}} - m_{\text{slope,mc,fc}} E_{\text{moist,fc}})\right) \quad (208)$$

m_{int} and m_{slope} are parameters, the value of which is modulated by both size class fc and by the effective fuel moisture class mc , defined by $E_{\text{moist,fc}}$. m_{int} and m_{slope} are defined for low-, mid-, and high-moisture conditions, the boundaries of which are also functions of the litter size class following [22] (page 802). The fuel burned, $f_{\text{cground,fc}}$ ($\text{Kg m}^{-2} \text{ day}^{-1}$) is calculated from $f_{\text{cdead,fc}}$ for each fuel class:

$$f_{\text{cground,fc}} = f_{\text{c,dead,fc}} (1 - M_f) \frac{F_{\text{fc}}}{0.45} \quad (209)$$

Where 0.45 converts from carbon to biomass. The total fuel consumption, $f_{\text{ctot,patch}}$ (Kg m^{-2}), used to calculate fire intensity, is then given by

$$f_{\text{ctot,patch}} = \sum_{fc=1}^{fc=4} f_{\text{c,ground,fc}} + f_{\text{c,ground,lgrass}} \quad (210)$$

There is no contribution from the 1000 hour fuels to the patch-level $f_{\text{ctot,patch}}$ used in the fire intensity calculation.

16.4 Fire Intensity

Fire intensity at the front of the burning area (I_{surface} , kW m^{-2}) is a function of the total fuel consumed ($f_{\text{ctot,patch}}$) and the rate of spread at the front of the fire, ros_f (m min^{-1}) (Eqn 15 [32])

$$I_{\text{surface}} = \frac{0.001}{60} f_{\text{energy}} f_{\text{ctot,patch}} ros_f \quad (211)$$

where f_{energy} is the energy content of fuel (kJ/kg - the same for all fuel classes). Fire intensity is used to define whether an ignition is successful. If the fire intensity is greater than 50kw/m then the ignition is successful.

16.5 Fire Duration

Fire duration is a function of the fire danger index with a maximum length of $F_{\text{dur,max}}$ (240 minutes in [32] Eqn 14, derived from Canadian Forest Fire Behaviour Predictions Systems)

$$D_f = \min\left(F_{\text{dur,max}}, \frac{F_{\text{dur,max}}}{1 + F_{\text{dur,max}} e^{-11.06fdi}}\right) \quad (212)$$

16.6 Fire Danger Index

Fire danger index (f_{di}) is a representation of the effect of meteorological conditions on the likelihood of a fire. It is calculated for each gridcell as a function of the Nesterov Index . f_{di} is calculated from NI as

$$f_{di} = 1 - e^{\alpha NI} \quad (213)$$

where $\alpha = 0.00037$ following [34].

16.7 Area Burned

Total area burnt is assumed to be in the shape of an ellipse, whose major axis f_{length} (m) is determined by the forward and backward rates of spread (ros_f and ros_b respectively).

$$f_{\text{length}} = F_d(ros_b + ros_f) \quad (214)$$

ros_b is a function of ros_f and windspeed (Eqn 10 [32])

$$ros_b = ros_f e^{-0.72W} \quad (215)$$

The minor axis to major axis ratio l_b of the ellipse is determined by the windspeed. If the windspeed (W) is less than 16.67 ms^{-1} then $l_b = 1$. Otherwise (Eqn 12 and 13, [32])

$$l_b = \min\left(8, f_{\text{tree}}(1.0 + 8.729(1.0 - e^{-0.108W})^{2.155}) + (f_{\text{grass}}(1.1 + 3.6W^{0.0464}))\right) \quad (216)$$

f_{grass} and f_{tree} are the fractions of the patch surface covered by grass and trees respectively. The total area burned (A_{burn} in m^2) is therefore (Eqn 11, [32])

$$A_{\text{burn}} = \frac{n_f \frac{3.1416}{4l_b} (f_{\text{length}}^2)}{10000} \quad (217)$$

where n_f is the number of fires.

16.8 Crown Damage

c_k is the fraction of the crown which is consumed by the fire. This is calculated from scorch height H_s , tree height h and the crown fraction parameter F_{crown} (Eqn 17 [32]):

$$c_k = \begin{cases} 0 & \text{for } H_s < (h - hF_{\text{crown}}) \\ 1 - \frac{h - H_s}{h - F_{\text{crown}}} & \text{for } h > H_s > (h - hF_{\text{crown}}) \\ 1 & \text{for } H_s > h \end{cases}$$

The scorch height H_s (m) is a function of the fire intensity, following [3], and is proportional to a plant functional type specific parameter $\alpha_{s,ft}$ (Eqn 16 [32]):

$$H_s = \sum_{\text{FT}=1}^{\text{NPFT}} \alpha_{s,p} \cdot f_{\text{biomass,ft}} I_{\text{surface}}^{0.667} \quad (218)$$

where $f_{\text{biomass,ft}}$ is the fraction of the above-ground biomass in each plant functional type.

16.9 Cambial Damage and Kill

The cambial kill is a function of the fuel consumed $f_{c,\text{tot}}$, the bark thickness t_b , and τ_1 , the duration of cambial heating (minutes) (Eqn 8, [22]):

$$\tau_1 = \sum_{\text{fc}=1}^{\text{fc}=5} 39.4 F_{p,c} \frac{10000}{0.45} (1 - (1 - f_{c,\text{dead,fc}})^{0.5}) \quad (219)$$

Bark thickness is a linear function of tree diameter dbh_{coh} , defined by PFT-specific parameters $\beta_{1,\text{bt}}$ and $\beta_{2,\text{bt}}$ (Eqn 21 [32]):

$$t_{b,\text{coh}} = \beta_{1,\text{bt,ft}} + \beta_{2,\text{bt,ft}} dbh_{\text{coh}} \quad (220)$$

The critical time for cambial kill, τ_c (minutes) is given as (Eqn 20 [32]):

$$\tau_c = 2.9t_b^2 \quad (221)$$

The mortality rate caused by cambial heating τ_{pm} of trees within the area affected by fire is a function of the ratio between τ_1 and τ_c (Eqn 19, [32]):

$$\tau_{pm} = \begin{cases} 1.0 & \text{for } \tau_1/\tau_c \geq 2.0 \\ 0.563(\tau_1/\tau_c) - 0.125 & \text{for } 2.0 > \tau_1/\tau_c \geq 0.22 \\ 0.0 & \text{for } \tau_1/\tau_c < 0.22 \end{cases} \quad (222)$$

Table 18: Parameters needed for fire model.

Parameter Symbol	Parameter Name	Units	indexed by
$\beta_{1,bt}$	Intercept of bark thickness function	mm	<i>FT</i>
$\beta_{2,bt}$	Slope of bark thickness function	mm cm ⁻¹	<i>FT</i>
F_{crown}	Ratio of crown height to total height	none	<i>FT</i>
α_{fmc}	Fuel moisture parameter	°C ⁻²	<i>fc</i>
β_{fuel}	Fuel Bulk Density	kG m ⁻³	<i>fc</i>
$\sigma_{fuel,fc}$	Surface area to volume ratio	cm ⁻¹	<i>fc</i>
m_{int}	Intercept of fuel burned	none	<i>fc</i> , moisture class
m_{slope}	Slope of fuel burned	none	<i>fc</i> , moisture class
M_f	Fuel Mineral Fraction		
$F_{dur,max}$	Maximum Duration of Fire	Minutes	
f_{energy}	Energy content of fuel	kJ/kg	
α_s	Flame height parameter		<i>FT</i>

16.10 Additional features of tech note to be included

- Table of default parameter values and their potential ranges
- Comprehensive list of all patch and cohort attributes.