Environmental drivers of drought deciduous phenology in the Community Land Model

K.M. Dahlin[1,2], R.A. Fisher[2], & P.J. Lawrence[2]

[1] Department of Geography, Michigan State University, East Lansing, Michigan, USA

[2] Climate and Global Dynamics Division, National Center for Atmospheric Research*, Boulder, Colorado, USA

* NCAR is sponsored by the National Science Foundation.

Correspondence to: K.M. Dahlin (kdahlin@msu.edu)

Abstract

Seasonal changes in plant leaf area have a substantial impact on global climate. If and when leaves are present affects surface roughness and albedo, and the gas exchange occurring between leaves and the atmosphere affects carbon dioxide concentrations and the global water system. Thus, correct predictions of plant phenological processes are important for understanding the present and future states of the Earth system. Here we compare plant phenology as estimated in the Community Land Model (CLM) to that derived from satellites in drought deciduous regions of the world. We reveal a subtle but important issue in the CLM: anomalous green-up during the dry season in many semi-arid parts of the world owing to rapid upwards water movement from wet to dry soil layers. We develop and implement a solution for this problem by introducing an additional criterion of minimum cumulative rainfall to the leaf-out trigger in the drought deciduous algorithm. We discuss some of the broader ecological impacts of this change and highlight some of the further steps that need to be taken to fully incorporate this change into the CLM framework.

1 Introduction

Ecosystems change with the seasons in response to environmental cues. Some of those cues are fixed, like day length, while others are climate-driven and therefore vary from year to year. The combination of fixed and climate-driven phenological cues poses an interesting problem in the face of climate change – climate related drivers of phenology (temperature and rainfall patterns) are likely to change (Lau et al., 2013), while fixed cues will remain unchanged. Phenological shifts due to climate change have already been identified (e.g. Parmesan & Yohe, 2003). Phenology can refer to a large number of patterns and behaviors in plants and animals that shift with the seasons. Here, however, because we are focused on land surface model simulations, we use phenology specifically to refer to intraannual variation in leaf area index (LAI). Leaf area can vary significantly within a year and is, therefore, a critical control on land-atmosphere feedbacks (Lawrence et al., 2012).

Recent advances have greatly improved our ability to predict seasonal patterns in northern temperate deciduous forests (Richardson et al., 2012), but our understanding of phenological patterns in stress or drought deciduous plants (also called 'raingreen') remains weak (Guan et al., 2014; Jenerette et al., 2010; Ma et al., 2013). The semi-arid ecosystems that host the majority of drought deciduous woody plants have relatively low biomass but make up a large fraction of global land area (~30%; Scholes & Hall, 1996). Their extensiveness alone makes them important to global radiation budgets, but additionally these systems are likely very sensitive to climate change given their apparent bistability (Scholes & Hall, 1996; Staver *et al.*, 2011). In semi-arid ecosystems leaf-out is typically thought to be a function of water availability (Reich, 1995; White et al., 1997), however, some woody plants do leaf-out several weeks before the first rains of the season (Archibald and Scholes, 2007).

In an Earth system modeling context, the timing and magnitude of plant phenology, and how these processes may change, is critical for approximating the energy and carbon balances of the planet. Prognostic phenology has only recently been incorporated in to Earth system models, however, and its fidelity, particularly in semi-arid regions, remains poorly tested (Blyth et al., 2011; Lawrence et al., 2011; Randerson et al., 2009). Lawrence *et al.* (2012) found that the prognostic phenology in the Community Land Model version 4 (CLM4(CN)) degraded estimates of latent heat flux and other biophysical properties in comparison to using prescribed, satellite-derived phenology (CLM4SP). Wang *et al.* (2013) compared intraannual variation in the fraction of absorbed photosynthetically active radiation (fAPAR) in CLM4CN to satellite-derived estimates and found substantial differences in regional averages, zonal means, and interannual trends. It is difficult, however, to isolate the impact of the drought deciduous phenology algorithm using these regional and zonal estimates.

Satellite-derived estimates of greenness, fAPAR, and LAI have greatly improved our ability to study the environmental drivers of phenology (Reed et al., 2009), however, the majority of studies have focused on northern deciduous and boreal forests (e.g. Delbart *et al.*, 2006; White *et al.*, 2009; Yang *et al.*, 2012). While fewer studies have focused on remote sensing of phenology in semi-arid systems, Zhang *et al.* (2005) found a strong relationship between greenness onset and the start of the rainy season across the semi-arid parts of Africa. They found a weaker relationship, however, between dormancy and the end of rainy seasons, and they attribute this weakness to differences in soil properties. Similarly, Ma *et al.* (2013) found a strong relationship between greenness and rainfall in northern Australia in both seasonal timing and amplitude and Bradley *et al.* (2011) found a close relationship between rainfall and seasonality in Amazonian savannas. Interestingly, in Africa Zhang *et al.* (2005) also showed a strong relationship between latitude and both green-up and dormancy onset, even in the narrow band of the Sahelian and Sub-Sahelian region, suggesting a possible link

between phenology and subtle changes in photoperiod at least in northern Africa. Recently Guan *et al.* (2014) showed a relationship between woody plant cover and phenological timing in African savannas.

In this study we address three questions related to the representation of drought deciduous phenology in the CLM. (1) How well does the CLM capture phenological patterns of LAI among different drought deciduous plant functional types (PFTs) as compared to satellite-derived estimates?; (2) Which parameters in the current version of the CLM have the most leverage on drought deciduous phenology?; and (3) Do changes in the phenology algorithms in the CLM improve the model's representation of seasonal cycles regionally?

2 Methods

2.1 Model Description

The CLM is the terrestrial component of the Community Earth System Model (CESM; Lawrence *et al.*, 2011); it simulates biogeophysical and biogeochemical processes including radiation interactions with vegetation and soil, heat transfer in soil and snow, hydrology, and plant photosynthesis and respiration. In this paper we use the most recent release of the Community Land Model with active biogeochemistry, CLM4.5BGC (Oleson et al., 2013). Henceforth, references in this paper to the "CLM" will refer to CLM4.5BGC.

The CLM is run here on a 1.25^{o} x 0.9375^{o} grid, and each grid cell is, where applicable, divided into fractions representing vegetated land, lakes, glaciers, and urban areas. Within the vegetated fraction of a grid cell there may be multiple PFTs representing a coarse division of biodiversity along its major axes of variation: trees/shrubs/grass, broadleaf/needleaf, C3/C4 photosynthesis mechanisms and phenological habit (evergreen, cold deciduous and stress/drought

93 deciduous). There are currently 15 non-crop PFTs in the CLM, four of which follow the drought
94 deciduous phenology algorithm (Oleson et al., 2013). **Figure 1** shows where these different PFTs
95 dominate the globe. Over time the relative cover of the PFTs may shift, as may the overall fraction
96 of vegetation, depending on shifts in land use, though these shifts are minor in recent decades. In
97 the simulations used in this paper these shifts in PFT fractions and cover are prescribed from
98 satellite observations (Lawrence and Chase, 2007) as opposed to emerging from vegetation
99 competition (Bonan et al., 2003).

100 In the CLM, drought deciduous plants are represented by the 'stress deciduous' phenology
101 type, as distinct from the evergreen or 'seasonal' (cold) deciduous phenology types. This designation
102 allows for plants to lose their leaves either via the impact of cold, via the impact of drought, or via
103 the onset of short days thus allowing the model to simulate, for example, grass vegetation growing in
104 an environment that is both seasonally cold and seasonally dry. If the triggers for offset are not
105 reached in a given year, drought deciduous vegetation will follow the evergreen phenology
106 algorithm, gaining and losing fixed fractions of carbon with each time step. This stress deciduous
107 algorithm, described in more detail below and in Oleson *et al.* (2013), was developed in part from
108 White *et al.* (1997), though that study was particularly focused on grass phenology.

109 The deciduous algorithms are hierarchical, such that plants classified as 'stress deciduous' but
110 growing at high latitudes or in cold climates will follow the same onset/offset rules as
111 cold/seasonally deciduous plants. From the beginning of a dormant period a 'freezing day
112 accumulator' is activated whereby time steps with a temperature below freezing (0° C) are summed
113 and if this sum exceeds 15 days then the plants will follow both the winter deciduous and drought
114 deciduous algorithms. Leaf onset can only be triggered if day length is greater than 6 hours, a
115 latitude-specific sum of growing degree days has been reached (described in Oleson *et al.* (2013)) and
116 the soil wetness criteria described below have been met.

In seasonally dry, warm regions (the focus of this paper) where day length is never less than 6 hours, leaf onset for the stress deciduous phenology type is determined by soil wetness. At the end of the previous offset period an accumulated soil water index (SWI) is set to zero and accumulation is calculated as:

$$SWI^n = \begin{cases} SWI^{n-1} + f_{day} & \text{for } \psi_{soil\,3} \geq \psi_{threshold} \\ SWI^{n-1} & \text{for } \psi_{soil\,3} < \psi_{threshold} \end{cases} \quad (1)$$

Where *n* and *n-1* refer to the values in the previous and current time steps, $\Psi_{\text{soil 3}}$ is the soil water potential (MPa) in the third soil layer (6.23 cm – 9.06 cm), $\Psi_{\text{threshold}}$ is -2 MPa, and f_{day} is a time step (30 minutes in CLM) as a fraction of a day. Onset is triggered when SWI exceeds 15 days.

The rate of leaf onset (fraction of onset per time step), which in the CLM is represented as the transfer of C and N from a storage pool to the 'display' leaf pool, is determined by the number of days prescribed for onset, fixed at 30 days. The rate (r_{onset}) at each time step is defined as:

$$r_{onset} = \begin{cases} \frac{2}{t_{onset}} & \text{for } t_{onset} \neq \Delta t \\ \frac{1}{\Delta t} & \text{for } t_{onset} = \Delta t \end{cases} \quad (2)$$

where t_{onset} is time remaining in the current onset period in seconds and Δt is the length of a time step (1800 seconds). The flux of C out of the storage pool is then defined as the amount in the C storage pool at that time step multiplied by r_{onset}. These functions result in a linearly decreasing flux out of the transfer pool, so the rate of increase in LAI over the onset period steadily decreases as C moves from the storage pool to the display pool (see Fig 14.1 in Oleson *et al.* (2013)). During the onset period C and N are also transferred from storage pools for fine roots, live and dead stem, and live and dead coarse roots into these components' respective displayed growth pools. During the growing season, C and N taken up by the plant are accumulated in transfer pools, to be used in the next growing season.

As long as the leaf onset period is complete, leaf offset can be triggered by short (<6 hr) day length, a period of cold temperatures (described in Oleson *et al.* (2013)) or if the soil dryness criteria described below has been met.

The offset soil wetness index (OSWI) can potentially start accumulating time steps once the previous leaf onset phase is complete. The algorithm differs slightly from the onset trigger, in that OSWI can increase or decrease as described below.

$$OSWI^n = \begin{cases} OSWI^{n-1} + f_{day} & \text{for } \psi_{soil\,3} \leq \psi_{threshold} \\ \max(OSWI^{n-1} - f_{day}, 0) & \text{for } \psi_{soil\,3} > \psi_{threshold} \end{cases} \quad (3)$$

where $\Psi_{\text{threshold}}$ is -2 MPa, and leaf offset is triggered when OSWI equals 15 days.

Similar to the rate of leaf onset, leaf offset rate is a function of the amount of time left in the offset period, fixed at 15 days:

$$r_{offset} = \frac{2\Delta t}{t_{offset}^2} \quad (4)$$

Carbon fluxes into the litter pool are only calculated for leaves and fine roots (stems and coarse roots cannot shrink). Nitrogen fluxes into the litter pool reflect retranslocation of N prior to offset. See Oleson *et al.* (2013)) for more details.

The model runs used in the global simulations described here ran for 45 years, and were started from an equilibrium baseline state generated by a standard CLM spin up run (as described in detail by Koven et al., 2013), cycling meteorological conditions of 1948-1972. The present-day run (1965 – 2010) used CRU-NCEP meteorological reanalysis data (N. Vivoy, pers. comm.; data available at: http://dods.ipsl.jussieu.fr/igcmg/IGCM/BC/OOL/OL/CRU-NCEP/) and transient CO_2 concentrations to drive the model. Soil type and land cover are prescribed in the model, and recent work has suggested that the soil resistance parameterization may be unrealistic in arid

159 ecosystems (Swenson and Lawrence, 2014). More details on CLM are available in Oleson et al.
160 (2013).

161 **2.2 Satellite derived LAI**

162 We compared the model-derived estimates of LAI to those estimated from the Advanced
163 Very High Resolution Radiometer sensors (AVHRR) onboard the National Oceanic and
164 Atmospheric Administration satellites. These data are available bimonthly and span from 1981 to
165 2011. They are supplied at 1/12 degree resolution. A detailed description of the development of the
166 LAI product (hereafter LAI3g) is in Zhu *et al.* (2013).

167 To ensure the most appropriate comparison possible, the LAI3g dataset was rescaled to
168 match the mean monthly LAI output from the CLM. First, the two LAI3g maps generated for each
169 month were averaged, then the LAI3g pixels were aggregated (averaged) to match the size of a CLM
170 grid cell (~165 pixels per grid cell). If more than 80% of the grid cell did not have values in LAI3g
171 (mostly applicable at high latitudes), the entire grid cell was removed from further analysis. Finally,
172 the aggregated LAI3g data was resampled using a nearest neighbor approach to align with the CLM
173 grid for further analysis. All spatial and statistical analyses were performed in R (R Core Team, 2013)
174 using the ncdf (Pierce, 2011), raster (Hijmans and van Etten, 2013) and rgdal (Bivand et al., 2013)
175 packages.

176 **2.3 Comparing LAI3g to CLM LAI**

177 We first compared LAI3g to the CLM output for 1982 (the first full year of LAI3g) to 2010
178 (the last available year of the CRU-NCEP forcing dataset for CLM), aggregating values by zones
179 based on dominant PFT and hemisphere. To aggregate into PFT classes, we only considered grid
180 cells dominated by a single drought deciduous PFT (>50% cover), permitting six possible
181 comparisons with a sufficient number of grid cells for comparison: northern hemisphere (NH)

182 temperate C3 grasses (n = 180), NH C4 grasses (n = 234), tropical deciduous trees (n = 242),
183 southern hemisphere (SH) deciduous shrubs (n = 68), SH C3 grasses (n = 160), and SH C4 grasses
184 (n = 271). Note that the actual number of grid cells compared from year to year varied slightly with
185 changes in land cover and, in the case of LAI3g, available data. The counts listed here are the
186 averages for each PFT. To visually assess the comparison between LAI3g and CLM we plotted the
187 monthly means and standard deviations for these seven regions. We also computed the R^2 and root
188 mean squared error (RMSE) across all 29 years' monthly values to assess CLM's ability to fit the
189 seasonality and the magnitude of the LAI3g values. Due to the temporal coarseness of these data
190 and the irregular seasonal patterns found in many of our areas of interest, we did not fit a
191 continuous function to these data.

192 **2.4 Point Simulations and Parameter Sensitivity Tests**

193 Given the observed mismatches between LAI3g seasonality and CLM predictions among
194 drought deciduous woody plants (see Discussion), we conducted an analysis of how the parameters
195 determining phenology in the model affect the model outcome. The use of global or regional
196 simulations to assess the sensitivity of models to their structural and parametric assumptions is
197 problematic on account of both the computational requirements to do so and the high-
198 dimensionality of the model outputs, which can hinder understanding. To avoid these issues, we
199 conducted a sensitivity test to the major model parameters that control seasonality in LAI at the
200 point scale. In these point simulations we focused on low-latitude drought-deciduous ecosystems,
201 and selected six locations dominated by tropical deciduous trees, grasses, or a combination of the
202 two (**Table 1**).

203 The phenology model contains three empirical parameters that collectively describe the leaf
204 onset and offset algorithms.

1. Critical soil water potential ($\Psi_{threshold}$) values for leaf onset and offset (default = -2 MPa)
2. Soil water potential days to onset/offset (SWI/OSWI) threshold (default = 15 days)
3. Quantity of carbon assimilation which is directed to current leaf growth rather than storage ($F_{current}$) (default = 0.0)

To determine what impact the choice of these parameters has on the model outcome, we conducted a Latin-hypercube analysis (McKay et al., 1979), beginning the model from a spun-up state with default parameters. We then perturbed the parameters and ran the model forwards until a new LAI equilibrium condition was detected. Because the nitrogen cycle is active in CLM, soil biogeochemical equilibrium can in some circumstances (particularly at high latitudes) take many decades or even centuries to achieve, therefore we set a threshold for the new equilibrium state as the absence of a trend in LAI resulting in a 2% increase over a five year period. Given the high temperature and relatively low biomass and productivity of the ecosystems in question, in our simulations LAI equilibrium was in practice typically reached after the first 5 year period of the simulation. As with the global simulations, we used the CRU-NCEP reanalysis forcing data, extracted at the six points of interest, to drive the point simulations.

The Latin hypercube methodology is a ~~particularly efficient~~computationally inexpensive means of investigating a multi-dimensional parameter space, because each run of the model perturbs every parameter, and the algorithm ensures that the distribution of sampled points is distributed efficiently (but not uniformly) through parameter space. For this study, since we were comparing the model output to monthly satellite observations, we did not manipulate the number of days for onset (30) or offset (15) to occur. We also did not consider the growing degree days and day length parameters because the focus of this study was on tropical and subtropical regions where these components of the algorithm are not active.

For the critical soil moisture potential threshold, $\Psi_{threshold}$, we investigated values from 0 to -3.5MPa (where the default value is -2 MPa). The upper end of this range is the maximum possible value for saturated soils, whereas the bottom end of the range was determined from a prior set of sensitivity tests which determined that sensitivity below this range was very low (i.e. the soil moisture potential in the third layer rarely drops below -3.5 MPa, and so leaves remain on continuously for those simulations at our locations of interest). For the number of days of onset, SWI, we followed a similar protocol~~,~~ and found that the range of sensitivity was focused between 5 and 35 days (where the default is 15 days). For the fraction of displayed assimilated carbon ($F_{current}$) we varied the values between 0 and 0.5, (where the default is zero). Sensitivity of average LAI to this parameter was low in all cases, but it has an impact on the intraannual cycle, since the LAI is unchanging through a single growing season if $F_{current} = 0$.

To assess the performance of the different models in the Latin hypercube test we originally plotted the coefficients of determination between the different models' LAI values and the LAI3g data at those points. However, this result did not illustrate any clear optimum in model performance either for the parameters of the existing model, nor for the rainfall threshold. We illustrate this using the time-series data in Fig. 4, which highlights the unusual behavior of the model, and to assess whether the extra green-up period during the dry season had been eliminated in any of the parametric permutations. We ascribe the lack of a clear parametric signal to two effects. First, the LAI3g data were necessarily aggregated to monthly values, meaning that the primarily sub-monthly variation between ensemble members was masked. Second, the timing of the secondary- leaf-on period in the dry season was the emergent property of the oscillatory (and thus somewhat chaotic) dynamics of the soil-vegetation feedback on soil moisture. We thus conclude that the model deficiency is caused by structural, not parametric, issues.

Once we determined that we could not eliminate the dry season green-up by changing the existing model parameters, we considered four possible additions to the model. The first three are described here but, for brevity, are not quantified in the results. First, we considered that using the third soil layer in CLM may be an arbitrary choice of soil depth, and that usage of the soil moisture potential derived drought index ('BTRAN', (Oleson et al., 2013)), which is weighted by vertical root fraction across the whole rooting depth profile, might provide a more physiologically relevant metric and be less prone to increases due to upwards moisture diffusion in the dry season. However, since the exponential root profile in the CLM weights the top soil layers (including layer 3) more strongly than the lower layers with fewer roots, this metric was just as prone to increasing water potential during the dry season as soil water potential in the third soil layer.

Second, we implemented leaf onset as a function of a total column soil moisture content threshold rather than soil moisture potential. We postulated that the redistribution of water causes the erroneous behavior and that this would not impact total column moisture. However, the establishment a single global threshold for total soil moisture is challenging, as a number of different variables impact soil moisture, including the variation in soil water retention capacities between different land points, and by the interaction between leaf area, evaporation rate and deep soil moisture content. Variation in rainfall and evaporation rates affects the equilibrium water content of deep soils, which changes the total column soil moisture content between locations and years, but not the physiologically relevant upper soil moisture potential. Therefore, we abandoned this possible driver of drought deciduous phenology.~~metric of phenological trigger.~~

Third, we considered a metric of triggering leaf flush by the rate of change of total column soil moisture, rather than soil moisture potential. However, this methodology also generates erroneous behavior, on account of the ability of the CLM hydrology model to extract water from the water table or aquifer along a water potential gradient. Thus, when water potential is low in the

bottom soil layer in the dry season, the rate of change of total soil moisture can be positive without any input from rainfall.

2.5 Rainfall Model

To correct biases uncovered in the model output (described below) we introduced a simple trigger into the model, that time-averaged 10-day precipitation must exceed a given threshold before leaf onset is triggered. This approach requires the addition of a new parameter, *rain_threshold*, into the model, which is the threshold over which the sum of precipitation over 10 days must be for leaf-on to occur. Leaf onset is thus triggered if 10 day rain is higher than *rain_threshold* and if the SWI is greater than 15 days.

We then used a Latin hypercube approach again to determine the sensitivity of the model to *rain_threshold* at our six chosen geographical points. We considered a range of rainfall rates, requiring that it rain 20 mm over the course of 5 to 60 days in order for plants to begin growing leaves. To test the global impact of these parameter changes we ran CLM with the new rainfall-based trigger and compared the results both at several points and globally.

2.6 Global Simulations

We used a number of different metrics to globally compare CLM to the LAI3g data and, later, to the modified version of the model (CLM-MOD). First we compared maps of maximum annual LAI and differences between the three maps. We also developed an algorithm to count the number of LAI peaks per year in all three data sets on grid cells with an LAI range greater than one, by counting the number of times per year that the difference between one month's LAI and the next was negative, then taking the mode across all 29 years. Finally, we calculated the coefficient of determination (R^2) in each grid cell, comparing the monthly LAI3g data to CLM and CLM-MOD to

identify areas with strong agreement between the remotely sensed data and the models, and areas with weak relationships.

The recent focus on land model benchmarking has led to a number of additional suggested methods for assessing seasonality in models compared to data (e.g. Randerson *et al.*, 2009, Kelley *et al.*, 2013), however, none of the proposed metrics would capture the central issue addressed in this paper – model output with two or more peaks per year, data with only one – as they begin with the unstated assumption that seasonality is unimodal over the course of a year, as do measures of the start and end of the growing season. In Randerson *et al* (2009) seasonality is assessed by identifying the month of peak LAI and comparing that to MODIS LAI (MOD15A2), and in Kelley et al (2014) several more complicated metrics are introduced (equations 7-9) to again produce single numbers to compare a model's seasonality to a benchmark data set. In these examples, as in other benchmarking studies, the focus is on producing a single number, which, while useful, can miss important details.

3. Results

3.1 Seasonal Patterns in CLM

We found generally good agreement between LAI3g and CLM averaged across grass-dominated regions. In a comparison of monthly values from 1982 to 2010 for the single PFT dominated regions in **Fig. 1**, R^2 values ranged from 0.54 to 0.9 (**Table 2**) with the majority of the grass R^2 greater than 0.7. **Figure 2** shows the monthly values across all years, and we see similar results – generally good correspondence, especially in seasonal pattern, between LAI3g and the CLM runs in the grass-dominated regions. The root mean squared error (RMSE) values in **Table 2**, as well as **Fig. 2B** and **C** show that CLM does not always capture the appropriate LAI values in grasslands, but the seasonal cycle is reasonably correct.

In contrast, CLM does not successfully capture phenological patterns or values in areas dominated by woody drought deciduous vegetation. Among tropical deciduous trees CLM predicted LAI appears to be both too high and out of phase with the satellite observations (**Fig. 2E**) while CLM shows no apparent seasonality among deciduous shrubs in the southern hemisphere (**Fig. 2F**), while LAI3g shows a slight cycle ranging from 0.4 to 0.7 LAI.

3.2 Point simulations & sensitivity tests

To look more closely at seasonal patterns in drought deciduous locations we selected six points around the globe across a range of latitudes dominated by a mixture of broadleaf deciduous tropical trees, C3 and C4 grasses (**Table 1**), (all of which use the same stress deciduous phenology algorithm). To better understand the phenological patterns, we re-ran CLM using the same methods as described above but recording daily outputs of relevant parameters including LAI, soil water potential, rainfall, and others. Plots of the seasonal cycles at these specific points using daily model output (solid green lines in **Fig. 3**) revealed a pattern whereby CLM appears to put leaves on during the "brown season" in the LAI3g data in some of the points in addition to during the LAI3g green season. We note, however, that some areas in reality do have two separate growing seasons per year (e.g. **Fig. 3E**). Despite the lack of rainfall, soil water potential in the third soil layer in CLM rises during the dry season and is extremely variable in the dry season, on account of periods of high transpiration when plants leaf out (blue dot-dashed line in **Fig. 3**).

We used the output from the Latin hypercube approach at these six points to vary the parameters of interest (days to onset/offset, critical soil water potential, carbon assimilation) to assess whether modification of parameter values could ameliorate the problem of plants leafing out during the dry season in CLM. We found, however, that simply varying the parameters of the

existing model within the parameter space investigated (and assuming no large non-linearities in the model response surface) did not remove the dry season leaf out in the model (**Fig. 4**).

In order to address this issue, we considered a number of structural perturbations to the leaf-on and leaf-off algorithms (described in the discussion below), but ultimately decided on adding a new parameter, *rain_threshold*, to the model. We then used the same Latin hypercube approach to determine the best fitting values for this parameter (**Fig. 5**). This additional leaf-on criterion, set so that 20 mm of rain must accumulate over 10 days in order for leaf onset to occur, led to a removal of the "brown season" leaf out in CLM (dashed green line in **Fig. 3**) without preventing two green seasons per year, as is possible in some semi-arid regions (e.g. parts of Ethiopia, **Fig. 3E**). While this new rainfall threshold improved model performance both at our points and globally (see below), we note that the model did not appear to be particularly sensitive to the amount of rain that fell, as long as some rain did fall, but this threshold, and the drought deciduous algorithm as a whole, deserves more research into seasonal drivers.

3.3 Global simulations

To test how well the additional rainfall parameter performed globally, we ran CLM with the new rainfall parameter for 45 years (CLM-MOD) from the same equilibrium baseline state as was used in the first run described. Measures of maximum LAI (**Fig. 6**) and mean LAI (data not shown) in CLM-MOD showed closer matches to LAI3g than CLM. While CLM values remain far too high in the evergreen tropics, the maximum LAI values in deciduous savanna regions did increase appropriately in CLM-MOD to better match the LAI3g data.

To test whether the poor fit between CLM and LAI3g was due to multiple annual LAI peaks in CLM we counted the number of peaks per year in each data set (**Fig. 7**). We found that in the observations, only areas in the humid tropics had multiple peaks in the LAI3g data ("peaks" in these

365 cases being relatively small fluctuations), while CLM showed multiple peaks per year throughout
366 many of the savanna regions of the world. CLM-MOD has more areas with only one peak,
367 particularly in Sub-Saharan Africa. To quantify these changes to the model we constructed
368 confusion matrices to compare the peak counts in LAI3g to those in CLM and CLM-MOD (Table
369 3) for grid cells with >50% drought deciduous cover (Fig. 1). Overall, CLM-MOD had a slightly
370 poorer performance, matching the number of peaks in the LAI3g dataset 42.5% of the time, while
371 CLM matched LAI3g 43.7% of the time. However, these unweighted summary numbers mask
372 improvements in CLM-MOD. CLM only correctly predicted a single peak 8.9% of the time, while
373 CLM-MOD correctly predicted single peaks 59% of the time, and ~~never did~~ CLM-MOD never
374 predicted more than two peaks in a year, matching the LAI3g data. The overall degradation in CLM-
375 MOD is due to fewer correctly identified grid cells with zero or two peaks.

376 We compared monthly data and mapped the point-wise coefficients of determination (R^2)
377 globally to consider how well CLM LAI seasonality matched the LAI3g dataset (**Fig. 8A**). There
378 were moderate to good relationships (R^2 >0.4) in the higher latitudes for the standard model, but
379 notably poorer relationships in the lower latitudes, particularly in savanna regions. In contrast, a
380 comparison between LAI3g and CLM-MOD showed improvements in savanna regions, with the
381 most dramatic improvements in Sub-Saharan Africa (**Fig. 8B**). Predicting the phenology of the
382 Brazilian Cerrado continues to be a challenge in CLM-MOD, shifting from two peaks in CLM to no
383 peaks in CLM-MOD, but the heterogeneity revealed in the LAI3g dataset suggests that this region
384 may need closer consideration and a separate phenology algorithm.

385

386 # 4. Discussion

387 ## 4.1 Comparing LAI3g to CLM LAI

By comparing the satellite LAI3g data to output from CLM we found that while the model performed reasonably well in temperate grasslands both in seasonal cycle and in magnitude, it performed poorly in areas of tropical grasslands, mixed grassland and drought deciduous trees (savannas) and areas dominated by drought deciduous trees. Closer examination of individual grid cells in tropical regions revealed that these points often have a leaf flush during the dry season in the model, which is not the case in the satellite data, or in reality. This additional leaf flush not only impacts the phenological cycle, but also affects the overall amount of carbon stored in plants and their maximum LAI, as plants spend their stored carbon unnecessarily in the dry season, leaving less carbon available during the wet season for growing leaves. This addition of leaf carbon in the dry season also may affect the fire cycle in varying ways around the dry tropics. While these runs of the model were not coupled to a dynamic atmosphere, we expect that this dry season leaf flush could also impact the climate, potentially having an unrealistic cooling effect by moving more water in to the atmosphere during what should be a very dry time of year, but also darkening the land surface, possibly leading to a slight warming.

The mechanism behind the dry season leaf flush is an increase in soil water potential in the dry season to levels above the prescribed leaf-out threshold. These increases derive from the assumption in CLM that all of the land surface ~~in CLM~~ sits on top of an unconfined aquifer. In most cases this aquifer is either irrelevant because plenty of soil water is available or it is essential to plant survival in areas where aquifers do exist in the real world. In semi-arid systems, however, this extra pool of soil water becomes problematic in the dry season. The top soil layers dry out due to soil evaporation and, when plants are active, evapotranspiration, establishing a water potential gradient~~,~~ which causes water to be transferred by mass flow from the aquifer up through the soil column to the shallow soil layers~~,~~ until eventually the moisture potential reaches the trigger for plants to leaf out. Per the drought deciduous phenology algorithm, once leaf out is triggered it must

412 be completed, so plants begin to grow leaves but then the increased evapotranspiration rate quickly
413 draws the soil moisture down below leaf off threshold levels, so leaf drop begins again, typically as
414 soon as the leaf out period (30 days) has ended. The degree to which aquifers in reality contribute to
415 dry season evapotranspiration is largely unconstrained because there are no global data sets for
416 depth to water table, making it impossible to non-arbitrarily define where plants should have access
417 to ground water and where they should not. Refinements of the soil water algorithms in CLM and
418 access to new data sources like the NASA Soil Moisture Active Passive mission (SMAP; Entekhabi
419 *et al.*, 2014) will likely improve this part of the model, but like many aspects of the CLM, more
420 global-scale data is needed.

421 **4.2 Soil water and rainfall in CLM**

422 To address the erroneous dry season leaf flush we tested a number of different model
423 alterations, beginning with the least invasive – adjusting existing parameters – and ending with
424 adding an additional rule to the drought deciduousness algorithm. We experimented with four
425 alternative methodologies for triggering leaf onset, described in the methods (section 2.4), but for
426 brevity we have only shown results from the last and most effective approach.

427 The hydrological issues in CLM are complex, and derive from the need to operate an
428 internally consistent global model of the water cycle in the absence of critical data at the appropriate
429 scale (depth to water table, the unsaturated hydraulic conductivity of deep soils, etc). In an ideal
430 case, improvements in hydrology might allow the existing phenological model to operate correctly.
431 However, here we took a more pragmatic approach and so we partially decoupled the soil hydrology
432 and the phenology models, allowing rainfall inputs to directly impact on leaf phenology without
433 interacting with the assumptions of the hydrology model. Leaving the condition of soil water
434 potential in the third soil layer in place, we then added an additional condition which was that the

435 rainfall accumulated over the last 10 days should be higher than a threshold value (20 mm). Thus, if
436 soil moisture rose above the threshold level, but little or no rain had fallen, plants would not put on
437 leaves. The new model performs better both for the point simulations and in global simulations,
438 both in terms of the seasonal cycle of LAI, where the average point-wise coefficient of
439 determination (R^2) between modelled and observed monthly satellite LAI of drought deciduous
440 dominated points (>50% drought deciduous cover) is significantly higher for the new model (0.31
441 vs 0.13). While there was no substantial change in the overall peak count accuracy (Table 3), CLM-
442 MOD had zero drought deciduous dominated grid cells with >2 peaks and a substantial
443 improvement in the identification of single-peak grid cells (8.9 to 59%). The added rainfall trigger
444 did, however, reduce the number of zero peak and two peak grid cells correctly identified. This
445 result highlights the need for more research into the diversity of drought deciduous phenology
446 drivers around the world.

447 **4.3 Impacts of modifications to the model**

448 This relatively small change to the drought deciduous phenology algorithm had wide ranging
449 impacts within the CLM. Because carbon was not being unnecessarily spent to grow leaves during
450 the dry season, which was then not replenished since there was not enough water to maintain
451 photosynthesis, CLM-MOD showed substantially higher overall carbon stores in savanna regions
452 (**Fig. 10**, blue lines). Over time, this increase in vegetation carbon could lead to more realistic soil
453 carbon levels, which have been shown to be too low in savanna regions in CLM (Wieder et al.,
454 2013).

455 Fire is a critical component of savanna ecology and has been a focus of recent efforts to
456 improve the CLM (Li et al., 2014). Our change to the drought deciduous phenology algorithm does
457 have an impact on the fire cycle, but unfortunately, though not surprisingly, it degrades the fire

458 model's performance relative to a global fire data set (GFED4; Giglio *et al.*, 2013). Comparing the
459 average total annual fire fractions for each grid cell with drought deciduous cover greater than 50%
460 across the time period for which we have both GFED4 data and CLM output (1996-2010) we find a
461 correlation between GFED4 and CLM of 0.35 (global correlation = 0.44), and a correlation of 0.23
462 with CLM-MOD (global correlation = 0.33). This degradation of fire model performance is not
463 surprising, however, given that the fire model was developed using CLM4.0CN with the erroneous
464 dry season green up and a different forcing dataset. As shown in **Fig. 9**, fraction of area burned per
465 grid cell decreases in many areas in CLM-MOD, likely due to the fact that less fuel is being produced
466 in the dry season, and seasonality shifts as well (**Fig. 10**, red lines). Future work will include
467 exploring the impacts of this change to CLM on fire and other ecosystem properties.

468 Two other outstanding questions about LAI in CLM remain. First, in savanna regions in the
469 CLM LAI drops to zero during the dry season, implying that across an entire grid cell all vegetation
470 is perfectly drought deciduous (**Figs. 3-6**). Reality is, of course, far more complex, as reflected in the
471 LAI3g dataset which rarely drops below 1.0 in savanna regions when aggregated to the CLM grid.
472 The focus of this study was on improving the timing and magnitude of peak LAI, however,
473 improving dry season values is also a concern. This is a deeper question in the CLM, as it relates to
474 the overall land cover data. It is possible, for example, that there is evergreen vegetation in these grid
475 cells, while the land cover classification determines that all grid cells with a significant seasonal LAI
476 signal are 100% 'drought deciduous'. Even if drought deciduous and evergreen vegetation types did
477 not co-exist in the same ecosystem type, within grid cell spatial heterogeneity might also allow their
478 coexistence within a whole grid cell (e.g. riparian areas or areas with shallow ground water that are
479 able to stay green year round). Second, though also not the focus of this study, it is clear from **Fig. 6**
480 that across the mesic regions of the terrestrial biosphere CLM is dramatically overestimating LAI.
481 This issue, often masked when only mean annual values or zonal means are considered, deserves

more attention, and it is likely that recent detailed studies of carbon allocation (e.g. Doughty *et al.*, 2014) could improve this part of the model.

A question also remains as to whether our new representation of leaf phenology, in spite of its improved performance, constitutes a better predictive model of current and future ecosystem behavior. In general, we hope to construct ecosystem models that represent hypotheses of how plants function, that we might test against observations. In this case, we find that the existing hypothesis - that plants respond to the soil water potential of the upper soil - does not adequately represent phenological patterns. However, this explanation is complicated because the predictions depend also on the properties of soil water in the model. Given a perfect representation of soil moisture, we might find that the existing leaf-on hypothesis is a good approximation of average vegetation behavior. However, at present the coupling of these two complex systems produces unexpected results. By tying the vegetation behavior to the actual climate drivers we are reducing the complexity of the problem, however, we are also reducing the capacity of the model to be responsive to the nuances of climate drivers. For example, the same rainfall amount in high and low humidity regimes will have different impacts on net soil moisture.

Ideally, models should represent the mechanisms by which ecological processes operate in as much fidelity as we understand. The representation of drought phenology is interesting; however, as we suspect that there are many different phenological strategies in the tropics that the CLM classifies with the same algorithm (e.g. Archibald & Scholes, 2007). This means that, in the absence of the representation of these numerous phenological strategies in the model, we are really representing the net behavior of ecosystems, rather than the exact mechanisms pertaining to a single species. The fact that CLM-MOD improved model performance most significantly in Africa and less so in Australia and South America by some metrics (**Fig. 7**) suggests that evolutionary differences between plants could play a significant role in determining phenological patterns between continents. In a higher-

506 fidelity land surface model, we might ideally allow numerous phenological algorithms to compete for
507 light and water resources, and the ecosystem LAI profile would reflect the net behavior of the
508 successful algorithms. This type of modelling is now theoretically possible (e.g. Fisher et al., 2010,
509 2015), and will be investigated in future versions of the CLM.

510

511 ## 5. Conclusions

512 By comparing satellite derived estimates of LAI to LAI values produced by the latest version
513 of the CLM we revealed a small but significant issue in the CLM – the tendency for leaves to flush
514 during the dry season in drought deciduous PFTs due to unrealistic upwards movement of water
515 through the soil column. We tested a number of different approaches to address this issue, however
516 we found that tying leaf flushing to rainfall directly produced results that better matched the satellite
517 data. While this change to the drought deciduous phenology algorithm does not reflect our
518 understanding of how plants respond to their environment in the real world, without better data on
519 soil water movement at scales relevant to global land surface modeling it is difficult to rely on the
520 soil water model to drive plant physiology. Changing the drought deciduous phenology algorithm to
521 remove dry season leaf flushes improved overall LAI values in savanna systems as well as changed
522 the amount of carbon stored in these systems and altered the fire cycle. We also emphasize that this
523 issue would have been impossible to detect with a standard 'benchmarking' type of metric for
524 measuring seasonality and was difficult to identify until daily model outputs were reported and
525 analyzed (i.e. Fig. 3). Future work will include exploring different drought deciduous phenology
526 algorithms for different PFTs and testing the importance of this change in a coupled Earth system
527 model.

528

Acknowledgements

The authors thank the members of the Terrestrial Sciences Section at NCAR for helpful discussions of this work, and we thank Dr. Ranga Myneni and his group for providing the LAI3g dataset. KMD was funded by an Advanced Study Program Postdoctoral Fellowship at NCAR. We would also like to acknowledge high-performance computing support from Yellowstone (ark:/85065/d7wd3xhc) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

REFERENCES

Archibald, S. and Scholes, R. J.: Leaf green-up in a semi-arid African savanna – separating tree and grass responses to environmental cues, J. Veg. Sci., 18, 583–594, 2007.

Bivand, R., Keitt, T. and Rowlingson, B.: rgdal: Bindings for the geospatial data abstraction library, [online] Available from: http://cran.r-project.org/package=rgdal, 2013.

Blyth, E., Clark, D. B., Ellis, R., Huntingford, C., Los, S., Pryor, M., Best, M. and Sitch, S.: A comprehensive set of benchmark tests for a land surface model of simultaneous fluxes of water and carbon at both the global and seasonal scale, Geosci. Model Dev., 4(2), 255–269, doi:10.5194/gmd-4-255-2011, 2011.

Bonan, G. B., Levis, S., Sitch, S., Vertenstein, M. and Oleson, K. W.: A dynamic global vegetation model for use with climate models : concepts and description of simulated vegetation dynamics, Glob. Chang. Biol., 9, 1543–1566, doi:10.1046/j.1529-8817.2003.00681.x, 2003.

Bradley, A. V., Gerard, F. F., Barbier, N., Weedon, G. P., Anderson, L. O., Huntingford, C., Aragão, L. E. O. C., Zelazowski, P. and Arai, E.: Relationships between phenology, radiation and precipitation in the Amazon region, Glob. Chang. Biol., 17(6), 2245–2260, doi:10.1111/j.1365-2486.2011.02405.x, 2011.

Delbart, N., Le Toan, T., Kergoat, L. and Fedotova, V.: Remote sensing of spring phenology in boreal regions: A free of snow-effect method using NOAA-AVHRR and SPOT-VGT data (1982–2004), Remote Sens. Environ., 101(1), 52–62, doi:10.1016/j.rse.2005.11.012, 2006.

Doughty, C. E., Malhi, Y., Arujo-Murakami, A., Metcalfe, D. B., Silva-Espejo, J. E., Arroyo, L., Heredia, J. P., Pardo-Toledo, E. and Mendizabal, L. M.: Allocation trade-offs dominate the response of tropical forest growth to seasonal and interannual drought, , 95(8), 2192–2201, 2014.

Entekhabi, D., Yueh, S., O'Neill, P. E., Kellogg, K. H., Allen, A., Bindlish, R., Brown, M., Chan, S., Colliander, A. and Crow, W. T.: SMAP Handbook., 2014.

Fisher, R., McDowell, N., Purves, D., Moorcroft, P., Sitch, S., Cox, P., Huntingford, C., Meir, P. and Ian Woodward, F.: Assessing uncertainties in a second-generation dynamic vegetation model caused by ecological scale limitations, New Phytol., 187, 666–681, doi:10.1111/j.1469-8137.2010.03340.x, 2010.

Fisher, R. a., Muszala, S., Verteinstein, M., Lawrence, P., Xu, C., McDowell, N. G., Knox, R. G., Koven, C., Holm, J., Rogers, B. M., Lawrence, D. and Bonan, G.: Taking off the training wheels: the properties of a dynamic vegetation model without climate envelopes, Geosci. Model Dev. Discuss., 8(4), 3293–3357, doi:10.5194/gmdd-8-3293-2015, 2015.

Giglio, L., Randerson, J. T. and Van Der Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), J. Geophys. Res. Biogeosciences, 118(November 2012), 317–328, doi:10.1002/jgrg.20042, 2013.

572 Guan, K., Wood, E. F., Medvigy, D., Kimball, J., Pan, M., Caylor, K. K., Sheffield, J., Xu, X. and
573 Jones, M. O.: Terrestrial hydrological controls on land surface phenology of African savannas and
574 woodlands, J. Geophys. Res. Biogeosciences, 119, 1652–1669, doi:10.1002/2013JG002572.Received,
575 2014.

576 Hijmans, R. J. and van Etten, J.: raster: Geographical data analysis and modeling, [online] Available
577 from: http://cran.r-project.org/package=raster, 2013.

578 Jenerette, G. D., Scott, R. L. and Huete, A. R.: Functional differences between summer and winter
579 season rain assessed with MODIS-derived phenology in a semi-arid region, J. Veg. Sci., 21(1), 16–30,
580 doi:10.1111/j.1654-1103.2009.01118.x, 2010.

581 Kelley, D. I., Prentice, I. C., Harrison, S. P., Wang, H., Simard, M., Fisher, J. B. and Willis, K. O.: A
582 comprehensive benchmarking system for evaluating global vegetation models, Biogeosciences, 10(5),
583 3313–3340, doi:10.5194/bg-10-3313-2013, 2013.

584 Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S., Collins, W. D., Bonan, G. B.,
585 Lawrence, D. M. and Swenson, S. C.: The effect of vertically resolved soil biogeochemistry and
586 alternate soil C and N models on C dynamics of CLM4, Biogeosciences, 10(11), 7109–7131,
587 doi:10.5194/bg-10-7109-2013, 2013.

588 Lau, W. K.-M., Wu, H.-T. and Kim, K.-M.: A canonical response of precipitation characteristics to
589 global warming from CMIP5 models, Geophys. Res. Lett., 40(12), 3163–3169,
590 doi:10.1002/grl.50420, 2013.

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

595 Lawrence, D. M., Oleson, K. W., Flanner, M. G., Fletcher, C. G., Lawrence, P. J., Levis, S.,
596 Swenson, S. C. and Bonan, G. B.: The CCSM4 Land Simulation, 1850–2005: Assessment of Surface
597 Climate and New Capabilities, J. Clim., 25(7), 2240–2260, doi:10.1175/JCLI-D-11-00103.1, 2012.

598 Lawrence, P. J. and Chase, T. N.: Representing a new MODIS consistent land surface in the
599 Community Land Model (CLM 3.0), J. Geophys. Res., 112(G1), G01023,
600 doi:10.1029/2006JG000168, 2007.

601 Li, F., Bond-Lamberty, B. and Levis, S.: Quantifying the role of fire in the Earth system – Part 2:
602 Impact on the net carbon balance of global terrestrial ecosystems for the 20th century,
603 Biogeosciences, 11(5), 1345–1360, doi:10.5194/bg-11-1345-2014, 2014.

604 Ma, X., Huete, A., Yu, Q., Coupe, N. R., Davies, K., Broich, M., Ratana, P., Beringer, J., Hutley, L.
605 B., Cleverly, J., Boulain, N. and Eamus, D.: Spatial patterns and temporal dynamics in savanna
606 vegetation phenology across the North Australian Tropical Transect, Remote Sens. Environ., 139,
607 97–115, doi:10.1016/j.rse.2013.07.030, 2013.

McKay, M. D., Beckman, R. J. and Conover, W. J.: Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, Technometrics, 21(2), 239–245, doi:10.1080/00401706.1979.10489755, 1979.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C., Levis, S., Li, F., Riley, W., Subin, Z., Swenson, S., Thornton, P. E., Bozbiyik, A., Fisher, R., Heald, C., Kluzek, E., Lamarque, J.-F., Lawrence, P., Leung, L., Lipscomb, W., Muszala, S., Ricciuto, D., Sacks, W., Sun, Y., Tang, J. and Yang, Z. L.: Technical description of version 4.5 of the Community Land Model (CLM), NCAR Tech. Note, 503+STR(June), doi:10.5065/D6RR1W7M, 2013.

Parmesan, C. and Yohe, G.: A globally coherent fingerprint of climate change impacts across natural systems., Nature, 421(6918), 37–42, doi:10.1038/nature01286, 2003.

Pierce, D.: ncdf: Interface to unidata netCDF files, [online] Available from: http://cran.r-project.org/package=ncdf, 2011.

Randerson, J. T., Hoffman, F. M., Thornton, P. E., Mahowald, N. M., Lindsay, K., Lee, Y.-H., Nevison, C. D., Doney, S. C., Bonan, G., Stöckli, R., Covey, C., Running, S. W. and Fung, I. Y.: Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models, Glob. Chang. Biol., 15(10), 2462–2484, doi:10.1111/j.1365-2486.2009.01912.x, 2009.

RCoreTeam: R: A language and environment for statistical computing, 2013.

Reed, B. C., Schwartz, M. D. and Xiao, X.: Remote sensing phenology: Status and the way forward, in Phenology of Ecoystem Processes: Applications in global change research, edited by A. Noormets, pp. 231–246, Springer., 2009.

Reich, P. B.: Phenology of tropical forests: Patterns, causes, and consequences, Can. J. Bot., 73, 164–174, 1995.

Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., Chen, J. M., Ciais, P., Davis, K. J., Desai, A. R., Dietze, M. C., Dragoni, D., Garrity, S. R., Gough, C. M., Grant, R., Hollinger, D. Y., Margolis, H. a., McCaughey, H., Migliavacca, M., Monson, R. K., Munger, J. W., Poulter, B., Raczka, B. M., Ricciuto, D. M., Sahoo, A. K., Schaefer, K., Tian, H., Vargas, R., Verbeeck, H., Xiao, J. and Xue, Y.: Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis, Glob. Chang. Biol., 18(2), 566–584, doi:10.1111/j.1365-2486.2011.02562.x, 2012.

Scholes, R. J. and Hall, D. O.: The carbon budget of tropical savannas, woodlands, and grassslands, in SCOPE 56 - Global Change: Effects on Coniferous forests and grasslands, edited by A. I. Breymeyer, D. O. Hall, J. M. Melillo, and G. I. Agren, Wiley., 1996.

Staver, A. C., Archibald, S. and Levin, S. A.: The global extent and determinants of savanna and forest as alternative biome states., Science, 334(6053), 230–2, doi:10.1126/science.1210465, 2011.

642 Swenson, S. C. and Lawrence, D. M.: Assessing a dry surface layer-based soil resistance
643 parameterization for the Community Land Model using GRACE and FLUXNET-MTE data, J.
644 Geophys. Res. Atmos., 119(10), 299–312, doi:10.1002/2014JD022314.Received, 2014.

645 Wang, K., Mao, J., Dickinson, R., Shi, X., Post, W., Zhu, Z. and Myneni, R.: Evaluation of CLM4
646 Solar Radiation Partitioning Scheme Using Remote Sensing and Site Level FPAR Datasets, Remote
647 Sens., 5(6), 2857–2882, doi:10.3390/rs5062857, 2013.

648 White, M. a., Thornton, P. E. and Running, S. W.: A continental phenology model for monitoring
649 vegetation responses to interannual climatic variability, Global Biogeochem. Cycles, 11(2), 217–234,
650 doi:10.1029/97GB00330, 1997.

651 White, M. a., de BEURS, K. M., Didan, K., Inouye, D. W., Richardson, A. D., Jensen, O. P.,
652 O'Keefe, J., Zhang, G., Nemani, R. R., van LEEUWEN, W. J. D., Brown, J. F., de WIT, A.,
653 Schaepman, M., Lin, X., Dettinger, M., Bailey, A. S., Kimball, J., Schwartz, M. D., Baldocchi, D. D.,
654 Lee, J. T. and Lauenroth, W. K.: Intercomparison, interpretation, and assessment of spring
655 phenology in North America estimated from remote sensing for 1982-2006, Glob. Chang. Biol.,
656 15(10), 2335–2359, doi:10.1111/j.1365-2486.2009.01910.x, 2009.

657 Wieder, W. R., Bonan, G. B. and Allison, S. D.: Global soil carbon projections are improved by
658 modelling microbial processes, Nat. Clim. Chang., 3(10), 909–912, doi:10.1038/nclimate1951, 2013.

659 Yang, X., Mustard, J. F., Tang, J. and Xu, H.: Regional-scale phenology modeling based on
660 meteorological records and remote sensing observations, J. Geophys. Res., 117(G3), G03029,
661 doi:10.1029/2012JG001977, 2012.

662 Zhang, X., Friedl, M. A., Schaaf, C. B. and Strahler, A. H.: Monitoring the response of vegetation
663 phenology to precipitation in Africa by coupling MODIS and TRMM instruments, J. Geophys. Res.,
664 110(D12), D12103, doi:10.1029/2004JD005263, 2005.

665 Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R. and Myneni,
666 R.: Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of Photosynthetically
667 Active Radiation (FPAR)3g Derived from Global Inventory Modeling and Mapping Studies
668 (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the Period 1981 to 2, Remote
669 Sens., 5(2), 927–948, doi:10.3390/rs5020927, 2013.

670

671

672 **Table 1.** List of locations for point simulations and percent cover of plant functional types (PFTs).
673 PFTs with no coverage at any point are not listed.

point.name	Latitude	Longitude	Bare ground	Broadleaf Evergreen Tree Tropical	Broadleaf Evergreen Tree Temperate	Broadleaf Deciduous Tree Tropical	C3 Grasses	C4 Grasses	Crops
Brasilia	-15	-51	0.46	1.69	0	16.52	8.83	62.35	10.15
western Brazil	-6	-39	2.66	0	0	35.4	9.04	40.6	12.3
South Chad	11	18	1.34	0	0	34.26	0.03	60.22	4.16
eastern Zambia	-13	32	0.22	0.56	1.27	26.4	37.81	26.39	7.35
south Ethiopia	5.5	40	8.75	0.13	0.02	63.1	19.12	5.42	3.47
Darwin Australia	-15	130.5	15.94	0	0	35.73	0	48.33	0

674

675 **Table 2.** R^2 and RMSE of AVHRR LAI3g v. CLM monthly data for all 29 yrs.

		NH C3 grass	NH C4 grass	tropical DT	SH DS	SH C3 grass	SH C4 grass
CLM	R^2	0.79	0.72	0.08	0.05	0.41	0.85
	RMSE	0.15	0.29	0.18	0.10	0.27	0.13

676 * PFT = plant functional type, NH = northern hemisphere, DS = deciduous shrub, DT =
677 deciduous tree, SH = southern hemisphere

678

Table 3. Confusion matrices comparing grid cell peak counts between LAI3g and the two model data sets. "%" rows and columns are the percent of the correct values (diagonal) compared to the sums for the respective rows and columns.

A. LAI3g vs. CLM						
		LAI3g				
		0	1	2	>2	%
CLM	0	436	164	145	0	58.5
	1	28	74	130	0	31.9
	2	196	555	499	0	39.9
	>2	7	42	35	0	0
	%	65.4	8.9	61.7	NA	43.7

B. LAI3g vs CLM-MOD						
		LAI3g				
		0	1	2	>2	%
CLM-MOD	0	365	279	277	0	39.6
	1	242	439	407	0	43.2
	2	60	63	125	0	50.4
	> 2	0	0	0	0	NA
	%	54.7	59.0	15.5	NA	42.5

685 **FIGURE CAPTIONS**

686 **Figure 1.** (A) Areas of the globe dominated by a single PFT (>50%) grouped where appropriate
687 (e.g. there are actually three grass PFTs). (B) Percent cover of drought deciduous PFTs within the
688 natural vegetation component of each grid cell; gray areas have zero percent cover of drought
689 deciduous vegetation. For visual clarity, grid cells with < 50% natural vegetation (e.g. grid cells that
690 are mostly water) are not shown in both maps.

691 **Figure 2.** Annual LAI cycles for LAI3g and CLM averaged for 1982-2010; shaded areas represent
692 one standard deviation. Each plot is averaged across a region as shown in Fig. 1. (A) Northern
693 hemisphere (NH) C3 grasses; (B) NH C4 grasses; (C) Southern hemisphere (SH) C3 grasses; (D) SH
694 C4 grasses; (E) Tropical deciduous trees; (F) SH broadleaved deciduous shrubs.

695 **Figure 3.** Seasonal cycles of rainfall (mm day^{-1}, gray bars); leaf area index (LAI, green lines and black
696 dots) and soil water potential in the third layer (MPa, blue lines) in CLM and CLM-MOD for one
697 year (2001).

698 **Figure 4.** Illustration of Latin hypercube (LH) variable exploration analysis results – here each line
699 represents one simulation all from one year of the LH analysis without the additional rainfall trigger.
700 Each line is from a model run with slightly different values for the variables considered. In actuality
701 100 simulations were performed, but for visual clarity we are showing a selection of 10 simulations.

702 **Figure 5.** Illustration of Latin hypercube (LH) variable exploration analysis results as with Fig. 4 –
703 here each line represents one simulation all from one year of the LH analysis with the additional
704 rainfall trigger.

705 **Figure 6.** Maximum annual LAI averaged across the 29 year time period (1982-2010) in LAI3g,
706 CLM, and CLM-MOD, and the differences between these three maps

Figure 7. Mode of annual peak count analysis for the three simulations. (A) LAI3g; (B) CLM; (C) CLM-MOD

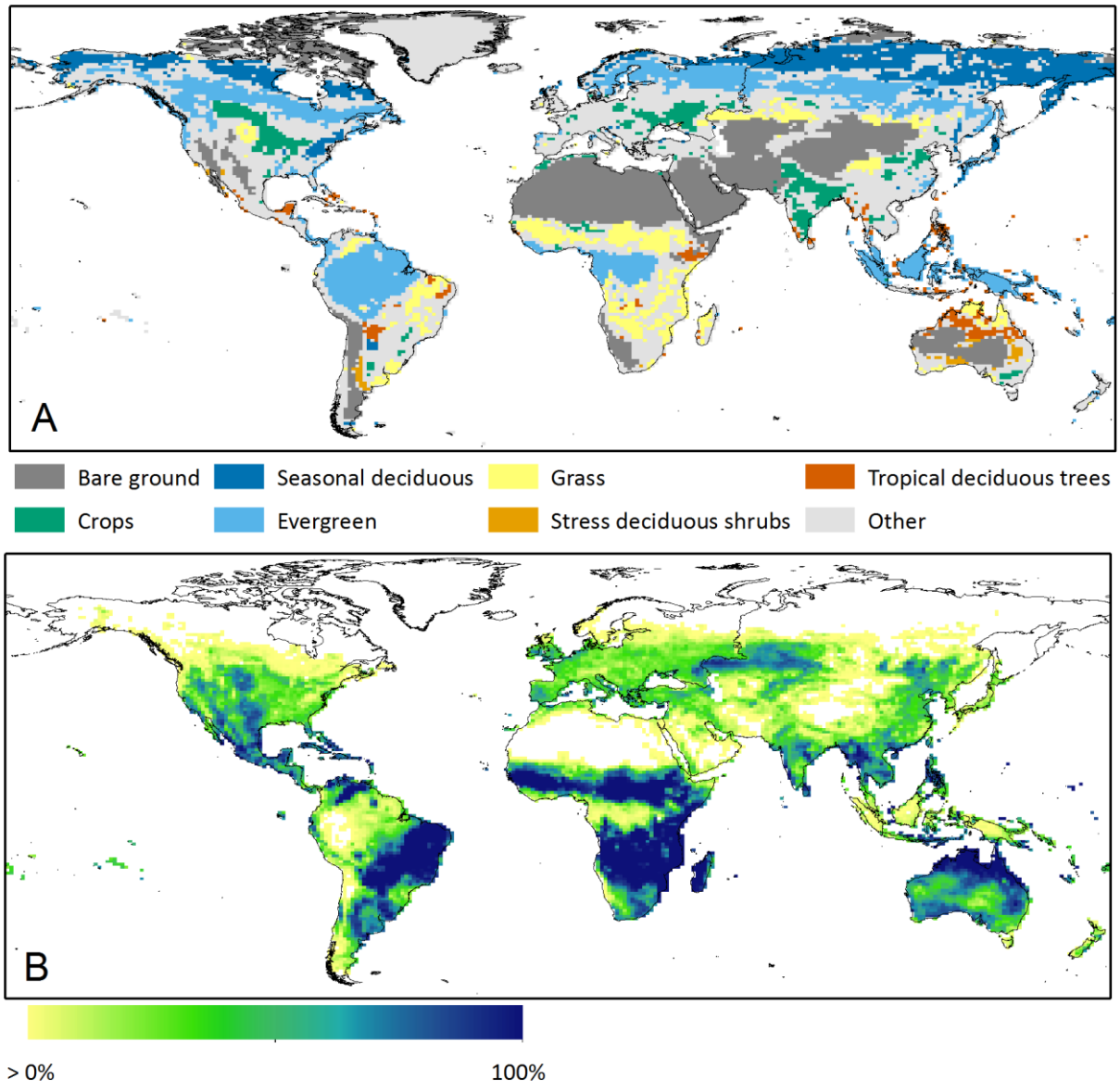
Figure 8. Coefficients of determination (R^2) between LAI3g and the two model versions.

Figure 9. Average burned area fraction per year across the time period where data was available (1996-2010) for GFED4, CLM, and CLM-MOD, and the differences between these maps.

Figure 10. Seasonal cycles of rainfall (mm day^{-1}; gray bars), total vegetation carbon (gC m^{-2}; blue lines), and grid cell burned fraction (day^{-1}; red lines) in CLM and CLM-MOD for one year (2001).

716

717



718

719 **Figure 1**

720

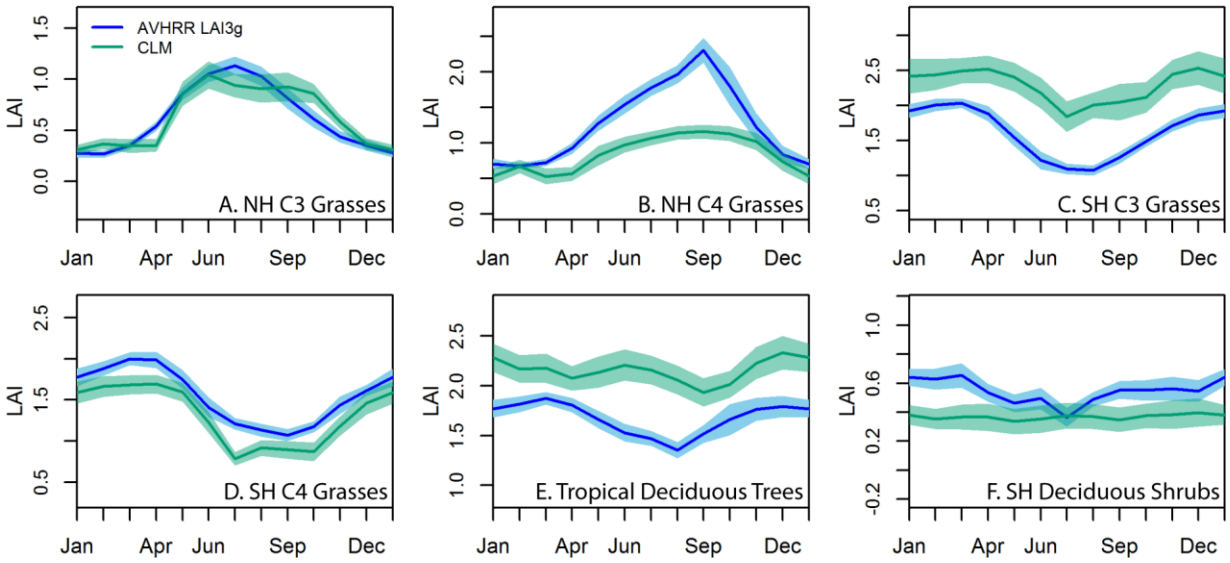

AVHRR LAI3g
CLM
LAI
A. NH C3 Grasses
B. NH C4 Grasses
C. SH C3 Grasses
D. SH C4 Grasses
E. Tropical Deciduous Trees
F. SH Deciduous Shrubs
Jan
Apr
Jun
Sep
Dec

Figure 2

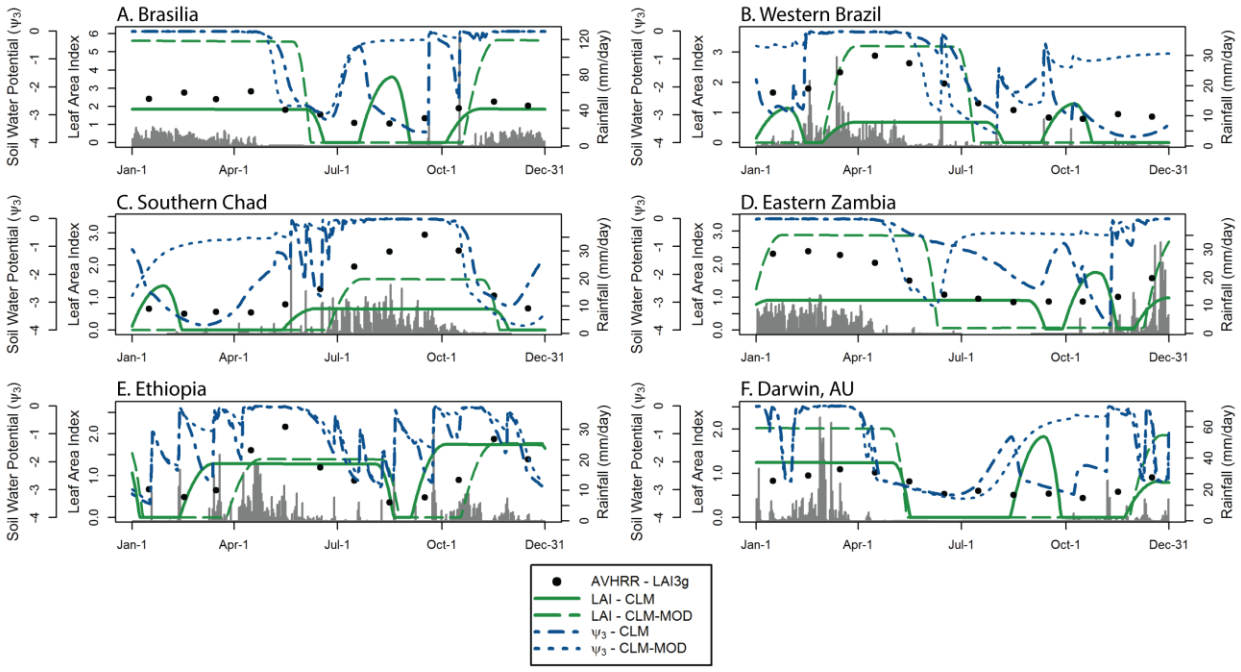

A. Brasilia
B. Western Brazil
C. Southern Chad
D. Eastern Zambia
E. Ethiopia
F. Darwin, AU
Soil Water Potential (ψ3)
Leaf Area Index
Rainfall (mm/day)
Jan-1
Apr-1
Jul-1
Oct-1
Dec-31
AVHRR - LAI3g
LAI - CLM
LAI - CLM-MOD
ψ3 - CLM
ψ3 - CLM-MOD

Figure 3

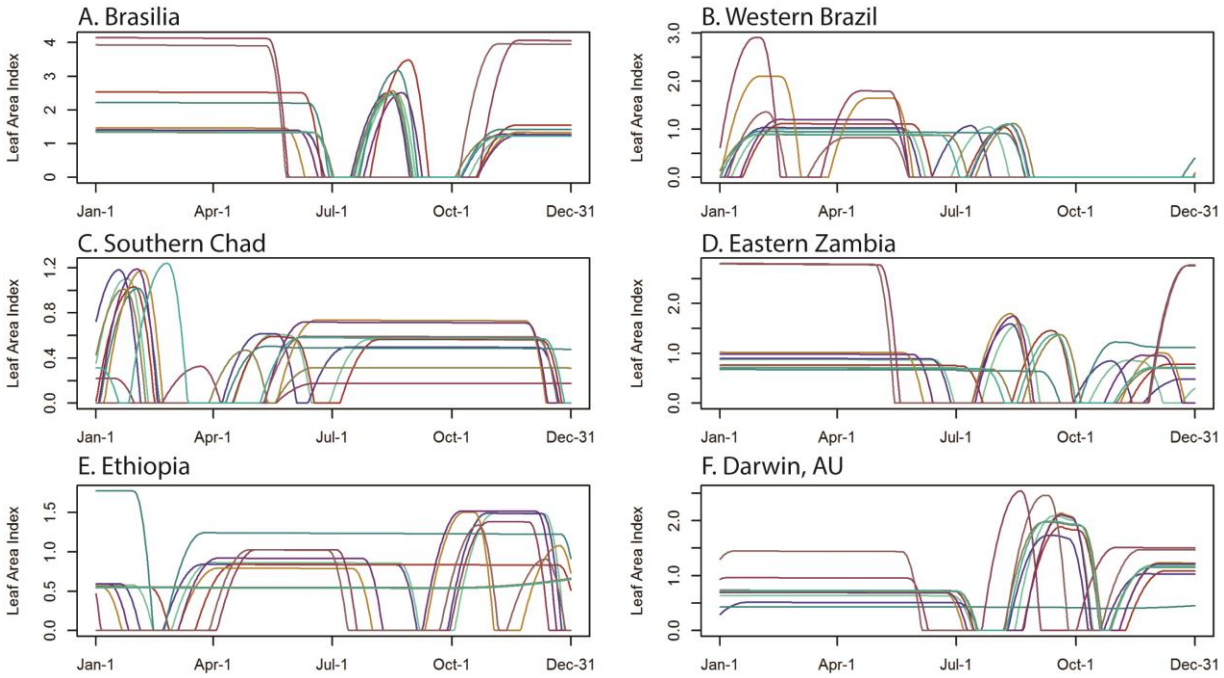

A. Brasilia
B. Western Brazil
C. Southern Chad
D. Eastern Zambia
E. Ethiopia
F. Darwin, AU
Leaf Area Index
Jan-1
Apr-1
Jul-1
Oct-1
Dec-31

Figure 4

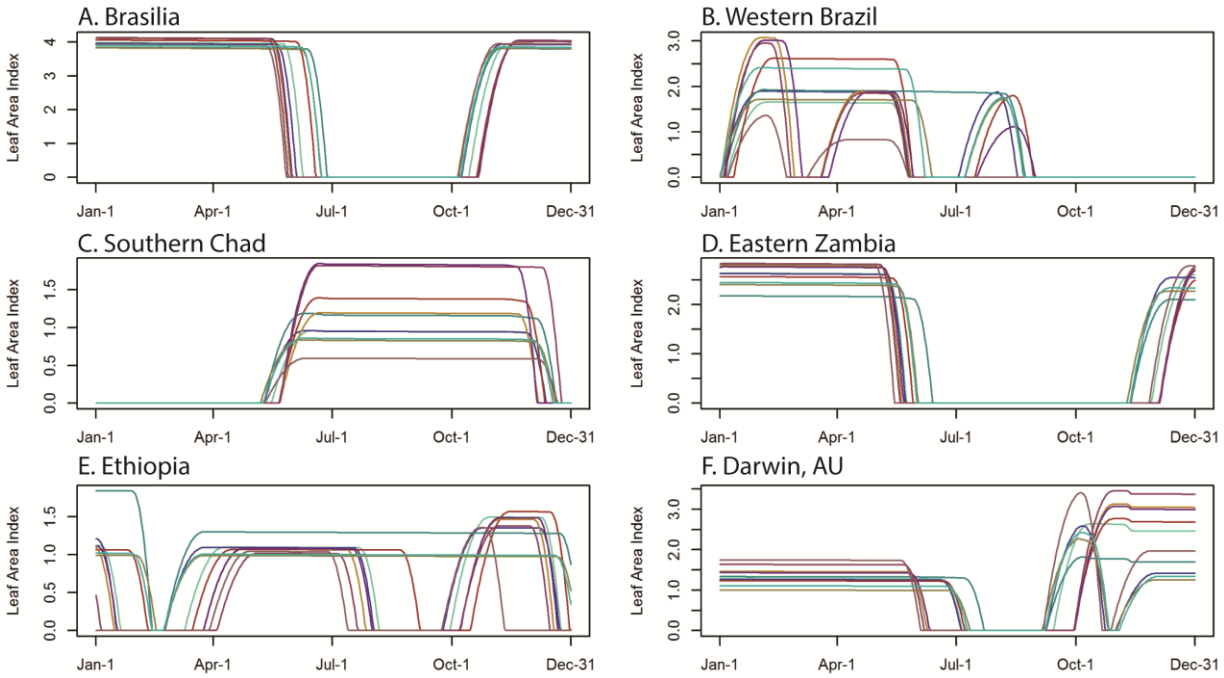

A. Brasilia
B. Western Brazil
C. Southern Chad
D. Eastern Zambia
E. Ethiopia
F. Darwin, AU
Leaf Area Index
Jan-1
Apr-1
Jul-1
Oct-1
Dec-31

Figure 5

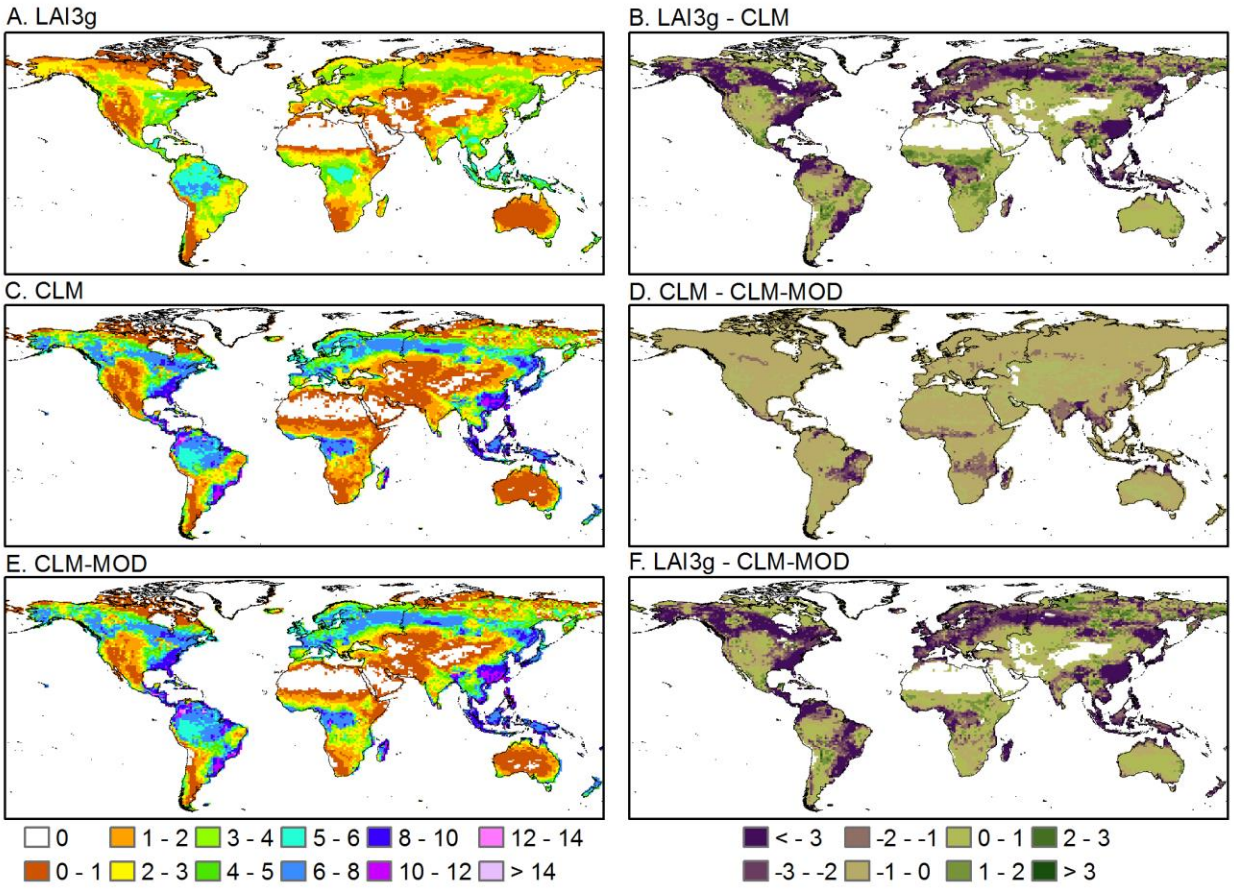

A. LAI3g
B. LAI3g - CLM
C. CLM
D. CLM - CLM-MOD
E. CLM-MOD
F. LAI3g - CLM-MOD
0
1 - 2
3 - 4
5 - 6
8 - 10
12 - 14
0 - 1
2 - 3
4 - 5
6 - 8
10 - 12
> 14
< - 3
-2 - -1
0 - 1
2 - 3
-3 - -2
-1 - 0
1 - 2
> 3

Figure 6

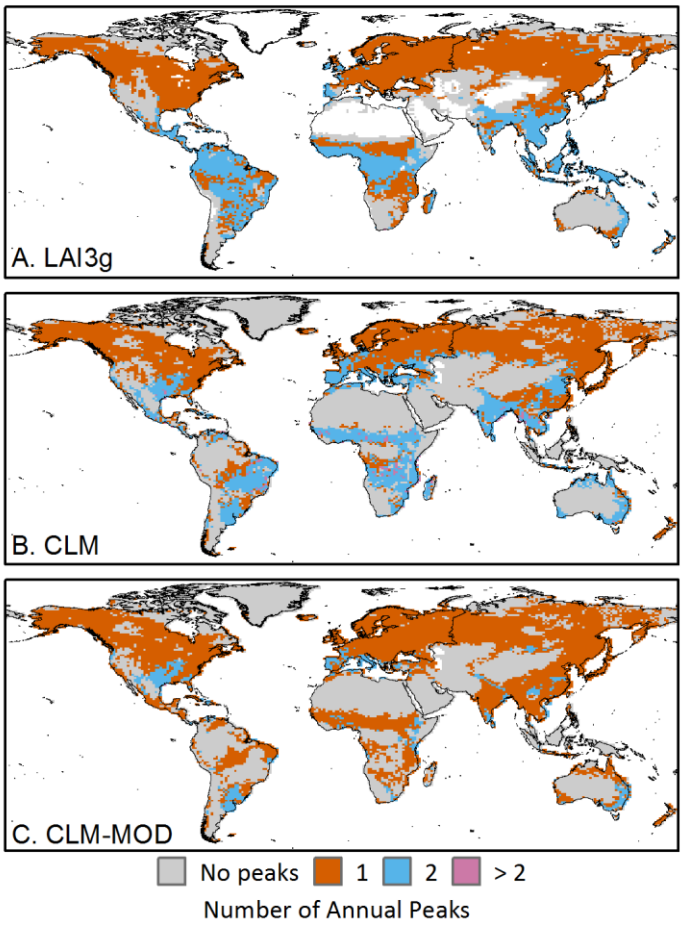


Figure 7

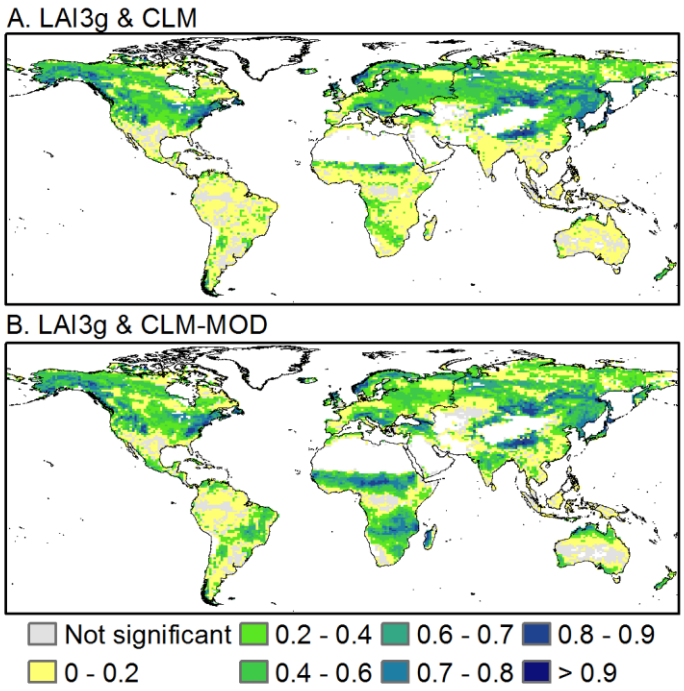

A. LAI3g & CLM
B. LAI3g & CLM-MOD
Not significant
0 - 0.2
0.2 - 0.4
0.4 - 0.6
0.6 - 0.7
0.7 - 0.8
0.8 - 0.9
> 0.9

Figure 8

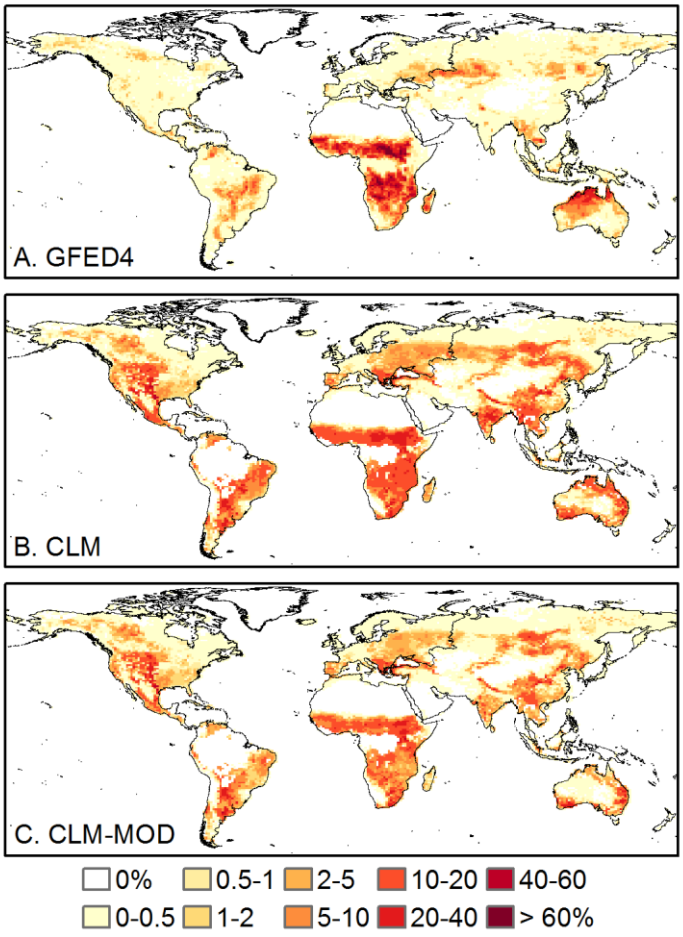


Figure 9

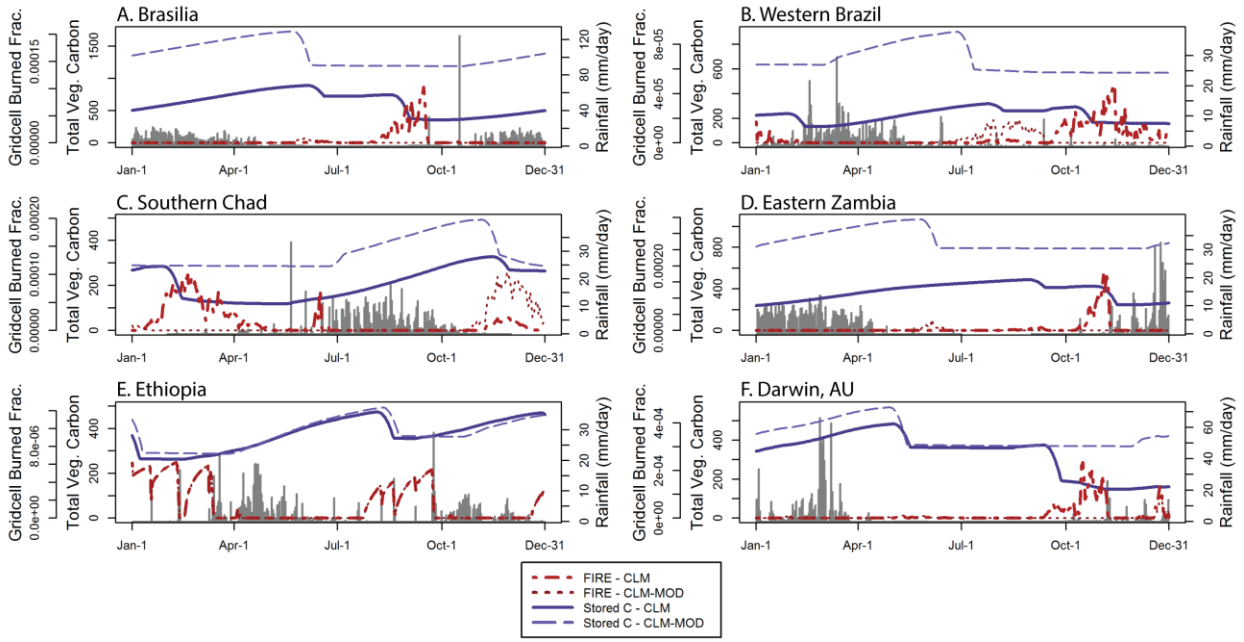

A. Brasilia
B. Western Brazil
C. Southern Chad
D. Eastern Zambia
E. Ethiopia
F. Darwin, AU
Gridcell Burned Frac.
Total Veg. Carbon
Rainfall (mm/day)
Jan-1
Apr-1
Jul-1
Oct-1
Dec-31
FIRE - CLM
FIRE - CLM-MOD
Stored C - CLM
Stored C - CLM-MOD

Figure 10