Phenology as a strategy for carbon optimality: a global model

S. Caldararu\textsuperscript{1,\,*}, D. W. Purves\textsuperscript{2}, and P. I. Palmer\textsuperscript{1}

\textsuperscript{1}School of GeoSciences, University of Edinburgh, Edinburgh, UK
\textsuperscript{2}Microsoft Research, Cambridge, UK
\* now at: Microsoft Research, Cambridge, UK

Received: 24 June 2013 – Accepted: 8 September 2013 – Published: 13 September 2013

Correspondence to: S. Caldararu (sicaldar@microsoft.com)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Phenology is essential to our understanding of biogeochemical cycles and the climate system. We develop a global mechanistic model of leaf phenology based on the hypothesis that phenology is a strategy for optimal carbon gain at the canopy level so that trees adjust leaf gains and losses in response to environmental factors such as light, temperature and soil moisture, to achieve maximum carbon assimilation. We fit this model to five years of satellite observations of leaf area index (LAI) using a Bayesian fitting algorithm. We show that our model is able to reproduce phenological patterns for all vegetation types and use it to explore variations in growing season length and the climate factors that limit leaf growth for different biomes. Phenology in wet tropical areas is limited by leaf age physiological constraints while at higher latitude leaf seasonality is limited by low temperature and light availability. Leaf growth in grassland regions is limited by water availability but often in combination with other factors. This model will advance the current understanding of phenology for ecosystem carbon models and our ability to predict future phenological behaviour.

1 Introduction

Leaf phenology refers to seasonal variations in leaf area, a direct constraint on carbon assimilation (White et al., 1999) and evapotranspiration (Wilson and Baldocchi, 2000), making it essential to understanding global and regional biogeochemical cycles. Phenological cycles are highly dependent on climate and the timing and spatial patterns of phenological dates may change significantly in response to changes in climate (Morin et al., 2009; Korner and Basler, 2010). As such, leaf phenology needs to be incorporated in global carbon and climate models, ideally in the form of simple equations based on biological processes, with predictive capabilities. Recent work has shown that existing phenology components of ecosystem models cannot correctly capture seasonal cycles as observed in flux tower measurements and that a better understanding of
phenology could improve current predictions of terrestrial systems (Richardson et al., 2012). Here we present a global, process based phenology model that aims to explain seasonal cycles as a function of environmental variables.

For cold deciduous vegetation, as occurs in temperate or boreal latitudes, the current understanding of spring phenology is that leaf budburst occurs after a given number of days with a temperature above a certain threshold (Kramer, 1994). Other potential factors include the photoperiod (day length requirement) and a chilling requirement necessary to prevent budburst after a warm period in winter (Chuine, 2000). Leaf senescence has been less intensely studied, but is believed to depend on either low temperatures or photoperiod (Hänninen et al., 1990; White et al., 1997; Delpierre et al., 2009), while some studies show that this is only dependent on day length and has a fixed date (Keskitalo et al., 2005). The effects of warming on leaf phenology are mostly considered to result in an early spring (e.g. Menzel et al., 2006; Thompson and Clark, 2008), but estimates of the potential magnitude of this change varies widely between studies (Korner and Basler, 2010). Also, the combination of the chilling effect and warming requirements can, in some species, lead to a late spring (Hanninen, 1990). Some studies have argued that, because of photoperiod constraints, this earlier budburst date cannot be proportional to spring warming as a very early date, even if warm, would not have the required daylength (Korner and Basler, 2010). Furthermore, an earlier budburst date is not necessarily directly related to an increase in overall productivity, as the seasonal response can be varied and associated changes in ecosystem respiration can lead to no net change, as shown by both measurement and modelling studies (Richardson et al., 2010; Parmentier et al., 2011).

The simplest method for describing the phenology component in climate and carbon models is to use prescribed budburst and senescence days (Sellers et al., 1986; Schaefer et al., 2008; Jain and Yang, 2005). Another method is to use satellite-derived vegetation data, which is well suited for large scale phenological studies because of its spatial and temporal coverage. Previous studies have used satellite vegetation indices, such as NDVI (normalised difference vegetation index) and EVI (enhanced vegetation
index) to determine budburst dates (Ju et al., 2006; Medvigy et al., 2009). Most of these studies use time-series techniques to determine onset and offset dates (Ludeke et al., 1996; Zhang et al., 2003). Any such approach using prescribed dates is incapable of projecting the potential impact of climate change on phenology. The most common approach to simulate climate change effects is to use a temperature dependency, often in the form of a growing degree day model, which uses the sum of days with temperatures above a given threshold, which is often fixed (White et al., 1997; Sitch et al., 2003; Krinner et al., 2005; Knorr et al., 2010). However, some models use a carbon efficiency approach to determine phenological cycles and patterns (Arora and Boer, 2005).

Most of the approaches used by global scale vegetation models are based on species level phenology, such as specific temperature thresholds, even though the spatial scales of such models can be quite coarse and one grid cell can include a mix of species and phenological types. The concept of landscape phenology was first introduced to refer to the phenological patterns captured by remote sensing data (Morisette et al., 2008; Schwartz, 1998) but it can also be applied to large scale modelling studies and a model that captures landscape rather than species level seasonality would be more appropriate for such models.

These model parametrisations often refer only to temperate deciduous forests, ignoring the large areas of dry and moist tropical forests that are often considered to lack a seasonal cycle (Cramer et al., 2001). Dry tropical forests and shrublands are generally thought to lose leaves during dry periods to prevent excessive water loss by plants, but leafing is often asynchronous between species and can occur during the dry season (Borchert, 1994; Reich and Borchert, 1982). The seasonally dry phenology is often represented through either prescribed dates for leaf out and leaf off (Schaefer et al., 2008) or as a threshold of available soil water, similar to the degree day approach (Knorr et al., 2010), but both these approaches cannot capture the more complex behaviour in these regions and cannot be used to predict future changes in phenology. Jolly et al. (2005) propose a model that uses empirical functions of tem-
temperature, daylength and water availability, to describe both temperate and dry tropical phenology, but does not include any seasonal cycle for the wet tropical forests.

In the case of moist tropical forests, studies have shown that these do have a weaker seasonal cycle, with a peak in the dry season (Myneni et al., 2007) due to an increase in solar radiation, especially in areas with deep-rooted trees and sufficient water (Nepstad et al., 1994). Caldararu et al. (2012) developed a mechanistic model of tropical leaf phenology for the Amazon basin and showed that these seasonal changes can be described as a response to variation in direct and diffuse radiation.

An alternative approach to threshold based phenology is that used by Kikuzawa (1996) to describe leaf habit based on the assumption of an optimal carbon assimilation strategy. This is independent of the environmental limiting factor or vegetation type, making it a more general approach (Schymanski et al., 2007). The assumption that plants are optimal and try to achieve maximum carbon gain has been previously tested both at the individual level (Ackerly, 1999; Le Roux et al., 2001) and at the ecosystem level (Niu et al., 2012).

In this paper we present a global process based phenological model, building on the tropical model of Caldararu et al. (2012), based on the hypothesis that phenology at a given location is a strategy for achieving an optimal carbon gain given the seasonal variation in light, temperature and water availability at that location. We fit this 14 parameter model globally at a resolution of 2° latitude by 2.5° longitude using leaf area index (LAI) data from the MODIS (Moderate Resolution Imaging Spectroradiometer) instrument (Sect. 2). We show that the model can be applied without any prior information about leaf habit (i.e. deciduous or evergreen) or biome and is able to explain and reproduce phenology at the landscape level in both temperate and tropical regions (Sect. 3). We then present the predicted LAI spatial and temporal patterns and use the fitted model to predict growing season metrics and the spatial distribution of factors which impact phenology across the globe (Sect. 4).
2 Data

2.1 MODIS LAI

We use leaf area index (LAI) data collection 5 from the MODIS (Moderate Resolution Imaging Spectroradiometer) Terra platform. The LAI/fPAR (fraction of absorbed photosynthetically active radiation) product is available at a 1 km spatial resolution (MOD15A) for the period 2000–present and at a temporal resolution of 8 days. The data is split into 1200 km by 1200 km tiles (10° latitude by 10° longitude at the equator). We use tiles for the entire globe for the chosen study (2001–2005) and evaluation periods (available at https://lpdaac.usgs.gov/).

The MODIS LAI retrievals are based on a reflectance algorithm (known as the main algorithm) which uses red and near infrared surface reflectance, illumination geometry and an eight-biome landcover map used to obtain information on vegetation structure and optical properties and soil optical properties (Knyazikhin et al., 1999). In cases where the main algorithm fails LAI values are calculated using an empirical relationship between NDVI and LAI (the back-up algorithm).

The quality flags associated with the LAI product provide information on the algorithm used, atmospheric conditions (cloud and aerosol presence) and snow cover. The data quality is affected mainly by cloud cover in tropical regions and snow at higher latitudes. Preliminary data analysis and ground validation studies (Cohen et al., 2006) have shown that values obtained using the back-up algorithm underestimate LAI, especially in high LAI regions such as the Amazon basin. Snow contaminated pixels also have low quality data. As such, we have eliminated all pixels that were derived using the back-up algorithm or were snow contaminated. We reproject all LAI data from its native sinusoidal projection to an orthogonal grid and spatially average to the GEOS 4 PAR data resolution (Sect. 2.2).

To spatially average the MODIS high resolution pixels, we need information on landcover type. The MODIS landcover product (MOD12Q1) provides 16 landcover classes under the IGBP classification. We have retained forest pixels classified as evergreen
(broadleaf and needleleaf), deciduous (broadleaf and needleleaf) and mixed and also open and closed shrublands and woody savannahs. Tropical forests are classed as > 90% evergreen, while mid-latitude forests are classed mostly as mixed, with no clear differences between temperate and boreal forests. We would expect a different leaf seasonality for boreal evergreen forests, with a lower seasonal amplitude, which is not reflected in the MODIS LAI data. This issue can be caused by poor snow detection in areas that are only partly snow covered (Klein et al., 1998; Beck et al., 2006). As such we aggregate all forest types into a mixed forest class. Since we do not have any previous information about the phenology of the different shrub landcover types, we also aggregate all three types into a mixed shrubland class. We need to differentiate between forest and shrubs within a phenology model as the two broad vegetation types generally have different rooting depth (Nepstad et al., 1994), which is important for describing soil water stress. This landcover information is only used for data preprocessing and the model does not require any further information the type of forest and its phenology type.

2.2 Environmental variables

The phenology model described in section 3 requires as inputs solar radiation, surface temperature and soil moisture. We use photosynthetically active radiation (PAR) data and surface temperature from assimilated meteorological data products of the Goddard Earth Observing System (GEOS-4) (Bey et al., 2001). The data is provided at a spatial resolution of 2° latitude × 2.5° longitude and a temporal resolution of 3 h, which we average to a one day resolution. To describe plant water availability within our model, we use volumetric soil moisture from the NCAR/NCEP reanalysis daily average surface flux data set (http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.surfaceflux.html) (Kalnay et al., 1996), which has a spatial resolution of 2.5° latitude × 2.5° longitude. This is currently one of the only available global datasets of soil moisture, as soil moisture is one of the most difficult variables to measure at large scales, together with most other soil variables, as it is determined by a combination of envi-
ronmental, above ground and below ground factors. This makes the reanalysis product difficult to validate at global scales and the few existing validation studies have proved inconclusive (Cheng-Hsuan et al., 2005).

Prior to model fitting we reproject all data onto the GEOS 4 orthogonal projection grid.

2.3 Ground based phenology data

We use ground based phenological measurements for model validation from the detailed record at Harvard Forest as well as measurements from a number of other flux sites (Table 1). The Harvard Forest data set (O’Keefe, 2000.) contains measurements of budburst, percent leaf development (percent leaves with 75 and 95% area developed), leaf colouring and senescence for all woody species at the site for the period 1990–2011. We use the mean percent leaf development at 75% and senescence over 2001–2006 for all species for comparison with our model results, which are LAI values across a larger area. Other phenological data used in this paper, obtained from the FLUXNET fair use data base, is less detailed and obtained through different types of measurements (Table 1).

3 Phenology model

We present a model of phenology based on the hypothesis that trees actively gain and lose leaves in order to achieve the maximum net carbon gain, that is, to achieve carbon optimality. We describe leaf gain and loss processes as a function of temperature, available light, soil moisture and leaf ageing. Figure 1 presents the model structure, described in detail below.
3.1 Leaf gain

We base the leaf gain mechanism on the tropical phenology model described in Caldararu et al. (2012). We assume that trees add leaves in order to achieve the optimal leaf area for light absorption, in response to available PAR so that in the absence of other constraints maximum LAI will occur at the time of peak solar radiation.

We define the target LAI, $LAI_{targ}$, as the optimal number of leaves that a tree will seek to achieve given a certain light level at the top of the canopy $I_0$ (Caldararu et al., 2012). This is calculated as:

$$LAI_{targ} = -\frac{1}{\alpha} \ln \left(\frac{C}{I_0}\right),$$  \hspace{1cm} (1)

where $\alpha$ is the attenuation coefficient and $C$ is a parameter representing the light compensation point, beyond which leaves are no longer able to photosynthesise. To calculate $I_0$ and the attenuation coefficient throughout the year, we account for variations in solar declination angles and extinction coefficients with both latitude and day of year (Brock, 1981). We calculate $I_0$ as the mean radiation over the previous $p$ days, where $p$ is a free parameter. Given this target value, trees will add new leaves if their LAI is lower than $LAI_{targ}$ at a given time $t$. The leaf production $P$ at location $x$ and time $t$ is then calculated as:

$$P(x, t) = \begin{cases} 
\text{gain}_{\text{max}}, & LAI_{targ}(x, t) - LAI(x, t - 1) > \text{gain}_{\text{max}} \\
LAI_{targ}(x, t) - LAI(x, t - 1), & 0 < LAI_{targ}(x, t) - LAI(x, t - 1) < \text{gain}_{\text{max}} \\
0, & LAI_{targ}(x, t) - LAI(x, t - 1) < 0 
\end{cases}$$  \hspace{1cm} (2)

Here the parameter $\text{gain}_{\text{max}}$ refers to the maximum leaf gain in a given time period and was introduced because trees have a limited leaf production rate.

To describe the role of temperature on phenology, important at higher latitudes, we include a temperature threshold of 0°C so that the conditions for leaf gain described above are only active if the average temperature over a number of $p$ days is above this limit.
3.2 Leaf loss

We assume that leaves are lost once the leaf becomes inefficient, that is, once the leaf assimilation is lower than its respiration and carbon maintenance cost. Depending on biome, the reason for a decrease in assimilation rate can either be a decrease in incoming solar radiation in winter (temperate regions), a decrease in water availability (seasonally dry regions) or, lacking any external constraints, simply leaf ageing (tropical regions).

In its simplest form we can describe carbon assimilation of a mature, unstressed leaf as a linear function of total absorbed PAR in the canopy, \( I_{\text{tot}} \), normalised by total LAI to obtain assimilation per unit leaf area:

\[
A_{\text{light}} = \frac{\phi I_{\text{tot}} - q}{\text{LAI}},
\]  

(3)

where \( \phi \) and \( q \) are parameters representing photosynthetic efficiency and overall canopy compensation point (the light level at which there is no photosynthesis anywhere in the canopy) respectively. At leaf level, carbon uptake saturates with incoming radiation and reaches a maximum value, \( A_{\text{max}} \). However, modelling studies have shown (Haxeltine and Prentice, 1996) that at the canopy level for time periods of one day or longer the relationship is linear due to both the distribution of nitrogen within the canopy and the differences in \( A_{\text{max}} \) and compensation points for leaves at different depths. As we are looking at large spatial scales over a sufficiently long time period (Sect. 2.1), we use the linear form. We calculate absorbed PAR, \( I_{\text{tot}} \) as a function of direct and diffuse PAR at the top of the canopy (\( I_0 \), see Sect. 3.1) and LAI, following the sun-shade model of dePury and Farquhar (1997).

As we do not use any prior information for the magnitude of carbon assimilation or the photosynthetic rate we normalise all assimilation values by setting the maximum assimilation rate \( A_{\text{max}} \) to one (unitless). For any values of \( I_{\text{tot}} \) that result in a rate greater than one, we set the assimilation to \( A_{\text{max}} \). 

15116
3.3 Water limitation

We know that, as soil water decreases, leaves are forced to partially or fully close their stomata, in order to avoid excessive water loss through transpiration, which leads to a lower carbon uptake. We define a water adjustment factor as:

\[ f_w = \frac{S - W_{\text{min}}}{W_{\text{max}} - W_{\text{min}}}, \quad (4) \]

where \( S \) is the water supply to the trees described as a function of soil moisture \( W_S \) (see below), \( W_{\text{max}} \) is the water level above which soil moisture has no impact on photosynthesis and \( f_w = 1 \) and \( W_{\text{min}} \) is the water level at which complete stomatal closure occurs and photosynthesis shuts down (\( f_w = 0 \)). For any water supply \( S \) greater than \( W_{\text{max}} \), \( f_w \) is set to 1, while for \( S \) values lower than \( W_{\text{min}} \), \( f_w \) is set to 0. Both \( W_{\text{min}} \) and \( W_{\text{max}} \) are dependent on the existing number of leaves, as shown below.

We express the water demand of a plant as the sum of the water used by the plant and the water lost through evapotranspiration and we assume that, under water stress conditions, trees adjust the number of leaves so that the water demand is equal to the soil water supply. The water available to the tree increases with soil moisture (Prentice et al., 1993), so that the supply, \( S \) is:

\[ S = s_1(W_S)^{s_2}, \quad (5) \]

given the two water extraction factors, \( s_1 \) and \( s_2 \).

We can express \( W_{\text{min}} \) and \( W_{\text{max}} \) (Eq. 4) in terms of water demand. \( W_{\text{min}} \), by definition, is the soil water level at which all stomata must be closed, so that there is no evapotranspiration and the water demand is equal to the water use, expressed as a function of the minimum water requirement per unit leaf area, \( u \):

\[ W_{\text{min}} = u \text{LAI} \quad (6) \]
$W_{\text{max}}$ is the soil water from which there is no water stress, so that no stomatal control is required and water demand is equal to water use plus the maximum evapotranspiration rate per unit leaf area, $\epsilon$:

$$W_{\text{max}} = uLAI + \epsilon LAI$$  \hfill (7)

Substituting these into Eq. (4), we calculate the water adjustment factor as a function of current LAI and soil moisture:

$$f_w = \frac{s_1(W_s)^{s_2}}{\epsilon LAI} - \frac{u}{\epsilon}$$  \hfill (8)

### 3.4 Leaf age effects

For each leaf age group we adjust the assimilation rate, as $A$ decreases with age. Following the leaf loss model for tropical regions (Caldararu et al., 2012), the age factor is for each cohort of age $a$:

$$f_a = \min(1, \exp^\mu(a_{\text{crit}} - a)),$$  \hfill (9)

where $\mu$ is the rate of decrease with age after a limit age $a_{\text{crit}}$. The total carbon assimilation, corrected for water and age effects is then

$$A_{\text{tot}} = A_{\text{light}} f_w f_a$$  \hfill (10)

Once this value reaches a threshold value $A_{\text{min}}$, the specific leaf cohort is lost. We then calculate leaf loss for each age cohort $\text{LAI}(x, t, a)$ as:

$$L(x, t, a) = \begin{cases} \text{LAI}(x, t, a), & A_{\text{tot}}(t, a) < A_{\text{min}} \\ 0, & A_{\text{tot}}(t, a) \geq A_{\text{min}} \end{cases}$$  \hfill (11)

The overall change in LAI at each time step $t$ and location $x$ is then:

$$\frac{d\text{LAI}(x, t)}{dt} = P(l_0(x, t), \text{LAI}(x, t - 1)) - \sum_{a=0}^{a_{\text{max}}} L(x, t, a),$$  \hfill (12)
where the production and loss processes are defined as in Eqs. (2) and (11). We fit the resulting 14 parameter model (Tables 2 and 3) to 5 yr of MODIS LAI data (2001–2005) using Filzbach, a Bayesian fitting algorithm (Caldararu et al., 2012).

3.5 Environmental limitation criteria

The fitted model parameters allow us to identify regions with a common limiting factor, i.e. light, soil moisture or leaf ageing using the three different triggers for leaf loss as follows. Light and temperature limited regions are regions where the light response assimilation $A_{\text{light}}$ is lower than the assimilation limit $A_{\text{min}}$, while in water limited areas $A_{\text{light}}f_w$ is lower than the limit. Age limited areas are regions where only $A_{\text{light}}f_wf_{\text{age}}$ is lower than the threshold. In practice, some regions will show a combination of these three limitations, in particular regions on the edge of wet tropical forests. We calculate the relative importance of these three factors by comparing the number of days in a year when any of the three conditions described above are true. For the purpose of assessing model performance for these different regions, we define a temperature and light, water and age regime as pixels where the relative importance is over 50% for temperature and light, water and age respectively.

4 Results and discussion

4.1 Model evaluation

Figure 2 shows that the model LAI generally agrees with the MODIS data, estimating the highest values of mean LAI in the tropical regions ($3.4 \pm 0.6 \text{ m}^2 \text{ m}^{-2}$) with values of up to $5.4 \text{ m}^2 \text{ m}^{-2}$ in the central Amazon and Central Africa. Temperate regions have a lower mean LAI ($1.2 \pm 0.3 \text{ m}^2 \text{ m}^{-2}$) with higher values in the deciduous Eastern US and Europe. The temperate regions exhibit a higher seasonal amplitude of $1.9 \text{ m}^2 \text{ m}^{-2}$ ($\pm 0.4 \text{ m}^2 \text{ m}^{-2}$), equivalent to 90% of the maximum LAI while tropical forests have a
much lower amplitude of $0.9 \pm 0.3 \, \text{m}^2 \, \text{m}^{-2}$, representing only 30% of the maximum, as expected for evergreen tropical regions.

We evaluate our results by running the model for 2006, a year that has not been used in the model fitting, to assess the predictive capability of the model. Figure 3 shows LAI timeseries for both the fitting and evaluation period for four major vegetation types. The model captures both the magnitude and seasonality of LAI at all four locations, with no major errors for the evaluation period.

Figure 4 shows the overall model error, expressed as root mean squared error (RMSE) relative to the mean LAI for both the fitting and evaluation periods. For the fitting period (Fig. 4a) the error is around 0.18, with higher values of up to 0.3 for shrubland regions. For the evaluation run, the RMSE is on average 0.25, slightly higher than that for the fitting period and follows the same spatial pattern, with higher values for shrublands. A direct comparison of observed and predicted mean and amplitude for the fitting and evaluation period (Figs. 6–7) shows that the model correctly captures the seasonal mean for all three different limitation regimes ($R^2 = 0.99$). The model LAI reproduces the observed amplitude for regions which are temperature and light limited ($R^2 = 0.99$) and a less good agreement over water and age limited regions ($R^2 = 0.67$ and 0.71 respectively). Figure 5 shows a map of the relative difference between model and observed mean and amplitude. Our model underestimates the mean LAI value mostly at higher latitudes by 25% (compared to only 2% in the tropics). This bias increases with latitude, from 8.8% at 30°N up to 37.5% at 54°N, but these values represent at most 0.5 m² m⁻². This issue is related to the extremely low values observed by MODIS in the boreal regions, which the model cannot correct. The evaluation run shows similar differences. The model tends to underestimate the seasonal amplitude in tropical regions by about 0.1 for both the fitting and evaluation period.

Our model is a process-based mechanistic model and has the advantage that, in addition to capturing the observed seasonal cycle, can be used to explore other aspects of leaf phenology such as the environmental limiting factors for leaf growth across the globe.
4.2 Growing season

The length of the growing season is a valuable indicator of changes in phenology in response to climate. The definition of the growing season varies with the type of data used, which makes ground validation of phenology models particularly difficult. For direct observations of budburst, the start of the growing season can be defined as the date of first budburst or percent of open buds and refers to single leaves. Canopy level studies report the canopy development level, often through measurements of radiation. When using satellite data, the start of the growing season, at landscape scale has previously been defined using a threshold for vegetation indices or as the inflection point based on a fitted curve (White et al., 1997). In studies which use flux tower data, the growing season is often defined as the period in which GPP is higher than respiration rates, once again at landscape level (Richardson et al., 2010), which is especially useful for evergreen boreal forests, which maintain leaf cover all year round.

Figure 8 shows the seasonal changes in canopy cover (percent canopy present of maximum LAI) for the model and MODIS data together with equivalent ground observations at the Harvard Forest site. The discrepancy between the model prediction and observed phenology at the start of the growing season, this decreases rapidly with progressivley better agreement further in the season. Furthermore, in autumn, where the MODIS data shows a sharper decline than the model prediction, the ground based observations show a similar pattern to our model, indicating a certain ability of the model to correct for errors in satellite data in this particular case. Figure 9 shows mean seasonal cycles of model and MODIS LAI together with the start and end of growing season predicted by the model, data and ground based observations at four flux sites. The observed start date of the growing season is generally earlier than the model by up to 8 days, while the MODIS start date is again later than the model by approximately 8 days. The observed difference can be caused by the different definitions of the growing season, as noted above. Whilst an actual observation of canopy development, as available at Harvard Forest, agrees well with our model, a start and end of season
observation does not offer sufficient information to fully assess the fit. The observed discrepancy in start date is also consistent with the observed behaviour at the Harvard forest site at the start of the season. While such differences can cause errors in our prediction of primary production, as noted by Richardson et al. (2012), the amount of carbon assimilated at the start of the growing season is very low compared to that assimilated at the peak of the season, where our model agrees well with both the MODIS and ground based data.

One further complication of large-scale models such as ours and point observations of vegetation is landscape heterogeneity. Previous studies have introduced the concept of landscape phenology (Morisette et al., 2008) or green wave phenology (Schwartz, 1998). The measured satellite LAI (or vegetation index) represents the vegetation behaviour for the entire grid cell, including all species both in the understory and overstory, often averaging across multiple vegetation types within the same biome. The general phenological behaviour at landscape scales is that species in the understory either leaf out early or keep leaves later in the autumn as an adaptation to their low light environment, as this maximises the amount of absorbed solar radiation, in the absence of leaves in the overstory (Richardson and O’Keefe, 2009). A similar pattern is shown by seedlings compared to adult trees (Seiwa, 2000). This would lead us to expect that the start date of the growing season predicted by our model is on average earlier than that observed in budburst dates for single species. We believe this is the main reason for the earlier start date, along with other early leafing species. The differences between landscape and single species phenology is even more pronounced in areas that include both deciduous and evergreen species, such as the high latitude boreal forests or high altitude mountain regions. Ground measurements in evergreen forests in the area show a higher LAI and less pronounced seasonal cycle than that observed in the MODIS LAI. This is the behaviour most commonly associated with coniferous evergreen species. However, the satellite observations also include the deciduous component, resulting in a seasonal cycle more similar to that of temperate forests. A similar problem is encountered in areas that include a mosaic of grasses and forests which result in a lower
LAI than expected if the pixel is classified as forest, or a higher LAI if it is classified as grassland. These observed differences are due to measurements at a different spatial scale. Furthermore, carbon cycle models are often on large spatial scales which would make observations, and predictions of landscape phenology a suitable input.

The model tends to predict the timing of 50% canopy development 16 days earlier than the observed MODIS LAI (equivalent to 2 timesteps), while the date of 75% development has an error of only 8 days (1 timestep). At the end of the growing season, the discrepancy is 16 days later for 50% canopy lost and 24 days for 75% lost. These thresholds are not necessarily an indication of the overall shape of the seasonal cycle, as can be seen in Figure 9 and the overall assessment of the model fit must be done in conjunction with estimates of LAI mean and amplitude, as shown above. It must be noted that while for species level phenology such errors are very high, the time scales used for coarse resolution studies are often different, reflecting data availability and inter-species variation.

Figure 10 shows the average length of the growing season for the 5 yr of the study period. We estimate that in tropical areas the length of the season covers the whole year, indicating that the trees are active throughout the year, whilst at higher latitudes the growing season is on average 225 days, decreasing with latitude (Fig. 11b). Figure 11a shows the variation in start and end date of the growing season for both forests and shrubs. The start date for forest pixels varies from day 68 at 30° N to day 120 at 66° N. The shrubs follow the same trend with generally later start days. The end day of the growing season varies from day 285 to day 341.

4.3 Phenological limiting factors

Figure 12 shows environmental limitations to phenology calculated using the fitted model parameters (Sect. 3.5). The model shows that the Amazon basin and parts of central Africa and South-East Asia are limited only by leaf ageing, indicating that the vegetation in these areas is wet tropical forests, not limited by seasonal water availability as previously discussed in Caldararu et al. (2012). The drier sub-tropical areas
around these forests are limited by a combination of water availability and leaf ageing. Only 48% of pixels limited by water to any extent are dominantly limited by water, with the rest being dominantly limited by leaf age. The dominantly water limited regions are concentrated in eastern Africa, while regions in South America and South East Asia are mostly driven by leaf ageing. There is a widely held assumption that the phenology of such forests is solely limited by water seasonality, but field studies have shown that leafing often occurs during the dry season and differs between species (Borchert, 1994; Reich and Borchert, 1982), which points at a further limitation.

In contrast to the overwhelming effect of temperature assumed in most phenology models for the higher latitudes, according to our analysis vegetation in these regions is limited not only by temperature, but also light availability, and the deciduous forests in Europe and eastern US show some influences of leaf age, which agrees with field observations which show that autumn senescence has a fixed date (Keskitalo et al., 2005). This results agrees with the hypothesis of Korner and Basler (2010), who stipulate that the length of the growing season cannot increase indefinitely in response to higher temperatures because of other constraints such as daylength.

The posterior parameter values determine the observed limitations in the model (Fig. A1–A3). Depending on the phenological limitation, some of the parameters will have little or no impact on the seasonal cycle. For example, in a high-latitude region that experiences periods of cold, but no seasonal drought, we would expect the parameters pertaining to water stress to be poorly constrained. This issue is reflected in the confidence intervals for each parameter (Fig. A4). The two parameters used to calculate plant water extraction, $s_1$ and $s_2$, are a measure of how much soil water is available for use by plants and reflect both the soil structure and root depth (Fig. A2). The water demand and evapotranspiration parameters determine the extent to which carbon assimilation is affected by water availability. High water use implies a high sensitivity to available soil water, something that would lead to pronounced drought-driven phenology. However, we estimate that in drought-deciduous regions (e.g. the African savannah) plants are generally well adapted to low water conditions and exhibit a wa-
ter limited seasonal cycle not because of their high water sensitivity but because of the very low soil moisture during the dry season.

Figure S3b shows the average age of the oldest leaf at any one point in time. This average leaf age is not a parameter of the model, but emerges when the model is simulated in a given region. As expected, leaves in tropical areas, which are age limited, have longer leaf lifespans, while leaves in temperate regions never have leaves older than 1 yr and mostly younger than 6 months (approximately equal to the growing season). We find that the leaf lifespan is not identical to the limit age parameter $\text{age}_{\text{crit}}$, particularly in temperate forests, that are temperature and light limited, where the average $\text{age}_{\text{crit}}$ is 1.4 yr (Fig. A3a) but leaves always drop at the end of the favourable season, making the effective leaf age equal to the growing season.

The four water related parameters are less well constrained over regions that are not impacted by water stress, while the age related parameters are less well constrained within temperate regions as their effects are not observed. The diffuse compensation point is better constrained in tropical forests where the combined seasonal cycles of the direct and diffuse PAR drive changes in LAI (Caldararu et al., 2012). Further information on plant traits would be needed to fully constrain all parameters at all locations.

5 Concluding remarks

We have shown that the model presented here, based on a carbon optimality hypothesis, is able to reproduce and predict phenology at global scales, as well as identify the climatic factors limiting leaf growth. We anticipate that the model can be used to improve existing representations of phenology in earth system models, thus improving our estimates of the global carbon budget. As a process based model, it can also be used for predicting future phenological behaviour under climate change scenarios. One interesting question to be answered is how the existing phenological limitations will change with climate. However, given that parameters for non-limiting environmental variables (e.g. the soil moisture parameters in cold regions) are not well constrained,
as explained above, it becomes apparent that the model in its current form cannot capture such regime shifts but can only predict changes in phenology in response to shifts in climate that do not alter the limiting factors, such as an increase in cloudiness in wet tropical regions or increases in temperature for temperate forests. To be able to predict more dramatic regime shifts the model would need information on plant behaviour which is not reflected in the existing data (e.g. drought response of cold temperate forests). There are two main approaches to solving this problem: using ground based measurements of plant traits such as leaf thickness and leaf nitrogen content to inform parameter values or Bayesian statistical methods that combine information from grid cells in the same geographical region or of the same plant functional type to obtain more information on the unconstrained parameters. Both these approaches constitute directions for future work, which will lead to a better understanding of potential changes in phenology across the globe. In the meantime, the model as presented here represents only the beginnings of a more physiological approach to predicting future leaf phenology worldwide.

Appendix A

Posterior parameter values and uncertainties

Figures A1–A3 show global distributions of posterior parameter values and Fig. A4 shows 95% confidence intervals for these values.

Acknowledgements. This project was funded through a Microsoft Research studentship. We would like to thank the developers of the MODIS LAI product for providing the data that forms the basis for our work. Phenology data for the Harvard Forest site was obtained through the LTER database and we would like to thank the team who collected this data for making it available. This work used eddy covariance data acquired by the FLUXNET community and in particular by the following networks: AmeriFlux (US Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (DE-FG02-04ER63917 and DE-FG02-
References


Phenology as a strategy for carbon optimality

S. Caldararu et al.


**Table 1.** Field sites where phenological data was available for model validation. Obtained from the FLUXNET fair-use database.

<table>
<thead>
<tr>
<th>Site</th>
<th>Name</th>
<th>Coordinates</th>
<th>Type of measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>US-Bar</td>
<td>Barlett Experimental Forest</td>
<td>44.66–71.29°</td>
<td>50% canopy development from fPAR data</td>
</tr>
<tr>
<td>US-MMS</td>
<td>Morgan Monroe State Forest</td>
<td>39.32–86.41°</td>
<td>NA</td>
</tr>
<tr>
<td>US-Ton</td>
<td>Tonzi Ranch</td>
<td>38.43–120.97°</td>
<td>60–70% of plants show the same phenology</td>
</tr>
<tr>
<td>US-UMB</td>
<td>Univ. of Michigan Biological Station</td>
<td>45.56–84.71°</td>
<td>LAI Licor-2000 measurements</td>
</tr>
<tr>
<td>US-Wcr</td>
<td>Willow Creek</td>
<td>45.81–90.08°</td>
<td>Changes in light interception</td>
</tr>
</tbody>
</table>
Table 2. Model parameters for leaf gain processes.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{\text{direct}}$</td>
<td>W m$^{-2}$</td>
<td>Leaf level light compensation point for direct PAR</td>
</tr>
<tr>
<td>$C_{\text{diffuse}}$</td>
<td>W m$^{-2}$</td>
<td>Leaf level light compensation point for diffuse PAR</td>
</tr>
<tr>
<td>$\rho$</td>
<td>days</td>
<td>Lag in response to incoming light</td>
</tr>
<tr>
<td>$\text{gain}_{\text{max}}$</td>
<td>m$^2$ m$^{-2}$</td>
<td>Maximum gain</td>
</tr>
</tbody>
</table>
Table 3. Model parameters for leaf loss processes.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi )</td>
<td>( \mu \text{mol s}^{-1} \text{W}^{-1} )</td>
<td>Photosynthetic efficiency</td>
</tr>
<tr>
<td>( q )</td>
<td>( \mu \text{mol m}^{-2} \text{s}^{-1} )</td>
<td>Canopy level compensation point</td>
</tr>
<tr>
<td>( s_1 )</td>
<td>–</td>
<td>Plant water uptake parameter</td>
</tr>
<tr>
<td>( s_2 )</td>
<td>–</td>
<td>Plant water uptake parameter</td>
</tr>
<tr>
<td>( e )</td>
<td>mm</td>
<td>Evapotranspiration per unit leaf area</td>
</tr>
<tr>
<td>( u )</td>
<td>mm</td>
<td>Plant water use per unit leaf area</td>
</tr>
<tr>
<td>( a_{\text{crit}} )</td>
<td>yr</td>
<td>Age after which leaves start ageing</td>
</tr>
<tr>
<td>( \mu )</td>
<td>yr(^{-1})</td>
<td>Decay constant of photosynthesis with age</td>
</tr>
<tr>
<td>( A_{\text{min}} )</td>
<td>( \mu \text{mol m}^{-2} \text{s}^{-1} )</td>
<td>Assimilation rate equal to leaf maintenance costs</td>
</tr>
</tbody>
</table>
Phenology as a strategy for carbon optimality

S. Caldararu et al.
Fig. 2. (a) Mean observed MODIS (top) and predicted (bottom) LAI and (b) relative annual amplitude expressed as the seasonal amplitude normalised by the maximum LAI value. Values calculated for the period 2001–2005.
Fig. 3. Predicted (black) and observed (blue) LAI for (a) tropical wet forests (6° S 55° W), (b) tropical dry forests (14° S 20° E), (c) temperate deciduous (46° N 15° E) and (d) temperate evergreen (54° N 120° E). Gray shaded area represents 95% confidence intervals calculated from the parameter posterior distribution. The blue shaded area (2006) represents the model evaluation period.
Fig. 4. Root mean squared error (RMSE) of predicted LAI normalised by mean LAI for (a) the study period (2000–2005) and (b) the evaluation period (2006).
Fig. 5. Relative model bias in mean LAI (left) and seasonal amplitude (right) for the model training period 2001–2005 (top) and evaluation period 2006 (bottom).
Fig. 6. Comparison of observed and predicted mean LAI for the fitting (top, 2001–2005) and evaluation (bottom, 2006) periods for the three limiting phenological regimes (Sect. 3.5).
Fig. 7. As for Fig. 6 but for seasonal amplitude.
Fig. 8. Percent canopy development as shown by the model (black), MODIS data (blue) and ground observations (red dots) at the Harvard Forest site (US-Harv, −72.18°, 42.53°) for the period 2001–2006. Shaded grey areas represent 95% confidence intervals calculated from the parameter posterior distribution.
Fig. 9. Model and MODIS mean seasonal cycle. The start and end of the growing season are determined as the day on which the canopy reaches 50% development for model and MODIS; other metrics for ground based estimates (see Table 1). Error bars represent minimum and maximum yearly variation for the measurement period.
Fig. 10. Length of growing season. All values are calculated for an average LAI for the 2000–2005 study period.
Fig. 11. Variation in growing season (a) start date and end date and (b) length with latitude for forests and shrubs.
Fig. 12. Regions where leaf loss is driven by temperature and light availability (red), water (blue) or age (green) as predicted by the model.
Fig. A1. Posterior parameter values describing leaf gain and carbon assimilation: (a) direct compensation point, $C_{\text{direct}}$, (b) diffuse compensation point $C_{\text{diffuse}}$, (c) maximum gain $g_{\text{max}}$, (d) photosynthetic efficiency $\phi$, (e) canopy compensation point $q$ and (f) assimilation limit $A_{\text{min}}$. 
Fig. A2. Posterior parameter values describing soil water effects: (a) and (b) soil water extraction parameters $s_1$ and $s_2$, (c) water use $u$ and (d) evapotranspiration rate $\epsilon$. 
Fig. A3. (a) Posterior parameter value for leaf ageing limit $a_{\text{crit}}$ and (b) realised age of oldest leaves.
Fig. A4. Relative 95% confidence intervals for all parameters derived from the posterior distribution. For parameter descriptions see Tables 1 and 2 and Sect. 2 in the main text.