GEOCLIM: a global climatology of LAI, FAPAR, and FCOVER from VEGETATION observations for 1999-2010

Aleixandre Verger [1,2,3*], Frédéric Baret [3], Marie Weiss [3], Iolanda Filella [1,2], Josep Peñuelas [1,2]

[1] CREAF, Cerdanyola del Vallès 08193, Catalonia, Spain

[2] CSIC, Global Ecology Unit, Cerdanyola del Vallès 08193, Catalonia, Spain

[3] INRA UMR114 EMMAH, Domaine Saint-Paul, Avignon 84914, France

***** Corresponding author: E-Mail: verger@creaf.uab.cat; Tel.: + 34- 935-813-008; Fax: +34-935-814-151.

Abstract

Land-surface modelling would benefit significantly from improved characterisation of the seasonal variability of vegetation at a global scale. GEOCLIM, a global climatology of leaf area index (LAI), fraction of absorbed photosynthetically active radiation (FAPAR), both essential climate variables, and fraction of vegetation cover (FCOVER) is here derived from observations from the SPOT VEGETATION programme. Interannual average values from the GEOV1 Copernicus Global Land time series of biophysical products at 1-km resolution and 10-day frequency are computed for 1999 to 2010. GEOCLIM provides the baseline characteristics of the seasonal cycle of the annual vegetation phenology for each 1-km pixel on the globe. The associated standard deviation characterises the interannual variability. Temporal consistency and continuity is achieved by the accumulation of multi-year observations and the application of techniques for temporal smoothing and gap filling. Specific corrections are applied over cloudy tropical regions and high latitudes in the Northern Hemisphere where the low number of available observations compromises the reliability of estimates. Artefacts over evergreen broadleaf forests and areas of bare soil are corrected based on the expected limited seasonality. The GEOCLIM data set is demonstrated to be consistent, both spatially and temporally. GEOCLIM shows absolute differences lower than 0.5 compared with MODIS (GIMMS3g) climatology of LAI for more than 80% (90%) of land pixels, with higher discrepancies in tropical and boreal latitudes. ECOCLIMAP systematically produces higher LAI values. The phenological metric for the date of maximum foliar development derived from GEOCLIM is spatially consistent (correlation higher than 0.9) with those of MODIS, GIMMS3g, ECOCLIMAP and MCD12Q2 with average differences within 14 days at the global scale.

Keywords: Climatology; vegetation phenology; seasonal and interannual variability; biophysical variables; SPOT VEGETATION

1. Introduction

The state and dynamics of vegetation play key roles in the carbon cycle and global climate. A set of essential climatic variables was identified as both accessible from remote sensing observations and involved in key processes (GCOS 2010). Among those relating to land surfaces, the leaf area index (LAI) and the fraction of absorbed photosynthetically active radiation (FAPAR) can be derived from observations in the reflective solar domain. These biophysical variables of vegetation are crucial in several processes, including photosynthesis, respiration, and transpiration. LAI is defined as one half the total area of green elements per unit area of horizontal ground (Chen and Black 1992; GCOS 2010). It controls the exchanges of energy, water, and greenhouse gases between the land surface and the atmosphere. FAPAR is defined as the fraction of radiation absorbed by the canopy in the 400-700 nm spectral domain under specified conditions of illumination and is a main input in models of light-use efficiency (McCallum et al. 2009). The fraction of vegetation cover (FCOVER), defined as the fraction of the background covered by green vegetation as seen from the nadir, is also as a very pertinent variable that can be used in models of the surface-energy balance to separate the contribution of the soil from that of the canopy (Gutman and Ignatov 1998; Su et al. 2005).

LAI, FAPAR, and FCOVER are routinely estimated from sensors with medium resolution such as VEGETATION (Baret et al. 2013), Moderate Resolution Imaging Spectroradiometer (MODIS) (Myneni et al. 2002) and the Advanced Very High Resolution Radiometer (AVHRR) (Zhu et al. 2013). The European Copernicus Global Land Service delivers global LAI, FAPAR, and FCOVER products from SPOT VEGETATION data from 1999 to the present with a spatial sampling close to 1 km. The products, known as GEOV1 products, have benefitted from the development and validation of existing products (Baret et al. 2013). Camacho et al (2013) demonstrated that GEOV1 products were more accurate and precise than current products.

Some land-surface models (LSMs) for simulating terrestrial water and carbon cycles use the spatiotemporal variation of LAI or FAPAR, described by different lookup tables depending on the type of vegetation (Viterbo and Beljaars 1995). The availability of satellite data in the last two decades describing the state and evolution of vegetation has allowed a better integration of biophysical variables into LSMs. Previous studies have demonstrated the improved performance of LSMs due to a better characterisation of the seasonal and interannual variability of vegetation functioning provided by the assimilation of satellite data. In particular, data assimilation yields a more realistic parameterisation in phenological models and reduces the models' prediction errors to 21 and 15% for FAPAR and LAI, respectively (Stöckli et al. 2011). A number of studies have shown the potential of assimilating LAI observations to correct vegetation model states (Demarty et al. 2007; Gu et al. 2006) and the implications of introducing the observed seasonal (van den Hurk et al. 2003) and interannual (Guillevic et al. 2002) variability of LAI in the annual cycle of hydrological fluxes. Boussetta et al (2013) showed that the assimilation of a MODIS derived LAI monthly climatology, i.e. the interannual average of LAI time series (as opposed to a vegetation-dependent constant LAI), in a model of global numerical weather prediction improved the forecast of near-surface (screen-level) air temperature and relative humidity through its effect on evapotranspiration. Barbu et al (2014) more recently demonstrated the potential of the assimilation of GEOV1 LAI into an ISBA-A-gs land-surface model to improve the monitoring of droughts. A LAI climatology was also useful for the identification of anomalies and trends in global vegetation (Baret et al. 2012; Brandt et al. 2014; Verger et al. 2014b; Verger et al. 2013; Zhu et al. 2013). The climatology of the biophysical variables reveals the seasonality inherent to the land-cover type and improve land-cover classification (Verhegghen et al. 2014). A climatology gap filling can better cope with missing and noise-contaminated data than can standard temporal filters for most missing data or large gaps in a single annual time series of satellite data, which have a large impact on the accuracy of the phenological metrics extracted from the reconstructed time series (Guyon et al. 2011; Kandasamy et al. 2013; Verger et al. 2013). Extraction of phenological information is also sensitive to the temporal (Pouliot et al. 2011; Zhang et al. 2009) and spatial (Fisher and Mustard 2007; Kovalskyy et al. 2011) resolution of the satellite data. The climatology derived from time series of moderate spatial resolution sensors preserves the high temporal frequency mandatory for phenological studies (Guyon et al. 2011). Finally, the climatology information can make projections and improve the stability of near real time estimates (Jiang et al. 2010; Verger et al. 2014a).

Despite the significance of global phenology for earth system monitoring and modelling, there are few data sets that explicitly describe the annual vegetation cycle at global scale. Boussetta et al. (2013) derived a monthly LAI climatology from 2000-2008 MODIS observations to be used in a numerical weather prediction model as indicator of the leaf development stage. The ECOCLIMAP programme (Faroux et al. 2013; Masson et al. 2003) is a dual database at 1 km resolution that includes an ecosystem classification and a coherent set of land surface parameters (including LAI, FAPAR and FCOVER) that are primarily mandatory in meteorological modelling for soil/vegetation–atmosphere transfer schemes. Other studies focus on the time variation of vegetation indices to propose a global normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI) reference data set for land-surface phenology using 13 years of VEGETATION observations (Verhegghen et al. 2014) or to derive phenological metrics from VEGETATION (Guyon et al. 2011), MODIS (Ganguly et al. 2010; Zhang et al. 2003) or AVHRR (Atzberger et al. 2013) time series.

The aim of this study is to provide a global climatology of LAI, FAPAR, and FCOVER for describing the seasonal and interannual variability of the vegetation cycle at the global scale. The derived climatology, GEOCLIM, will take advantage of the improvements in accuracy and temporal consistency provided by GEOV1 over existing products (Camacho et al. 2013). We propose to build a current climatology using a limited set of recent annual time-series since climate and land cover are changing with time. The time series, however, are expected to be sufficiently long for depicting the baseline annual cycle of the vegetation and for encompassing anomalies (Verhegghen et al. 2014). The climatology should simulate the data as closely as possible, i.e. it does not use existing land-cover maps or a model to describe the seasonal dynamics, thereby preventing the introduction of possible artefacts due to the lack of realism of the model used. GEOV1 time series from 1999 to 2010, corresponding to 12 years of estimates of biophysical variables at a spatial resolution of 1 km and a frequency of 10 days are used. The climatology is computed for each pixel as the average for a given date in a year across all years of the time series. The associated standard deviation characterises the interannual variability.

We will first describe the methodology and the data sets used to produce GEOCLIM. We will then evaluate GEOCLIM based on its main spatiotemporal features and its performance relative to other climatology data sets derived from AVHRR and MODIS data, with special emphasis on seasonality and the derived phenology.

2. GEOCLIM implementation

The generation of GEOCLIM was achieved from GEOV1 time series for the period 1999-2010 based on the interannual means of biophysical variables and the application of specific corrections. The input data set and the steps required to produce GEOCLIM are described hereafter.

2.1. GEOV1 biophysical products

GEOV1 biophysical products provide global coverage of LAI, FAPAR, and FCOVER from 1998/12/24 to the present at a ground sampling distance of 1/112° (approximately 1 km at the equator) and 10-day steps. A neural-network machine-learning algorithm was used to estimate GEOV1 products (Baret et al. 2013). Directionally normalised VEGETATION reflectances (Roujean et al. 1992) from the top of the canopy in the red, near-infrared, and short-wave infrared bands derived from the CYCLOPES processing line (Baret et al. 2007) were used as inputs to the neural networks. Based on the validation results for the available biophysical products (Garrigues et al. 2008; Weiss et al. 2007), the MODIS and CYCLOPES products were selected for the training process. The selected products were combined after re-projection onto the VEGETATION Plate-Carrée 1/112° grid, smoothed over time, interpolated at the 10-day frequency, combined, and eventually re-scaled to better fit the expected range of variation. Further details for the training of the neural networks and the generation of the product are provided in Baret et al (2013). Recent validation studies indicated that GEOV1 outperformed existing products in both accuracy and precision (Camacho et al. 2013). GEOV1 products are freely available at www1.

2.2. Climatology generation

For the generation of GEOCLIM only the GEOV1 biophysical products with the best quality were selected according to the quality flags on snow, aerosol, reflectance input and biophysical output status (Baret et al. 2010). The climatology is defined as the interannual mean of the best quality

GEOV1 products accumulated for 1999-2010. It is generated at the pixel scale (1-km spatial resolution) and at a dekadal temporal step (a 10-day period, with 36 dekads per year) within a 30-day compositing window (±15 days). The average values are then computed from three adjacent dekads instead of only one dekad along the 12-year period, which allows an increase in the number of points (12*3=36 compared to only 12), provides more robust and continuous estimates, and induces fewer artefacts because the dynamics of the products are approximately linear between the three dekads (Baret et al. 2010), as shown by Camacho et al (2013) for the smoothness of the GEOV1 products. A temporal smoothing and gap-filling (TSGF) technique (Verger et al. 2011) was applied to correct the residual artefacts, especially when the GEOV1 products were systematically unavailable across the years due to cloud coverage, and to ensure continuity and consistency in GEOCLIM as phenological studies request. Gap filling was achieved by linear interpolation. Temporal smoothing relied on an adaptive Savitzky-Golay second-degree polynomial fitting by processing three valid values on either side of the date (Verger et al. 2011). The compositing window may be asymmetric due to possible missing data. TSGF technique was demonstrated to improve other existing temporal filters for reconstruction satellite LAI time series in terms of the accuracy as compared to the original data by ensuring robustness to noise and missing data, while preventing over-smoothing (Kandasamy et al. 2013; Verger et al. 2011; Verger et al. 2013). An example of a climatology computation is illustrated in Figure 1, and TSGF correction is illustrated in Figure 2a.

[Figure 1]

[Figure 2]

2.3. Correction of specific artefacts

The generated climatology was then corrected for specific problematic behaviours based on available expert knowledge:

- Some artefacts were observed at northern high latitudes during the winter: anomalous seasonality and unexpected increases in LAI (FAPAR, FCOVER) (Figure 2b) with an artificial maximum peak in winter (Figure 2b) and high interannual variability resulting in high standard deviations (Figure 3a). These artefacts were mainly due to snow cover or very poor conditions of illumination that limited the number of valid observations and the reliability of the bidirectional reflectance model applied for the correction of VEGETATION data (Roujean et al. 1992) (Figures 2b, 3b). The LAI (FAPAR, FCOVER) values are expected to be relatively stable and low due to the low temperatures, short days, and low illumination during winter at these high latitudes. To correct these artefacts at northern latitudes, the GEOCLIM inputs higher than the 20th percentile during winter (defined here as the period for which the ***sun zenith angle, SZA>70º*** at the time of VEGETATION overpass, i.e. around 10:30) were fixed at the minimum pixel values observed over the entire period. We used the minimum values by preferentially selecting the values computed from at least three valid observations because the quality of GEOCLIM inputs is highly correlated with the number of valid observations available for their composition (Figures 2b, 3). We used the minimum value computed over all dekads if none of the dekads verified this condition. The areas where this specific correction was applied are shown in Figure 4a. Similar approaches based on representative winter values and thresholds to fill gaps and correct values affected by snow or poor illumination at high latitudes were also considered by Beck et al. (2006); Delbart et al. (2005); Zhang et al. (2004).
- Significant artefacts were also detected at equatorial and tropical latitudes due to aerosol-cloud contamination that produced high instabilities, artificial seasonalities, and missing data in the GEOV1 products and, consequently, in the derived output (Figure 2c). The high standard deviations (Figure 3a) and the low number of available observations (Figures 2c,

3b) appeared to be good indicators of the high uncertainty associated with the computed output over these tropical areas. Most of these cases corresponded to ***evergreen broadleaf forests*** (EBFs) (cf. Figures 3, 4b), so a minimum seasonality and high LAI values should be observed. We thus identified a pixel as an EBF if the 90th percentile (P90) of the LAI output was >4.5 and the 20th percentile was >P90-1.5. This method for the detection of EBFs based only on GEOV1 products (Figure 4a) agreed well with the GLOBCOVER land-cover map (Defourny et al. 2009) (Figure 4b). For EBFs, the GEOCLIM values were fixed to the 90th percentile computed over the entire period (Figure 2c).

- Some artefacts were also detected in the raw output for ***areas of bare soil*** (BS) where the observed seasonality was of the same order of magnitude of the precision of the GEOV1 product (Figure 2d). A pixel was identified as BS if the 90th percentile of the LAI output was <0.05 (compare Figure 4a to the land-cover map in Figure 4b). For BS, the GEOCLIM values were fixed to the 50th percentile computed over the entire period (Figure 2d).

[Figure 3]

[Figure 4]

3. Evaluation of GEOCLIM

This section assesses the performance of GEOCLIM. The validation data sets are first described. The spatial and temporal consistencies of GEOCLIM are discussed and evaluated by comparison with climatologies derived from both AVHRR and MODIS data. For the sake of brevity, the results focus on LAI, because of the three variables LAI, FAPAR, and FCOVER, LAI is used most by the scientific community. The comparison was performed at 0.5º spatial sampling on a Plate-Carrée grid. The 0.5º spatial resolution corresponds to the typical resolution of global models and reduces computation time. For comparison purposes, the different LAI data sets were averaged at monthly time step because one-month corresponds to the lowest temporal sampling among the validation datasets.

The global phenology derived from the temporal seasonality of GEOCLIM is also investigated in this section. For simplicity, we focus on the date of maximum foliar development, i.e. the timing of the peak of the growing season in the LAI annual cycle (Brown et al. 2012; Jönsson and Eklundh 2002). In addition to the phenological metrics derived from the LAI climatologies, we use also the MCD12Q2 MODIS phenological product (Zhang et al. 2003).

3.1. Validation data sets

3.1.1. ECOCLIMAP product

ECOCLIMAP is a database at 1/112º resolution on a Plate-Carrée grid resolution that includes a classification of ecosystems and a consistent set of associated land-surface variables, including LAI, at 10-day temporal sampling (Faroux et al. 2013). We used the latest Open-ECOCLIMAP v1 version, available since June 2014 at www2. It combines the global database of the first version, ECOCLIMAP-I, and an upgraded version, ECOCLIMAP-II, for Europe. ECOCLIMAP-I contains 215 ecosystems obtained by combining existing land covers, climatic maps, and NDVI seasonal profiles from AVHRR data acquired between April 1992 and March 1993 (Masson et al. 2003). For each class of vegetation, the maximum and minimum LAI values are fixed based on in-situ knowledge, and the annual cycle of LAI is constrained by the NDVI AVHRR temporal profiles using a linear relationship between NDVI and LAI. The second version, ECOCLIMAP-II, contains 573 ecosystems across Europe based on more recent land-cover maps, and the annual LAI profiles are derived from MODIS Collection 5 for the years 2002-2006 (Faroux et al. 2013).

The ECOCLIMAP LAI data set at the original 1/112º resolution was aggregated at 0.5º spatial sampling and averaged at monthly temporal sampling.

3.1.2. MODIS climatology

The MODIS Collection 5 Boston University (BU) LAI product at 0.25º latitude/longitude grid is the extracted best quality of standard MODIS LAI product based on MOD15A2 and MOD13A2 quality flags (Samanta et al. 2011). The standard MODIS LAI products relies on a biome dependent look-up table inversion of a radiative transfer model which ingests red and near infrared bidirectional reflectance factor values, their associated uncertainties, the view-illumination geometry, and biome type (within eight types based on MOD12Q1 land cover map). Further details on the retrieval algorithm are provided in Myneni et al. (2002); Yang et al. (2006b). Valid 1 km 8-day values are averaged to obtain monthly LAI (Samanta et al. 2011). The monthly LAI 1km sinusoidal product is aggregated and projected onto a 0.25º Plate-Carrée projection.

We derived a monthly MODIS climatology as the interannual average from 2000 to 2010 MODIS BU LAI product. The 0.25º product was aggregated to 0.5º spatial resolution for comparison purposes.

3.1.3. GIMMS3g climatology

The GIMMS3g LAI product derived from AVHRR data is available at 15-day temporal steps and 1/12º spatial resolution for the period July 1981 to December 2011. The principles used for the generation of this LAI data set are based on the use of neural networks which were trained first with GIMMS NDVI3g and MODIS LAI products for the overlapping period 2000-2009. The trained neural network algorithm is then applied using the land-cover class, the latitude and longitude coordinates, and the NDVI3g as the input data to generate the full temporal coverage of the GIMMS3g LAI data set. Further details on the algorithm for GIMMS3g retrieval can be found in Zhu et al. (2013).

We derived the GIMMS3g climatology as the interannual mean of GIMMS3g LAI time series at 15-day temporal step for the period 1999-2010. We aggregated the 1/12º products at 0.5º spatial resolution and averaged at monthly temporal sampling for comparison purposes.

3.1.4. MCD12Q2 product

MCD12Q2 (Collection 5) (Ganguly et al. 2010; Zhang et al. 2003) provides global yearly vegetation phenologies at 500 m from 2001 to 2010 MODIS time series. The MCD12Q2 algorithm uses a series of piecewise logistic functions fitted over the annual cycle of EVI data (Zhang et al. 2003). Among the transition dates provided by the MCD12Q2 product, the parameter "onset of greenness maximum" is used here for comparisons with the parameter "peak of growing season" derived from the LAI climatologies , i.e. the date for which the climatology reach its maximum value in the LAI annual cycle. The onset of maximal greenness conceptually corresponds to the transition date at which the annual cycle of the vegetation reaches maturity. This date is thus expected to be earlier than the date of maximum vegetation. Zhang et al (2006) compared the MCD12Q2 parameter to *in-situ* measurements and found that it corresponded to the time at which 85-90% of the individual leaves reached their final size.

The MCD12Q2 500 m sinusoidal product was projected onto a 0.5º Plate-Carrée projection using the MODIS re-projection tool (www3). Yearly MCD12Q2 values from 2001 to 2010 were then averaged to provide a typical phenology for comparison with the phenological metrics derived from the LAI climatology data sets.

3.2. Spatiotemporal consistency

The GEOCLIM biophysical variables had highly consistent spatial and temporal patterns (Figure 5a), in agreement with the global distributions of biomes (Figure 4b). The seasonal patterns of GEOCLIM LAI also reflected the expected regimes of vegetation at the global scale. Evergreen broadleaf forests exhibited null seasonality (Figure 5d) in the tropical belt where LAI was near 5 throughout the year (Figure 5b). Deserts also had no seasonality (Figure 5d) where LAI was near zero (Figure 5b). These results were expected given the forcing applied for evergreen broadleaf forests and bare soils (cf. Section 2.3). As expected, deciduous broadleaf forests and crops had the highest seasonalities (Figure 5d). The observed seasonality in needleleaf forests (Figure 5d) with a LAI ≤4 (Figure 5a) and means near 2 (Figure 5b) agreed with the observed seasonality of the understory layer, which can reach a LAI of ~2 or more in summer but which is often near zero in winter (Chen et al. 1997; Jiao et al. 2014; Masson et al. 2003).

The areas with the highest interannual variabilities in GEOCLIM (Figure 5c) corresponded to cropland in the USA and Eurasia, with intrinsic variabilities due to crop rotation or management, but also regions of severe drought and fire in South America, Africa, and Asia, regions of land-cover change such as the deforestation in Amazonian and Indonesian tropical forests, and regions of extreme events such as drought and heat waves in Europe, eastern China, and Australia. High interannual variability, however, may also indicate a problem with the computed GEOCLIM value due to instabilities in the GEOV1 data or to insufficient available GEOV1 data, as observed in the Gulf of Guinea (compare Figures 3b and 5c). In most regions, the interannual variability (Figure 5c) was significantly lower than the seasonal variability (Figure 5d), demonstrating that GEOCLIM provided a baseline vegetation annual cycle that was representative of the current phenology and that smoothed most of the anomalies.

[Figure 5]

3.3. Comparison with ECOCLIMAP, MODIS and GIMMS3g climatologies

The map of annual mean differences between GEOCLIM and the LAI climatologies derived from AVHRR and MODIS data (Figure 6) shows LAI differences of ±0.5 for 54%, 83% and 91% of the land pixels as compared with ECOCLIMAP, MODIS and GIMMS3g, respectively. GEOCLIM produced systematically lower values than ECOCLIMAP for the remaining 46% of pixels, with larger differences for dense forests (northern boreal and tropical forests) but with significant differences also for crops (e.g. USA and eastern Asia) (Figure 6a). These systematic negative bias of GEOCLIM as compared to ECOCLIMAP was evident across latitudes and along the year (Figure 7). GEOCLIM produced also systematically lower values than MODIS (Figure 6b) over tropical forests with differences ~0.5 along the year (Figure 7b) and over northern deciduous broadleaf forest during the maximum growing leaf development (lower frequencies for the maximum values in Figure 8c). On the contrary GEOCLIM produced slightly higher LAI values than GIMMS3g (Figure 6c) over Amazon and Indonesian evergreen broadleaf forests and over boreal needle leaf forests in Russia and USA during the winter time (Figure 7d).

[Figure 6]

Despite the large discrepancies in the magnitude of LAI between the different datasets, due in part to the differences in sensors and retrieval algorithms, seasonality and its phasing generally agreed well (Figure 7). Seasonality was inverted in the Southern Hemisphere relative to the Northern Hemisphere (compare Figures 7a and 7c). In the Northern Hemisphere, LAI seasonality decreased in the length of season (active growth period) with latitude (compare Figures 7c and 7d). In the tropical latitudes (-20-10º) characterized by very limited seasonality GEOCLIM and GIMMS3g systematically showed lower values than MODIS and ECOCLIMAP (Figure 7b). The largest differences in terms of seasonality were in the Northern Hemisphere at high latitudes (40-70º) where ECOCLIMAP produced longer growing seasons as compared to other LAI datasets and higher values in the period of active growth (Figure 7d). Nevertheless, all the data sets agreed well for the base level of LAI during the dormancy period for the 40-70º latitudes validating *a posteriori* the reliability of the specific correction applied in winter (SZA>70º) to the GEOCLIM values for high northern latitudes (Section 2.3).

[Figure 7]

Histograms of the LAIs (Figure 8) indicated very similar distributions between the different LAI datasets for shrubs/savannah/bare soil. Some similarities in the position of the maximum frequency were also observed for crops and grassland. Some discrepancies, however, were observed: ECOCLIMAP produced low frequencies for LAIs of zero at the expense of higher intermediate values, while GEOCLIM, MODIS and GIMMS3g produced a smoother transition. Forests had the largest discrepancies between the different data sets. GEOCLIM and GIMMS3g produced a bimodal distribution for deciduous broadleaf forests, with a peak for low values (LAI near 1) corresponding to the dormant period of the vegetation in winter and a second mode for the period of active growth with values higher than 6 in few occasions. MODIS produced also a peak for LAI near 1 and a smooth transition up to maximum values around 6.5. ECOCLIMAP produced an even distribution for deciduous forests, but with unrealistic peaks. Evergreen broadleaf forests had relatively consistent narrow distributions between GEOCLIM, ECOCLIMAP and MODIS but with significant differences in the magnitudes (i.e. similar shapes but shifted distributions). The LAI modes were 5 for GEOCLIM, and 6 for ECOCLIMAP and MODIS. GIMMS3g produced broader distributions with the LAI mode ~4. Needleleaf forests had similar distributions for GEOCLIM, MODIS and GIMMS3g but with higher frequencies for low values compared to those in ECOCLIMAP. The LAI mode around 1 for GEOCLIM (MODIS and GIMMS3g) for deciduous broadleaf forests (Figure 8c) and needleleaf forests (Figure 8e) corresponded to the winter LAI value and reproduced the expected seasonality in northern high latitudes (Figure 7d) while ECOCLIMAP produced unrealistic LAI distributions and peaks' locations (Figures 8c and 8e).

[Figure 8]

3.4. Assessment of global phenology

The spatial pattern of the phenology derived by GEOCLIM (Figure 9) reflected the distributions of climate and biome type (Figure 4b). Seasonality was strongly dependent on temperature at northern latitudes >30º, and the timing of maximum greenness had a clear latitudinal gradient indicating a delay in the date of peak development with latitude (Figure 10). In other regions, seasonality had more complex spatial patterns that were driven mostly by biome type, land use, and the seasonal variation in rainfall (Figure 9).

The phenological metrics (Figure 10) were spatially consistent in the timing of the maximum of the growing season as derived from the different data sets and particularly between GEOCLIM and MODIS (GIMMS3g) with uncertainties of around 14 days in terms of RMSE, bias of less than 1 day, a correlation higher than 0.95 and a slope of the linear regression close to the unity (Table 1). The phenology derived from ECOCLIMAP was also highly spatially consistent with GEOCLIM (correlation about 0.9, slope close to the unity and bias about 6 days, Table1) but it diverged to some degree (uncertainties of about one month in terms of RMSE), mostly in the Southern Hemisphere (Figure 10) in regions with a limited seasonality (Figure 5d). As expected (Section 3.1.4), the phenological phase for the "onset of greenness maximum" retrieved in MCD12Q2 occurred earlier than the peak date in GEOCLIM (bias of 14 days, Table 1) due to differences in the definitions of the phenological metrics. The phenological metrics derived from GEOCLIM constitutes an intermediate solution across latitudes between MODIS, GIMMS3g, ECOCLIMAP and MCD12Q2 for the date of maximum foliar development (Figure 10)

[Figure 9]

[Figure 10]

[Table 1]

4. Discussion

Twelve years (1999-2010) of data from GEOV1 LAI, FAPAR, and FCOVER products were used to compute GEOCLIM outputs for the interannual average seasonal cycle at a pixel scale. The main assumptions were that (i) no land-cover change or abrupt disturbance leading to a change in the phenological annual cycle occurred for the period considered and (ii) the time series were sufficiently long to reduce the sensitivities of the averages to anomalies. Specific correction were applied at northern high latitudes, bare soils and evergreen broadleaf forests to overcome problems associated, respectively, with strong bidirectional effects and snow cover, precision and signal to noise ratio, and aerosol-cloud contamination (Section 2.3). The identification of bare soil and evergreen broadleaf forests was completely driven by the data avoiding possible miss-classification errors introduced by external land cover map information though a good spatial consistency with GLOBCOVER map was observed. In these problematic cases, GEOCLIM was forced to fixed values derived from the input data at the pixel level under the following hypothesis: (i) minimum vegetation activity in winter time at northern latitudes, and no seasonality in (ii) desert areas and (iii) evergreen broadleaf forests where the vegetation is respectively low (LAI~0) and high (LAI~5) throughout the year. The last hypothesis constitutes an oversimplification of the reality because of the possible seasonality of evergreen broadleaf forests. The high uncertainty associated with the data due to poor atmospheric correction and very high cloud occurrence in equatorial and tropical latitudes prevented the extraction of meaningful phenology at the resolution of the individual pixels of 1 km. The high spatial and temporal resolution of forthcoming Sentinel2 sensors should improve the monitoring of vegetation in these problematic areas.

GEOCLIM was indirectly validated based on the comparison with AVHRR and MODIS derived climatologies of LAI. Multitemporal ground data would be preferable for validating GEOCLIM but unfortunately were rarely available. GEOCLIM showed a high agreement with MODIS (GIMMS3g) climatology of LAI and absolute differences were higher than the Global Climate

Observing System (GCOS 2010) requirements for accuracy, i.e. 0.5 LAI, only in northern boreal and tropical forests representing less than 20% (10%) of land pixels. GEOCLIM systematically produced lower values than MODIS over evergreen broadleaf forest as also observed in the comparison between GEOV1 and MODIS (Camacho et al. 2013; Fang et al. 2013). The difficult observational conditions in tropical latitudes with persistent clouds can cause irregularities in the solution and thus variable but systematic underestimations of LAI (Verger et al. 2011). The specific correction applied to GEOCLIM removed the instabilities in the solution but cannot correct possible biases in the magnitude of original GEOV1 products used as input data for GEOCLIM. Previous studies have also shown that GEOV1 products produce slightly higher values than MODIS for needleleaf forest in winter (Fang et al. 2013). The specific correction applied in GEOCLIM at northern high latitudes reduced these differences but may result in some underestimation of the seasonal amplitude in winter time. Accurate estimation of LAI in needleleaf forests in winter is challenging because contamination by clouds and snow limits the reliability of the reflectances used as inputs in the algorithms (Camacho et al. 2013). Further, the strong bidirectional effects of surface-reflectance at very high latitudes are not well simulated by the radiative transfer models currently used for product generation (Yang et al. 2006a). In addition, the understory and foliage clumping are not well accounted for (Jiao et al. 2014; Pisek et al. 2010).

LAI values were systematically higher for ECOCLIMAP than for GEOCLIM, MODIS and GIMMS3g. Boussetta et al (2013) reported similar higher LAI values for ECOCLIMAP than for MODIS. Garrigues et al (2008) also reported large positive biases for ECOCLIMAP compared with CYCLOPES, MODIS and GLOBCARBON and with ground measurements. The differences in the temporal period, input data and sensors (VEGETATION for GEOCLIM and AVHRR and MODIS for ECOCLIMAP, Section 3.1) can partially account for the significant discrepancies between GEOCLIM and ECOCLIMAP although the relatively good agreement of GEOCLIM with MODIS and GIMMS3g AVHRR derived LAI climatologies indicates that the major source of discrepancies are related to the retrieval algorithms. The linear relationship between NDVI and LAI used to retrieve the ECOCLIMAP product for pixels out of Europe (Masson et al. 2003) may have introduced some overestimation because the LAI-NDVI relationship is exponential and saturates at medium to high values (e.g. Myneni et al. (2002)). Since ECOCLIMAP assumes low spatial variability within each class of land cover, it is limited to capture the LAI spatial variability as compared to other LAI datasets (Garrigues et al. 2008). Nevertheless, identifying the source of the differences between ECOCLIMAP and the other LAI datasets being analyzed would require further attention and it is out of the scope of this paper.

Further research should focus on the development of improved LAI datasets with due attention to areas (boreal and tropical latitudes) and periods (winter time) where higher uncertainties exist (Fang et al. 2013). In these cases characterized by high level of noise and missing data, the use of the climatology and temporal smoothing and gap filling techniques applied at daily estimates of biophysical variables may increase the robustness of the solution as compared to the classical composition techniques (Verger et al. 2014a).

The phenological metrics derived from GEOCLIM was highly spatially consistent (correlation higher than 0.9) with MODIS and AVHRR derived phenologies, including ECOCLIMAP ones, for the date of maximum foliar development with differences lower than six days in all cases except when compared with MCD12Q2 product (systematic biases of 14 days) due to the differences in the definition. A standardization in the definitions of the phenological metrics appears necessary (White et al. 2009). Disentangling the mechanisms that govern the seasonal and interannual variability in phenology and vegetation-climate dynamics at the global scale would require further analysis.

5. Conclusions

This article has presented and provided a first quality assessment of GEOCLIM—a global climatology of LAI, FAPAR and FCOVER—from the multiannual time-series of 10-days and 1-km

GEOV1 products built from 1999 to 2010. Results showed GEOCLIM was temporally consistent for the seasonality across biomes and latitudes. GEOCLIM showed a high agreement with MODIS and AVHRR climatologies of LAI: differences within the Global Climate Observing System requirements, i.e. ±0.5 LAI, in more than 80% (90%) of GEOCLIM land pixels compared with MODIS (GIMMS3g). ECOCLIMAP systematically produced higher LAI values. Further research should focus on tropical and boreal latitudes where higher uncertainties exist in the LAI datasets. The phenology of the timing of maximum foliar development derived from GEOCLIM constituted an intermediate solution between those of GIMMS3g, MODIS, ECOCLIMAP and MCD12Q2: differences within 14 days and spatial correlation >0.9 at the global scale.

The GEOCLIM data set is continuous, both spatially and temporally, and it can be used for a wide range of land-biosphere applications. It may contribute to a better characterisation of the seasonal variability of vegetation in global land-surface models. It provides the baseline characteristics of the seasonal cycle of LAI, FAPAR, and FCOVER for the identification of anomalies and trends in global vegetation.

Acknowledgements

This research was partially supported by the European Earth observation programme Copernicus (388533), *Pôle Thématique Surfaces Continentales THEIA* (CNES 140570) and the FP7 geoland2 (218795), GIOBIO (32-566) and LONGLOVE (32-594) projects. This research was also supported by the Spanish Government grant CGL2013-48074-P, the Catalan Government grant SGR 2014-274, and the European Research Council Synergy grant ERC-2013-SyG-610028 IMBALANCE-P. Aleixandre Verger is the recipient of a *Juan de la Cierva* postdoctoral fellowship from the Spanish Ministry of Science and Innovation. The authors are indebted to Ranga Myneni, Taejin Park, the Boston University, MODIS, CNES, VITO, and CNRM-GAME (Météo France, CNRS) teams for providing access to the data.

References

Atzberger, C., Klisch, A., Mattiuzzi, M., & Vuolo, F. (2013). Phenological Metrics Derived over the European Continent from NDVI3g Data and MODIS Time Series. *Remote Sensing, 6,* 257-284

Barbu, A.L., Calvet, J.C., Mahfouf, J.F., & Lafont, S. (2014). Integrating ASCAT surface soil moisture and GEOV1 leaf area index into the SURFEX modelling platform: a land data assimilation application over France. *Hydrol. Earth Syst. Sci., 18*, 173-192

Baret , F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Weiss, M., Samain, O., Roujean, J.L., & Leroy, M. (2007). LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. Part 1: Principles of the algorithm. *Remote Sensing of Enviroment*, *110*, 275-286

Baret , F., Verger, A., Weiss, M., Makhmara, H., Lacaze, R., & Pacholczyk, P. (2012). Use of GEOV1 climatology based on 11 years of consistent observations: identification of trends at the global scale. In, *EGU General Assembly 2012*: Geophysical Research Abstracts, Vol. 14, EGU2012-7912

Baret , F., Weiss, M., & Kandasamy, S. (2010). Climatology of LAI, FAPAR and FCOVER products (V1). In (p. 30). Avignon (France): INRA-EMMAH

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholcyzk, P., & Smets, B. (2013). GEOV1: LAI, FAPAR Essential Climate Variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production. *Remote Sensing of Environment, 137*, 299-309

Beck, P.S.A., Atzberger, C., Hogda, K.A., Johansen, B., & Skidmore, A.K. (2006). Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sensing of Environment, 100*, 321

Boussetta, S., Balsamo, G., Beljaars, A., Kral, T., & Jarlan, L. (2013). Impact of a satellite-derived leaf area index monthly climatology in a global numerical weather prediction model. *International Journal of Remote Sensing, 34*, 3520-3542

Brandt, M., Verger, A., Diouf, A., Baret, F., & Samimi, C. (2014). Local Vegetation Trends in the Sahel of Mali and Senegal Using Long Time Series FAPAR Satellite Products and Field Measurement (1982–2010). *Remote Sensing, 6*, 2408-2434

Brown, M.E., de Beurs, K.M., & Marshall, M. (2012). Global phenological response to climate change in crop areas using satellite remote sensing of vegetation, humidity and temperature over 26 years. *Remote Sensing of Environment, 126*, 174-183

Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., & Weiss, M. (2013). GEOV1: LAI, FAPAR Essential Climate Variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products. *Remote Sensing of Environment, 137*, 310-329

Chen, J.M., & Black, T.A. (1992). Defining leaf area index for non-flat leaves. *Plant, Cell and Environment, 15*, 421-429

Chen, J.M., Rich, P., Gower, S.T., Norman, J.M., & Plummer, S. (1997). Leaf area index of boreal forests: theory, techniques and measurements. *Journal of Geophysical Research, 102*, 29429-29443

Defourny, P., Bicheron, P., Brockmann, C., Bontemps, S., Van Bogaert, E., Vancutsem, C., Pekel, J.F., Huc, M., Henry, C., Ranera, F., Achard, F., di Gregorio, A., Herold, M., Leroy, M., & Arino, O. (2009). The first 300 m global land cover map for 2005 using ENVISAT MERIS time series: a product of the GlobCover system,. In, *Proceedings of the 33rd International Symposium on Remote Sensing of Environment*. Stresa (Italy)

Delbart, N., Kergoat, L., Le Toan, T., Lhermitte, J., & Picard, G. (2005). Determination of phenological dates in boreal regions using normalized difference water index. *Remote Sensing of Environment, 97*, 26-38

Demarty, J., Chevallier, F., Friend, A.D., Viovy, N., Piao, S., & Ciais, P. (2007). Assimilation of global MODIS leaf area index retrievals within a terrestrial biosphere model. *Geophysical Research Letters, 34*

Fang, H., Jiang, C., Li, W., Wei, S., Baret, F., Chen, J.M., Garcia-Haro, J., Liang, S., Liu, R., Myneni, R.B., Pinty, B., Xiao, Z., & Zhu, Z. (2013). Characterization and intercomparison of global moderate resolution leaf area index (LAI) products: Analysis of climatologies and theoretical uncertainties. *Journal of Geophysical Research: Biogeosciences*, n/a-n/a

Faroux, S., Kaptué Tchuenté, A.T., Roujean, J.L., Masson, V., Martin, E., & Le Moigne, P. (2013). ECOCLIMAP-II/Europe: a twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models. *Geosci. Model Dev., 6*, 563-582

Fisher, J.I., & Mustard, J.F. (2007). Cross-scalar satellite phenology from ground, Landsat, and MODIS data. *Remote Sensing of Environment, 109*, 261-273

Ganguly, S., Friedl, M.A., Tan, B., Zhang, X., & Verma, M. (2010). Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product. *Remote Sensing of Environment, 114*, 1805

Garrigues, S., Lacaze, R., Baret, F., Morisette, J.T., Weiss, M., Nickeson, J.E., Fernandes, R., Plummer, S., Shabanov, N.V., Myneni, R.B., Knyazikhin, Y., & Yang, W. (2008). Validation and intercomparison of global Leaf Area Index products derived from remote sensing data. *Journal of Geophysical Research: Biogeosciences, 113,* G02028

GCOS (2010). Implementation Plan for the Global Observing System for Climate in support to UNFCCC (2010 Update). In (p. 186). Geneva, Switzerland: World Meteorological Organization

Gu, Y., Belair, S., Mahfouf, J.-F., & Deblonde, G. (2006). Optimal interpolation analysis of leaf area index using MODIS data. *Remote Sensing of Environment, 104*, 283

Guillevic, P., Koster, R.D., Suarez, M.J., Bounoua, L., Collatz, G.J., Los, S.O., & Mahanama, S.P.P. (2002). Influence of the Interannual Variability of Vegetation on the Surface Energy Balance—A Global Sensitivity Study. *Journal of Hydrometeorology, 3,* 617-629

Gutman, G., & Ignatov, A. (1998). The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *International Journal of Remote Sensing*, *19*, 1533-1543

Guyon, D., Guillot, M., Vitasse, Y., Cardot, H., Hagolle, O., Delzon, S., & Wigneron, J.-P. (2011). Monitoring elevation variations in leaf phenology of deciduous broadleaf forests from SPOT/VEGETATION time-series. *Remote Sensing of Environment, 115*, 615-627

Jiang, B., Liang, S., Wang, J., & Xiao, Z. (2010). Modeling MODIS LAI time series using three statistical methods. *Remote Sensing of Environment, 114*, 1432-1444

Jiao, T., Liu, R., Liu, Y., Pisek, J., & Chen, J.M. (2014). Mapping global seasonal forest background reflectivity with Multi-angle Imaging Spectroradiometer data. *Journal of Geophysical Research: Biogeosciences, 119,* 2013JG002493

Jönsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to timeseries of satellite sensor data. *IEEE transactions on Geoscience and Remote Sensing, 40*, 1824-1832

Kandasamy, S., Baret, F., Verger, A., Neveux, P., & Weiss, M. (2013). A comparison of methods for smoothing and gap filling time series of remote sensing observations: application to MODIS LAI products. *Biogeosciences, 10,* 4055-4071

Kovalskyy, V., Roy, D.P., Zhang, X.Y., & Ju, J. (2011). The suitability of multi-temporal web-enabled Landsat data NDVI for phenological monitoring – a comparison with flux tower and MODIS NDVI. *Remote Sensing Letters, 3,* 325-334

Masson, V., Champeaux, J.L., Chauvin, F., Meriguer, C., & Lacaze, R. (2003). A global database of land surface parameters at 1km resolution in meteorological and climate models. *Journal of Climate*, *16*, 1261-1282

McCallum, I., Wagner, W., Schmullius, C., Shvidenko, A., Obersteiner, M., Fritz, S., & Nilsson, S. (2009). Satellite-based terrestrial production efficiency modeling. *Carbon Balance and Management, 4:8*

Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J.L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G.R., Lotsch, A., Friedl, M., Morisette, J.T., Votava, P., Nemani, R.R., & Running, S.W. (2002). Global products of vegetation leaf area and absorbed PAR from year one of MODIS data. *Remote Sensing of Environment, 83*, 214-231

Pisek, J., Chen, J.M., Alikas, K., & Deng, F. (2010). Impacts of including forest understory brightness and foliage clumping information from multiangular measurements on leaf area index mapping over North America. *Journal of Geophysical Research: Biogeosciences, 115*, G03023

Pouliot, D., Latifovic, R., Fernandes, R., & Olthof, I. (2011). Evaluation of compositing period and AVHRR and MERIS combination for improvement of spring phenology detection in deciduous forests. *Remote Sensing of Environment, 115*, 158-166

Roujean, J.L., Leroy, M., & Deschamps, P.Y. (1992). A bidirectional reflectance model of the Earth's surface for the correction of remote sensing data. *Journal of Geophysical Research, 97*, 20455-20468

Samanta, A., Costa, M.H., Nunes, E.L., Vieira, S.A., Xu, L., & Myneni, R.B. (2011). Comment on “Drought-Induced Reduction in Global Terrestrial Net Primary Production from 2000 Through 2009”. *Science, 333*, 1093

Stöckli, R., Rutishauser, T., Baker, I., Liniger, M.A., & Denning, A.S. (2011). A global reanalysis of vegetation phenology. *Journal of Geophysical Research: Biogeosciences, 116*, G03020

Su, H., McCabe, M.F., Wood, E.F., Su, Z., & Prueger, J.H. (2005). Modeling Evapotranspiration during SMACEX: Comparing Two Approaches for Local- and Regional-Scale Prediction. *Journal of Hydrometeorology, 6*, 910-922 van den Hurk, B.J.J.M., Viterbo, P., & Los, S.O. (2003). Impact of leaf area index seasonality on the annual land surface evaporation in a global circulation model. *Journal of Geophysical Research: Atmospheres*, *108*, 4191

Verger, A., Baret, F., & Weiss, M. (2011). A multisensor fusion approach to improve LAI time series. *Remote Sensing of Enviroment, 115*, 2460-2470

Verger, A., Baret , F., & Weiss, M. (2014a). Near real time vegetation monitoring at global scale. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7*, 3473-3481

Verger, A., Baret, F., Weiss, M., Filella, I., & Peñuelas, J. (2014b). Quantification of leaf area index anomalies from global VEGETATION observations. In, *Global vegetation monitoring and modeling (available at https://colloque.inra.fr/gv2m/Oral-Sessions)*. Avignon (France)

Verger, A., Baret , F., Weiss, M., Kandasamy, S., & Vermote, E. (2013). The CACAO method for smoothing, gap filling and characterizing seasonal anomalies in satellite time series. *IEEE transactions on Geoscience and Remote Sensing, 51*, 1963-1972

Verhegghen, A., Bontemps, S., & Defourny, P. (2014). A global NDVI and EVI reference data set for land-surface phenology using 13 years of daily SPOT-VEGETATION observations. *International Journal of Remote Sensing, 35*, 2440-2471

Viterbo, P., & Beljaars, A.C.M. (1995). An Improved Land Surface Parameterization Scheme in the ECMWF Model and Its Validation. *Journal of Climate, 8*, 2716-2748

Weiss, M., Baret , F., Garrigues, S., Lacaze, R., & Bicheron, P. (2007). LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. part 2: Validation and comparison with MODIS Collection 4 products. *Remote Sensing of Environment, 110*, 317-331

White, M.A., de Beurs, K.M., Didan, K., Inouye, D.W., Richardson, A.D., Jensen, O.P., O'Keefe, J., Zhang, G., Nemani, R.R., van Leeuwen, W.J.D., Brown, J.F., de Wit, A., Schaepman, M., Lin, X., Dettinger, M., Bailey, A.S., Kimbal, J., Schwartz, M.D., Baldocchi, D.D., Lee, J.T., & Lauenroth, W.K. (2009). Intercomparison, interpretation, and assessment of spring phenology in

North America estimated from remote sensing for 1982-2006 *Global Change Biology, 15*, 2335–2359

Yang, W., Shabanov, N.V., Huang, D., Wang, W., Dickinson, R.E., Nemani, R.R., Knyazikhin, Y., & Myneni, R.B. (2006a). Analysis of leaf area index products from combination of MODIS Terra and Aqua data. *Remote Sensing of Environment, 104*, 297-312

Yang, W., Tan, B., Huang, D., Rautiainen, M., Shabanov, N., Wang, Y., Privette, J.L., Huemmrich, K.F., Fensholt, R., Sandholt, I., Weiss, M., Ahl, D.E., Gower, S.T., Nemani, R.R., Knyazikhin, Y., & Myneni, R. (2006b). MODIS leaf area index products: from validation to algorithm improvement. *IEEE transactions on Geoscience and Remote Sensing, 44*, 1885-1898

Zhang, X., Friedl, M.A., & Schaaf, C.B. (2006). Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. *Journal of Geophysical Research: Biogeosciences, 111*, G04017

Zhang, X., Friedl, M.A., & Schaaf, C.B. (2009). Sensitivity of vegetation phenology detection to the temporal resolution of satellite data. *International Journal of Remote Sensing, 30*, 2061-2074

Zhang, X., Friedl, M.A., Schaaf, C.B., & Strahler, A.H. (2004). Climate controls on vegetation phenological patterns in northern mid- and high latitudes inferred from MODIS data. *Global Change Biology, 10*, 1133-1145

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Reed, B.C., & Huete, A. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment, 84*, 471-475

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R., & Myneni, R. (2013). Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of Photosynthetically Active Radiation (FPAR)3g Derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the Period 1981 to 2011. *Remote Sensing, 5*, 927-948

WWW sites

www1: GEOV1 Biophysical Products.

http://land.copernicus.eu/global (last accessed 22 June2015)

www2: ECOCLIMAP code and data

https://opensource.cnrm-game-meteo.fr/projects/ecoclimap (last accessed 22 June2015)

www3: MODIS Reprojection Tool

https://lpdaac.usgs.gov/tools/modis_reprojection_tool (last accessed 22 June2015)

List of Figure Captions

Figure 1. Illustration of climatology computation over a grassland site (17.98ºS, 16.82ºE). (a) Time series of the GEOV1 LAI product for 1999-2010. (b) LAI values over the entire period (dots) in dekadal steps. The outputs (bold line) are computed as the interannual means LAI values over a 30-day compositing window at dekadal steps. The shaded areas correspond to 75% (dark grey), 85% (medium grey), and 95% (light grey) of the population of values for a given date. DOY, day of the year.

Figure 2. Illustration of TSGF correction for LAI for four GLOBCOVER biome classes (Defourny et al. 2009): (a) grassland, (b) needleleaf forest, (c) evergreen broadleaf forest, and (d) bare soil. The thin lines and grey intervals represent the raw model output computed as interannual means and the associated standard deviations. The thick lines represent the final GEOCLIM output corrected for artefacts. The dashed lines represent the number of available observations (Nb). The latitudes and longitudes of the sites are indicated. DOY, day of the year.

Figure 3. Global maps of (a) the maximum standard deviations (Max. SD) of interannual LAI values observed over the 36 dekads and (b) the minimum number (Min. Nb) of valid observations over the 36 dekads for computing GEOCLIM. The areas in grey correspond to pixels with no data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Figure 4. (a) Map of northern high latitudes for which the sun zenith angle, SZA>70º, areas of bare soil and evergreen broadleaf forest where specific corrections were applied in GEOCLIM. (b) Simplified GLOBCOVER (Defourny et al. 2009) land-cover map after aggregating the 22 original classes into six main land-cover classes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Figure 5. GEOCLIM global maps of (a) the maximum LAI at the peak of the growing season (Max. LAI), (b) the mean annual LAI (Mean LAI), (c) the standard deviation of interannual LAI values for the date of the peak (interannual variability) (SD Max. LAI), and (d) the standard deviation of the mean LAI annual cycle (seasonal variability) (SD Mean LAI). The areas in dark grey correspond to pixels with no data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Figure 6. Maps of the mean LAI differences between (a) GEOCLIM and ECOCLIMAP, (b) GEOCLIM and MODIS, and (c) GEOCLIM and GIMMS3g. The percentage of land pixels for each interval of mean LAI differences is indicated on the right of the colour bar. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Figure 7. Temporal profiles of GEOCLIM LAI averaged for 30º latitudinal bands and mean LAI differences as compared to ECOCLIMAP, MODIS and GIMMS3g. DOY, day of the year.

Figure 8. Distributions of GEOCLIM, ECOCLIMAP, MODIS and GIMMS3g LAIs per biomes based on the simplified GLOBCOVER (Defourny et al. 2009) land-cover map (Figure 4b): (a) shrubs/savannah/bare soil, (b) crops and grassland, (c) deciduous broadleaf forests, (d) evergreen broadleaf forests, (e) needleleaf forests, and (f) all biomes.

Figure 9. Global map of the day of the year (DOY) for maximum foliar development (date of peak of the growing season) derived from GEOCLIM. The areas of bare soil and evergreen broadleaf forests (Figure 4a) with insufficient seasonality for computing the phenological metrics are shaded in light grey. The areas in dark grey correspond to pixels with missing data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Figure 10. Latitudinal transects at resolution of 0.5 degrees of the average day of the year (DOY) for maximum foliar development derived from GEOCLIM and mean differences as compared to the phenological metrics derived from ECOCLIMAP, MODIS, GIMMS3g and MCD12Q2.

Table 1. The root mean square error (RMSE), bias, standard deviation (σ), correlation coefficient (*R*), and slope of the regression line through the origin for comparisons between the dates of maximum foliar development derived from GEOCLIM, ECOCLIMAP, and MCD12Q2 at a global scale at 0.5º. The areas of bare soil and evergreen broadleaf forests (Figure 4a) with insufficient seasonality for computing the phenological metrics and 10% outliers were not included.

	RMSE	Bias	σ	*R*	Slope
GEOCLIM – MCD12Q2	24.18	14.36	19.46	0.94	1.08
GEOCLIM – ECOCLIMAP	30.37	5.98	29.78	0.89	1.01
GEOCLIM – MODIS	14.38	0.97	14.35	0.98	1.00
GEOCLIM – GIMMS3g	17.65	-0.41	17.64	0.96	0.99

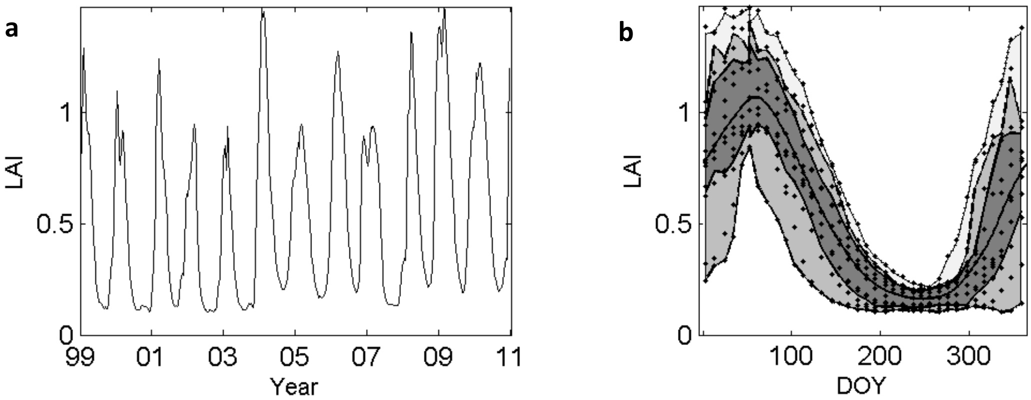


Figure 1. Illustration of climatology computation over a grassland site (17.98ºS, 16.82ºE). (a) Time series of the GEOV1 LAI product for 1999-2010. (b) LAI values over the entire period (dots) in dekadal steps. The outputs (bold line) are computed as the interannual means LAI values over a 30-day compositing window at dekadal steps. The shaded areas correspond to 75% (dark grey), 85% (medium grey), and 95% (light grey) of the population of values for a given date. DOY, day of the year.

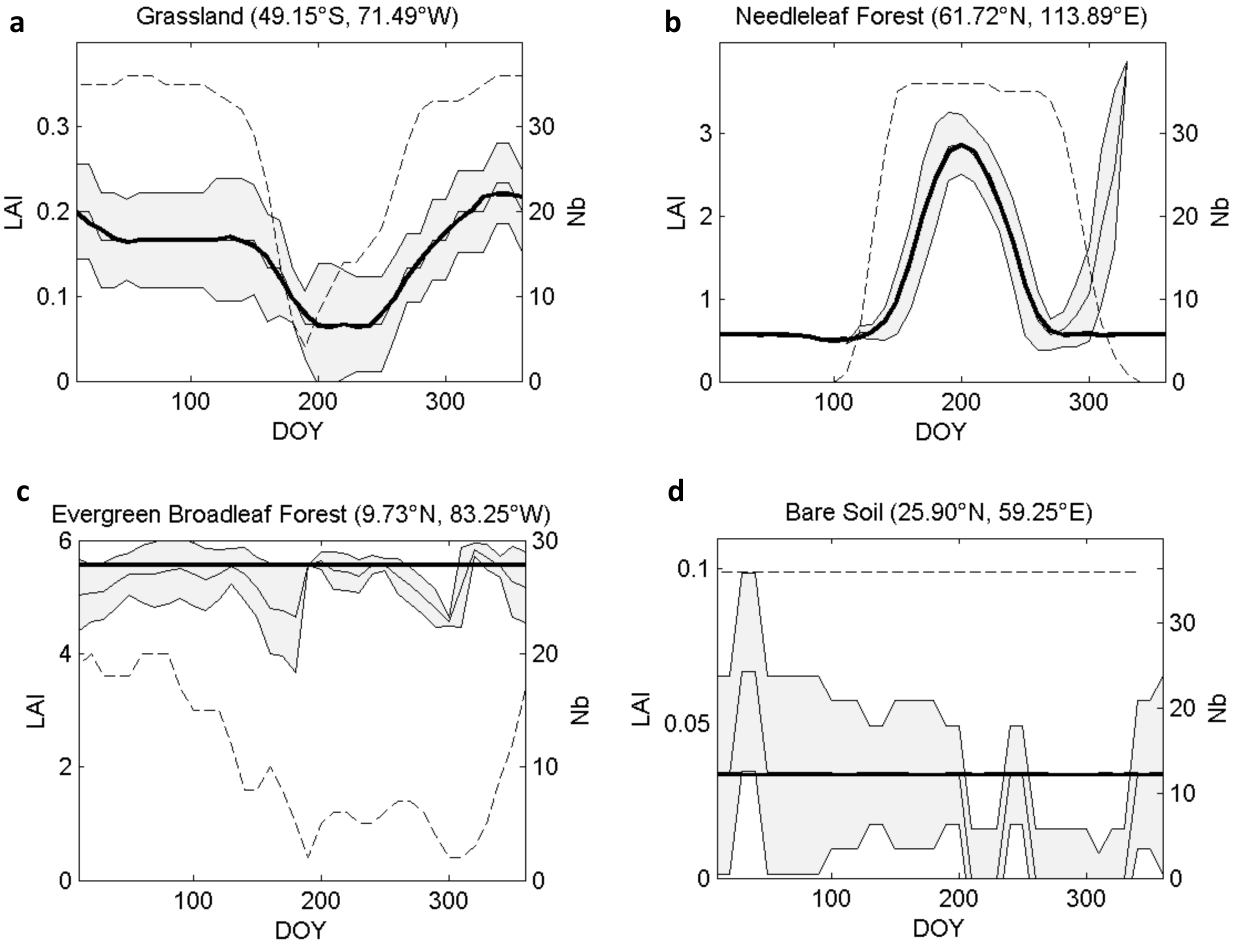


Figure 2. Illustration of TSGF correction for LAI for four GLOBCOVER biome classes (Defourny et al. 2009): (a) grassland, (b) needleleaf forest, (c) evergreen broadleaf forest, and (d) bare soil. The thin lines and grey intervals represent the raw model output computed as interannual means and the associated standard deviations. The thick lines represent the final GEOCLIM output corrected for artefacts. The dashed lines represent the number of available observations (Nb). The latitudes and longitudes of the sites are indicated. DOY, day of the year.

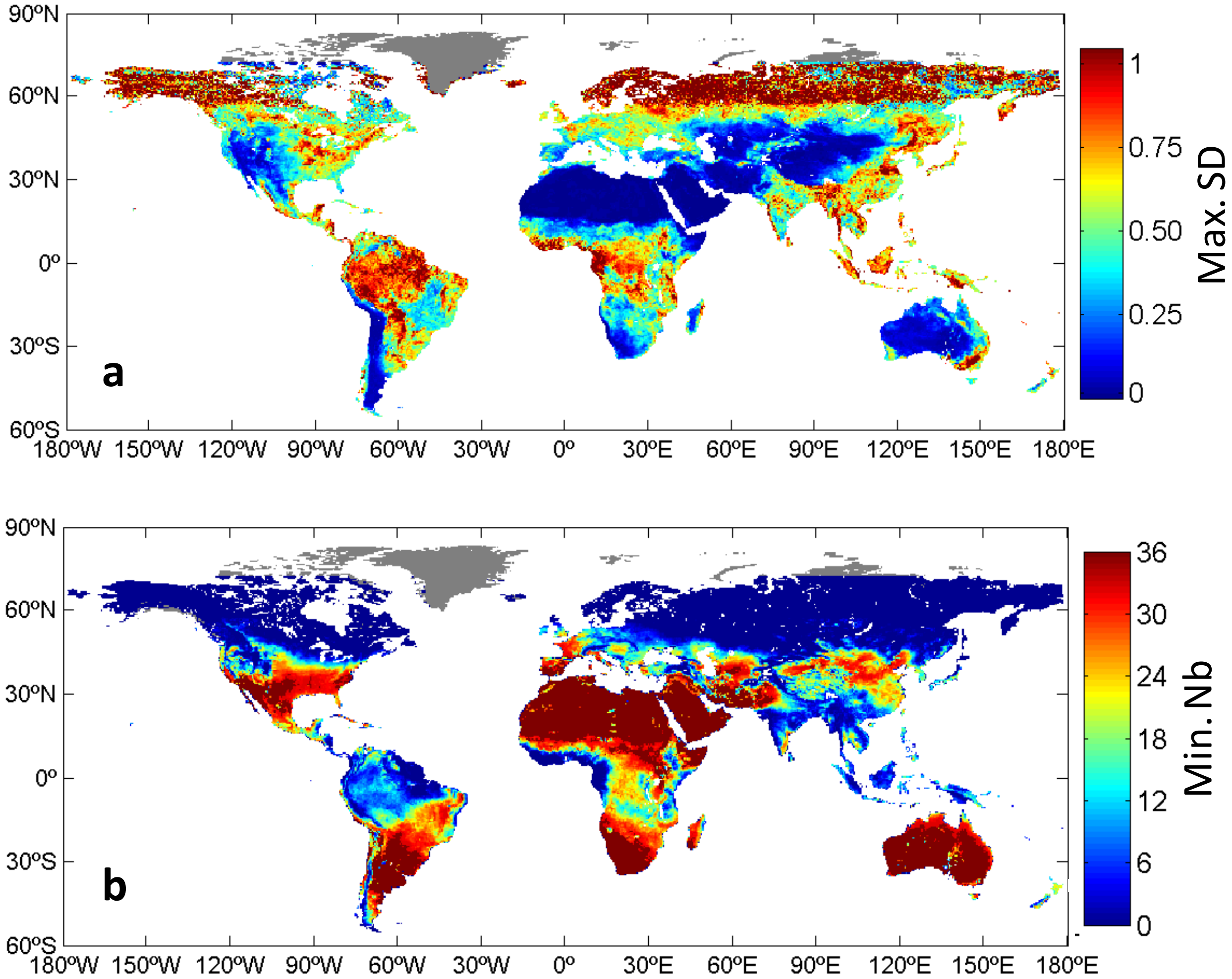


Figure 3. Global maps of (a) the maximum standard deviations (Max. SD) of interannual LAI values observed over the 36 dekads and (b) the minimum number (Min. Nb) of valid observations over the 36 dekads for computing GEOCLIM. The areas in grey correspond to pixels with no data.

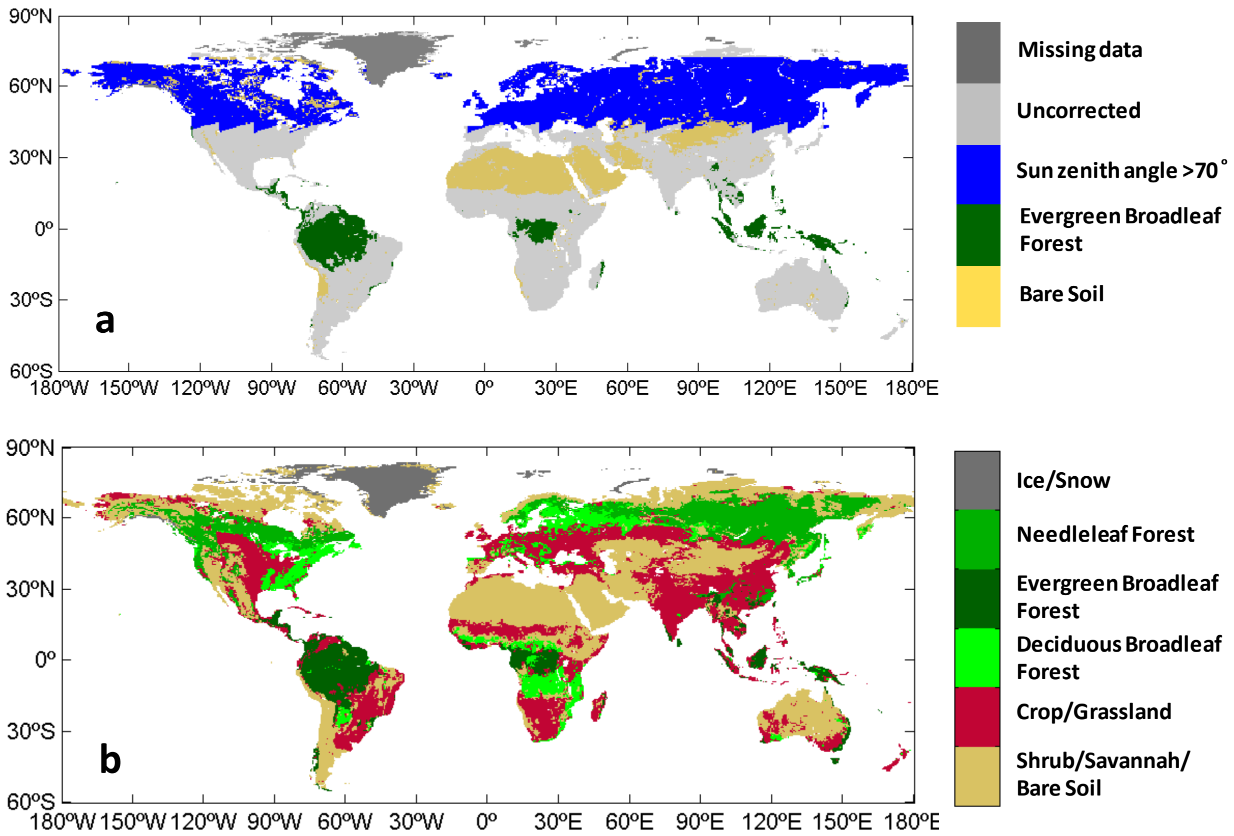


Figure 4. (a) Map of northern high latitudes for which the sun zenith angle, SZA>70º, areas of bare soil and evergreen broadleaf forest where specific corrections were applied in GEOCLIM. (b) Simplified GLOBCOVER (Defourny et al. 2009) land-cover map after aggregating the 22 original classes into six main land-cover classes.

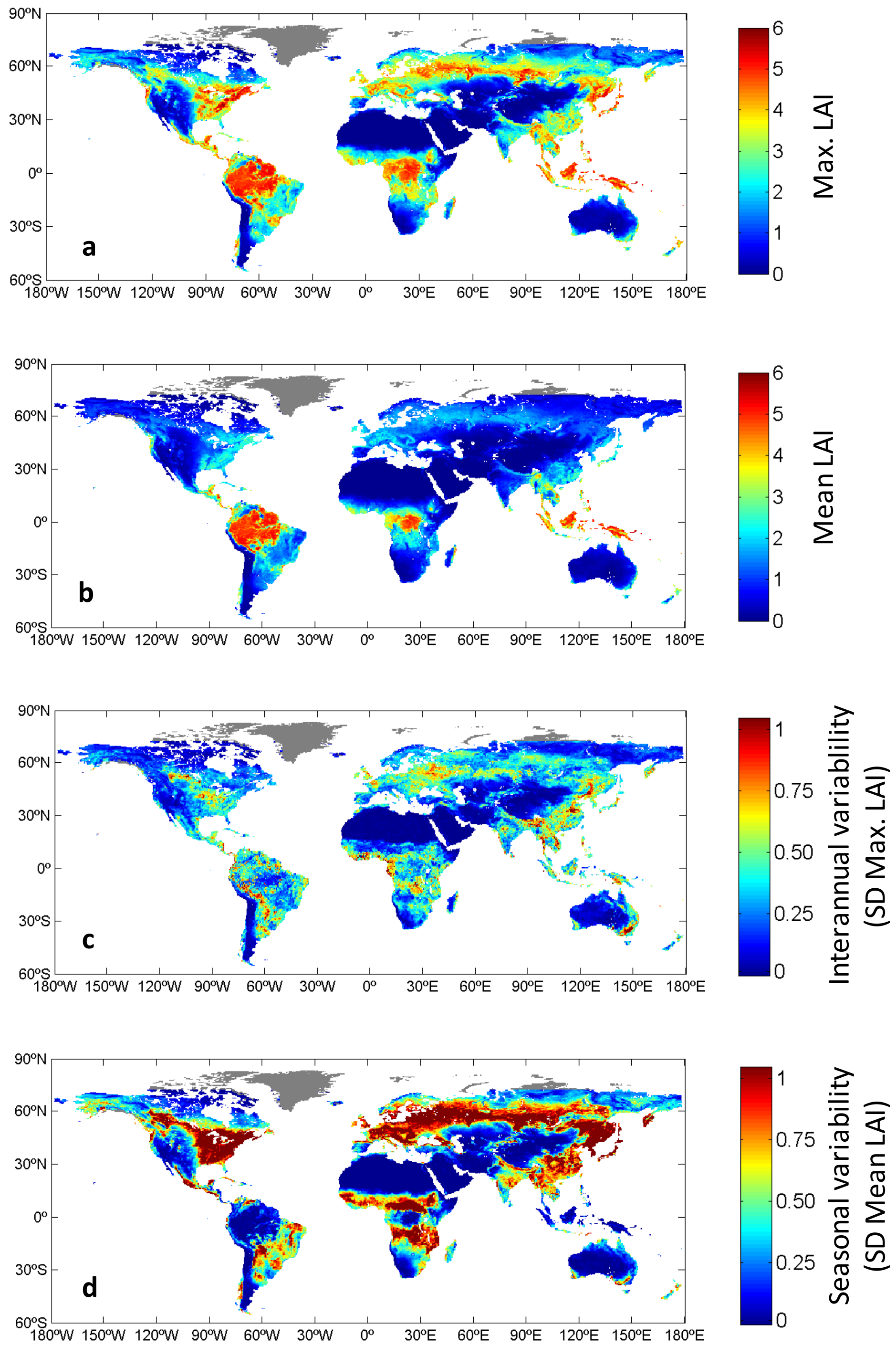


Figure 5. GEOCLIM global maps of (a) the maximum LAI at the peak of the growing season (Max. LAI), (b) the mean annual LAI (Mean LAI), (c) the standard deviation of interannual LAI values for the date of the peak (interannual variability) (SD Max. LAI), and (d) the standard deviation of the mean LAI annual cycle (seasonal variability) (SD Mean LAI). The areas in dark grey correspond to pixels with no data.

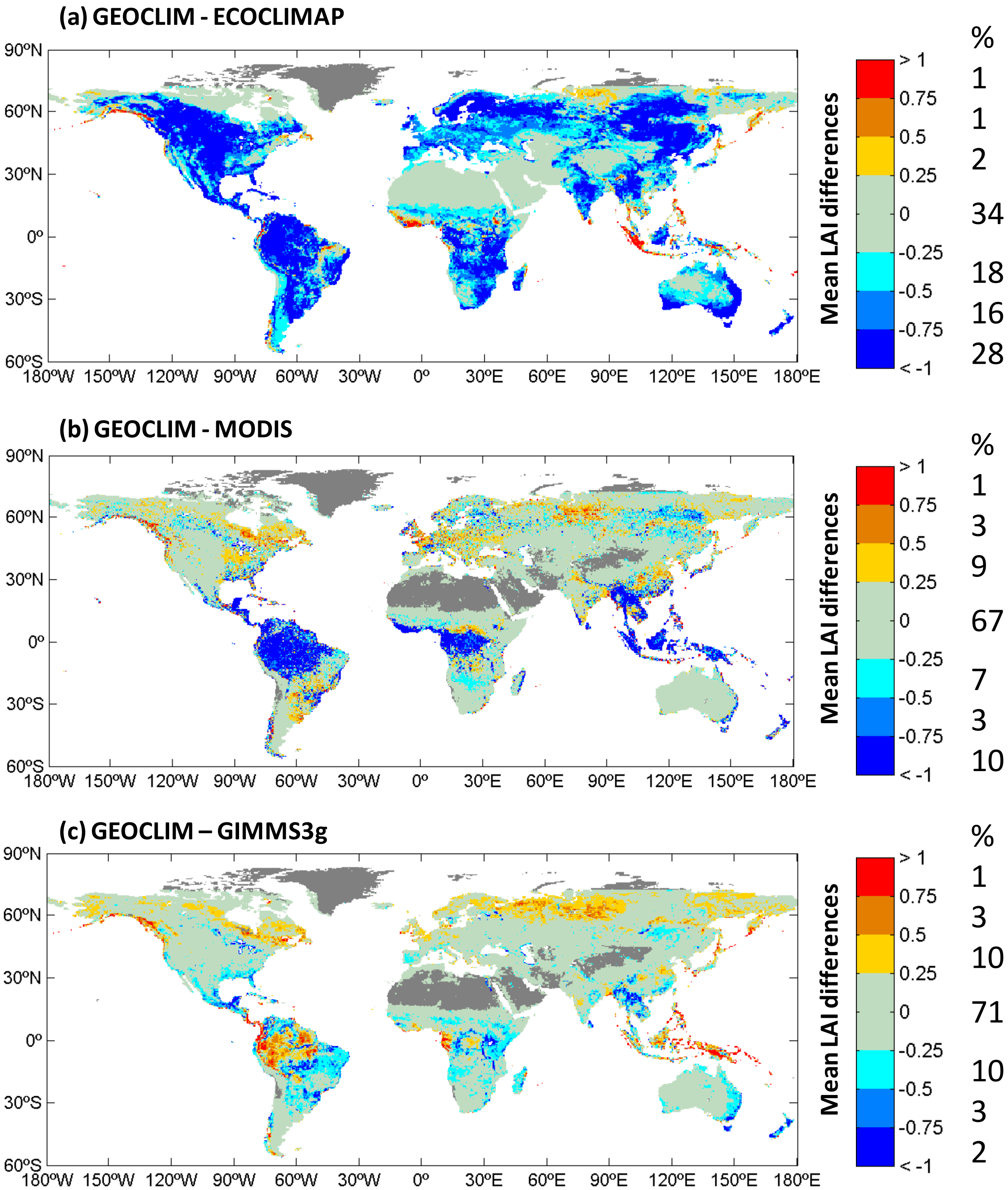


Figure 6. Maps of the mean LAI differences between (a) GEOCLIM and ECOCLIMAP, (b) GEOCLIM and MODIS, and (c) GEOCLIM and GIMMS3g. The percentage of land pixels for each interval of mean LAI differences is indicated on the right of the colour bar.

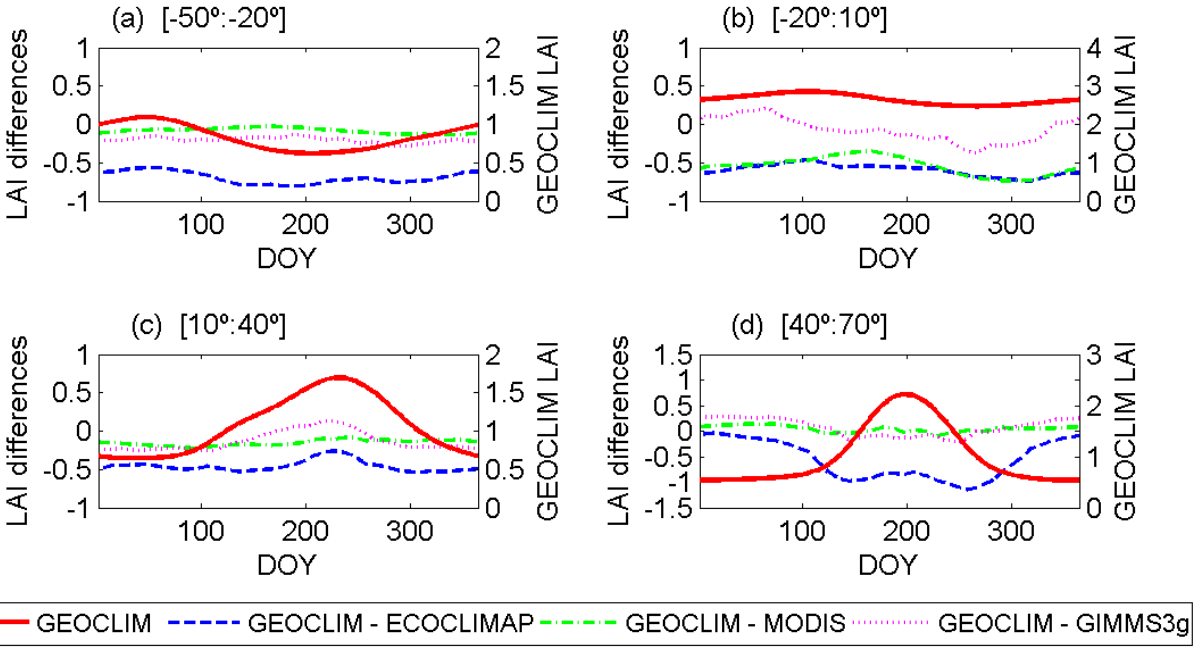


Figure 7. Temporal profiles of GEOCLIM LAI averaged for 30º latitudinal bands and mean LAI differences as compared to ECOCLIMAP, MODIS and GIMMS3g. DOY, day of the year.

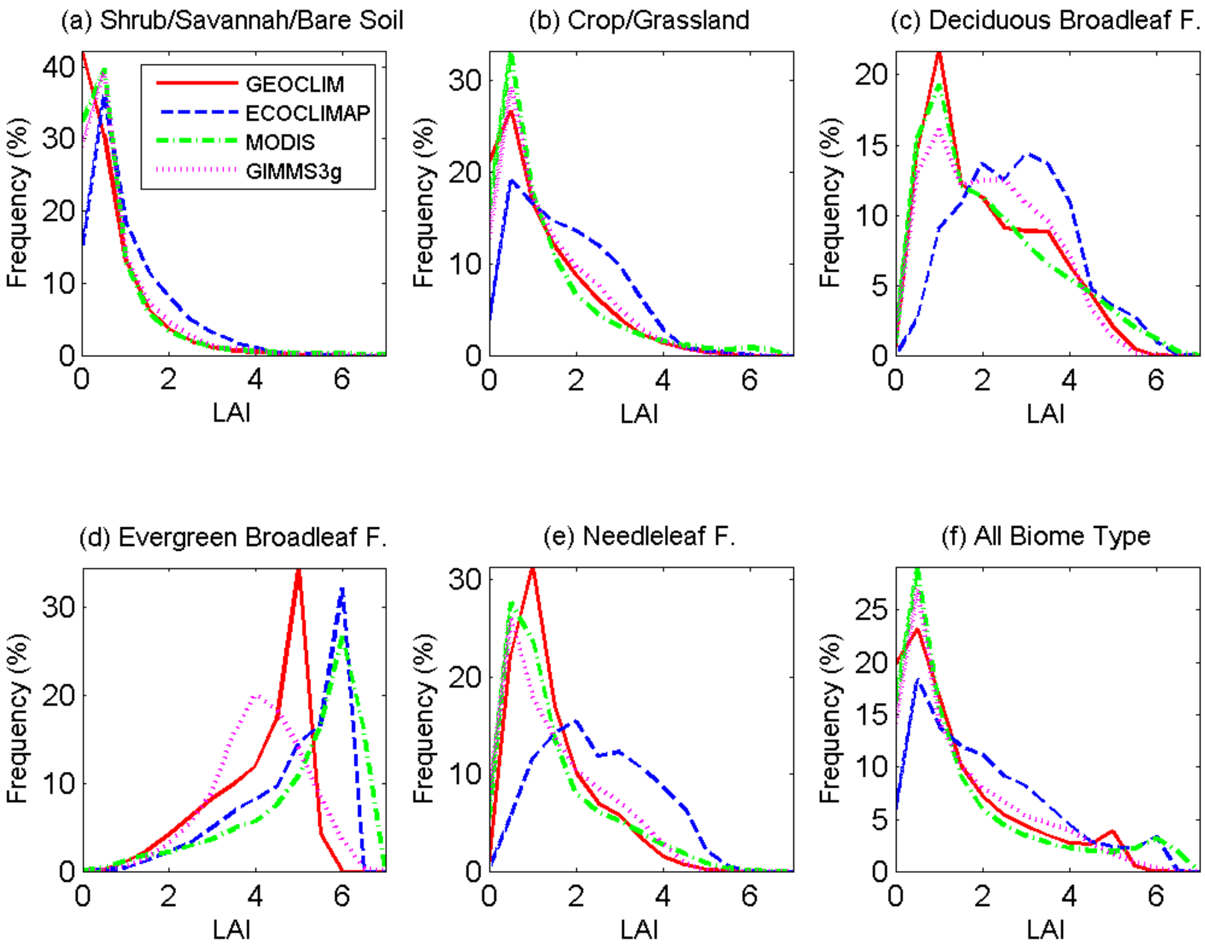


Figure 8. Distributions of GEOCLIM, ECOCLIMAP, MODIS and GIMMS3g LAIs per biomes based on the simplified GLOBCOVER (Defourny et al. 2009) land-cover map (Figure 4b): (a) shrubs/savannah/bare soil, (b) crops and grassland, (c) deciduous broadleaf forests, (d) evergreen broadleaf forests, (e) needleleaf forests, and (f) all biomes.

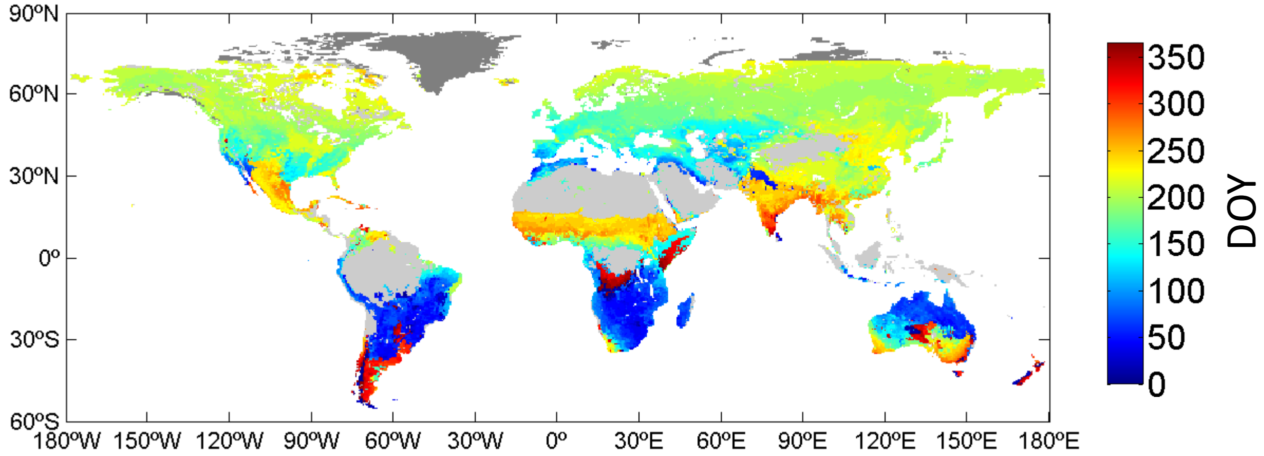
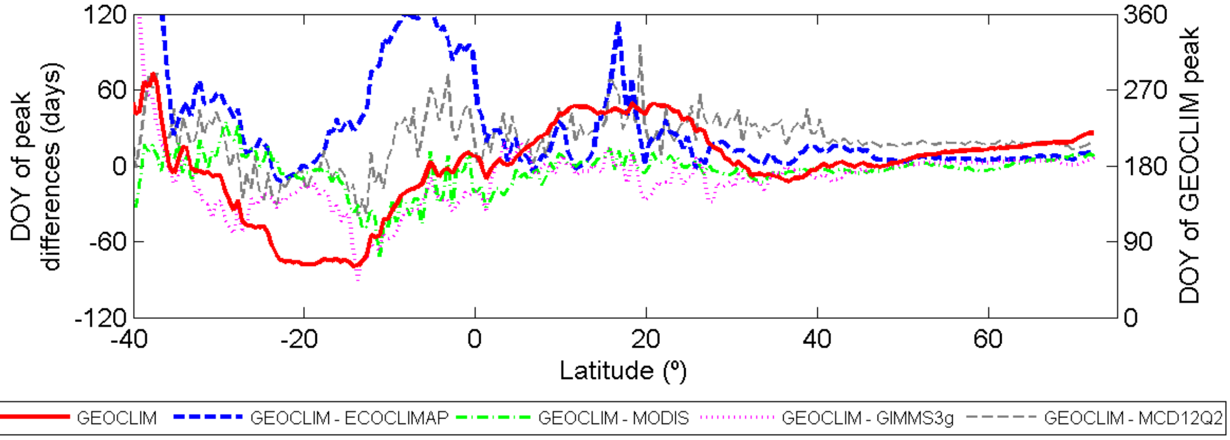


Figure 9. Global map of the day of the year (DOY) for maximum foliar development (date of peak of the growing season) derived from GEOCLIM. The areas of bare soil and evergreen broadleaf forests (Figure 4a) with insufficient seasonality for computing the phenological metrics are shaded in light grey. The areas in dark grey correspond to pixels with missing data.

826 

827 **Figure 10.** Latitudinal transects at resolution of 0.5 degrees of the average day of the year (DOY)
828 for maximum foliar development derived from GEOCLIM and mean differences as compared to the
829 phenological metrics derived from ECOCLIMAP, MODIS, GIMMS3g and MCD12Q2.

830