

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

3 **Aleixandre Verger**^{1,2,3*}, **Frédéric Baret**³, **Marie Weiss**³, **Iolanda Filella**^{1,2}, **Josep Peñuelas**^{1,2}

4 ¹ CREAF, Cerdanyola del Vallès 08193, Catalonia, Spain

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

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

7 * Corresponding author: E-Mail: verger@creaf.uab.cat; Tel.: + 34- 935-813-008; Fax: +34-935-
8 814-151.

9 **Abstract**

10 Land-surface modelling would benefit significantly from improved characterisation of the seasonal
11 variability of vegetation at a global scale. GEOCLIM, a global climatology of leaf area index (LAI),
12 fraction of absorbed photosynthetically active radiation (FAPAR), both essential climate variables,
13 and fraction of vegetation cover (FCOVER) is here derived from observations from the SPOT
14 VEGETATION programme. Interannual average values from the GEOV1 Copernicus Global Land
15 time series of biophysical products at 1-km resolution and 10-day frequency are computed for 1999
16 to 2010. GEOCLIM provides the baseline characteristics of the seasonal cycle of the annual
17 vegetation phenology for each 1-km pixel on the globe. The associated standard deviation
18 characterises the interannual variability. Temporal consistency and continuity is achieved by the
19 accumulation of multi-year observations and the application of techniques for temporal smoothing
20 and gap filling. Specific corrections are applied over cloudy tropical regions and high latitudes in
21 the Northern Hemisphere where the low number of available observations compromises the
22 reliability of estimates. Artefacts over evergreen broadleaf forests and areas of bare soil are

23 corrected based on the expected limited seasonality. The GEOCLIM data set is demonstrated to be
24 consistent, both spatially and temporally. GEOCLIM shows absolute differences lower than 0.5
25 compared with MODIS (GIMMS3g) climatology of LAI for more than 80% (90%) of land pixels,
26 with higher discrepancies in tropical and boreal latitudes. ECOCLIMAP systematically produces
27 higher LAI values. The phenological metric for the date of maximum foliar development derived
28 from GEOCLIM is spatially consistent (correlation higher than 0.9) with those of MODIS,
29 GIMMS3g, ECOCLIMAP and MCD12Q2 with average differences within 14 days at the global
30 scale.

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

33 **1. Introduction**

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

46 as a very pertinent variable that can be used in models of the surface-energy balance to separate the
47 contribution of the soil from that of the canopy (Gutman and Ignatov 1998; Su et al. 2005).

48 LAI, FAPAR, and FCOVER are routinely estimated from sensors with medium resolution such as
49 VEGETATION (Baret et al. 2013), Moderate Resolution Imaging Spectroradiometer (MODIS)
50 (Myneni et al. 2002) and the Advanced Very High Resolution Radiometer (AVHRR) (Zhu et al.
51 2013). The European Copernicus Global Land Service delivers global LAI, FAPAR, and FCOVER
52 products from SPOT VEGETATION data from 1999 to the present with a spatial sampling close to
53 1 km. The products, known as GEOV1 products, have benefitted from the development and
54 validation of existing products (Baret et al. 2013). Camacho et al (2013) demonstrated that GEOV1
55 products were more accurate and precise than current products.

56 Some land-surface models (LSMs) for simulating terrestrial water and carbon cycles use the
57 spatiotemporal variation of LAI or FAPAR, described by different lookup tables depending on the
58 type of vegetation (Viterbo and Beljaars 1995). The availability of satellite data in the last two
59 decades describing the state and evolution of vegetation has allowed a better integration of
60 biophysical variables into LSMs. Previous studies have demonstrated the improved performance of
61 LSMs due to a better characterisation of the seasonal and interannual variability of vegetation
62 functioning provided by the assimilation of satellite data. In particular, data assimilation yields a
63 more realistic parameterisation in phenological models and reduces the models' prediction errors to
64 21 and 15% for FAPAR and LAI, respectively (Stöckli et al. 2011). A number of studies have
65 shown the potential of assimilating LAI observations to correct vegetation model states (Demarty et
66 al. 2007; Gu et al. 2006) and the implications of introducing the observed seasonal (van den Hurk et
67 al. 2003) and interannual (Guillevic et al. 2002) variability of LAI in the annual cycle of
68 hydrological fluxes. Boussetta et al (2013) showed that the assimilation of a MODIS derived LAI
69 monthly climatology, i.e. the interannual average of LAI time series (as opposed to a vegetation-
70 dependent constant LAI), in a model of global numerical weather prediction improved the forecast

71 of near-surface (screen-level) air temperature and relative humidity through its effect on
72 evapotranspiration. Barbu et al (2014) more recently demonstrated the potential of the assimilation
73 of GEOV1 LAI into an ISBA-A-gs land-surface model to improve the monitoring of droughts. A
74 LAI climatology was also useful for the identification of anomalies and trends in global vegetation
75 (Baret et al. 2012; Brandt et al. 2014; Verger et al. 2014b; Verger et al. 2013; Zhu et al. 2013). The
76 climatology of the biophysical variables reveals the seasonality inherent to the land-cover type and
77 improve land-cover classification (Verhegghen et al. 2014). A climatology gap filling can better
78 cope with missing and noise-contaminated data than can standard temporal filters for most missing
79 data or large gaps in a single annual time series of satellite data, which have a large impact on the
80 accuracy of the phenological metrics extracted from the reconstructed time series (Guyon et al.
81 2011; Kandasamy et al. 2013; Verger et al. 2013). Extraction of phenological information is also
82 sensitive to the temporal (Pouliot et al. 2011; Zhang et al. 2009) and spatial (Fisher and Mustard
83 2007; Kovalskyy et al. 2011) resolution of the satellite data. The climatology derived from time
84 series of moderate spatial resolution sensors preserves the high temporal frequency mandatory for
85 phenological studies (Guyon et al. 2011). Finally, the climatology information can make projections
86 and improve the stability of near real time estimates (Jiang et al. 2010; Verger et al. 2014a).

87 Despite the significance of global phenology for earth system monitoring and modelling, there are
88 few data sets that explicitly describe the annual vegetation cycle at global scale. Boussetta et al.
89 (2013) derived a monthly LAI climatology from 2000-2008 MODIS observations to be used in a
90 numerical weather prediction model as indicator of the leaf development stage. The ECOCLIMAP
91 programme (Faroux et al. 2013; Masson et al. 2003) is a dual database at 1 km resolution that
92 includes an ecosystem classification and a coherent set of land surface parameters (including LAI,
93 FAPAR and FCOVER) that are primarily mandatory in meteorological modelling for
94 soil/vegetation–atmosphere transfer schemes. Other studies focus on the time variation of
95 vegetation indices to propose a global normalized difference vegetation index (NDVI) and

96 enhanced vegetation index (EVI) reference data set for land-surface phenology using 13 years of
97 VEGETATION observations (Verhegghen et al. 2014) or to derive phenological metrics from
98 VEGETATION (Guyon et al. 2011), MODIS (Ganguly et al. 2010; Zhang et al. 2003) or AVHRR
99 (Atzberger et al. 2013) time series.

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

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

118 **2. GEOCLIM implementation**

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

122 **2.1. GEOV1 biophysical products**

123 GEOV1 biophysical products provide global coverage of LAI, FAPAR, and FCOVER from
124 1998/12/24 to the present at a ground sampling distance of $1/112^\circ$ (approximately 1 km at the
125 equator) and 10-day steps. A neural-network machine-learning algorithm was used to estimate
126 GEOV1 products (Baret et al. 2013). Directionally normalised VEGETATION reflectances
127 (Roujean et al. 1992) from the top of the canopy in the red, near-infrared, and short-wave infrared
128 bands derived from the CYCLOPES processing line (Baret et al. 2007) were used as inputs to the
129 neural networks. Based on the validation results for the available biophysical products (Garrigues et
130 al. 2008; Weiss et al. 2007), the MODIS and CYCLOPES products were selected for the training
131 process. The selected products were combined after re-projection onto the VEGETATION Plate-
132 Carrée $1/112^\circ$ grid, smoothed over time, interpolated at the 10-day frequency, combined, and
133 eventually re-scaled to better fit the expected range of variation. Further details for the training of
134 the neural networks and the generation of the product are provided in Baret et al (2013). Recent
135 validation studies indicated that GEOV1 outperformed existing products in both accuracy and
136 precision (Camacho et al. 2013). GEOV1 products are freely available at www1.

137 **2.2. Climatology generation**

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

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

160 **[Figure 1]**

161 **[Figure 2]**

162 **2.3. Correction of specific artefacts**

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

165 - Some artefacts were observed at northern high latitudes during the winter: anomalous
166 seasonality and unexpected increases in LAI (FAPAR, FCOVER) (Figure 2b) with an
167 artificial maximum peak in winter (Figure 2b) and high interannual variability resulting in
168 high standard deviations (Figure 3a). These artefacts were mainly due to snow cover or
169 very poor conditions of illumination that limited the number of valid observations and the
170 reliability of the bidirectional reflectance model applied for the correction of
171 VEGETATION data (Roujean et al. 1992) (Figures 2b, 3b). The LAI (FAPAR, FCOVER)
172 values are expected to be relatively stable and low due to the low temperatures, short days,
173 and low illumination during winter at these high latitudes. To correct these artefacts at
174 northern latitudes, the GEOCLIM inputs higher than the 20th percentile during winter
175 (defined here as the period for which the *sun zenith angle*, $SZA > 70^\circ$ at the time of
176 VEGETATION overpass, i.e. around 10:30) were fixed at the minimum pixel values
177 observed over the entire period. We used the minimum values by preferentially selecting
178 the values computed from at least three valid observations because the quality of
179 GEOCLIM inputs is highly correlated with the number of valid observations available for
180 their composition (Figures 2b, 3). We used the minimum value computed over all dekads if
181 none of the dekads verified this condition. The areas where this specific correction was
182 applied are shown in Figure 4a. Similar approaches based on representative winter values
183 and thresholds to fill gaps and correct values affected by snow or poor illumination at high
184 latitudes were also considered by Beck et al. (2006); Delbart et al. (2005); Zhang et al.
185 (2004).

186 - Significant artefacts were also detected at equatorial and tropical latitudes due to aerosol-
187 cloud contamination that produced high instabilities, artificial seasonalities, and missing
188 data in the GEOV1 products and, consequently, in the derived output (Figure 2c). The high
189 standard deviations (Figure 3a) and the low number of available observations (Figures 2c,

190 3b) appeared to be good indicators of the high uncertainty associated with the computed
191 output over these tropical areas. Most of these cases corresponded to *evergreen broadleaf*
192 *forests* (EBFs) (cf. Figures 3, 4b), so a minimum seasonality and high LAI values should
193 be observed. We thus identified a pixel as an EBF if the 90th percentile (P90) of the LAI
194 output was >4.5 and the 20th percentile was $>P90-1.5$. This method for the detection of
195 EBFs based only on GEOV1 products (Figure 4a) agreed well with the GLOBCOVER
196 land-cover map (Defourny et al. 2009) (Figure 4b). For EBFs, the GEOCLIM values were
197 fixed to the 90th percentile computed over the entire period (Figure 2c).

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

203 [Figure 3]

204 [Figure 4]

205 3. Evaluation of GEOCLIM

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

213 time step because one-month corresponds to the lowest temporal sampling among the validation
214 datasets.

215 The global phenology derived from the temporal seasonality of GEOCLIM is also investigated in
216 this section. For simplicity, we focus on the date of maximum foliar development, i.e. the timing of
217 the peak of the growing season in the LAI annual cycle (Brown et al. 2012; Jönsson and Eklundh
218 2002). In addition to the phenological metrics derived from the LAI climatologies, we use also the
219 MCD12Q2 MODIS phenological product (Zhang et al. 2003).

220 **3.1. Validation data sets**

221 **3.1.1. ECOCLIMAP product**

222 ECOCLIMAP is a database at $1/112^\circ$ resolution on a Plate-Carrée grid resolution that includes a
223 classification of ecosystems and a consistent set of associated land-surface variables, including LAI,
224 at 10-day temporal sampling (Faroux et al. 2013). We used the latest Open-ECOCLIMAP v1
225 version, available since June 2014 at www2. It combines the global database of the first version,
226 ECOCLIMAP-I, and an upgraded version, ECOCLIMAP-II, for Europe. ECOCLIMAP-I contains
227 215 ecosystems obtained by combining existing land covers, climatic maps, and NDVI seasonal
228 profiles from AVHRR data acquired between April 1992 and March 1993 (Masson et al. 2003). For
229 each class of vegetation, the maximum and minimum LAI values are fixed based on in-situ
230 knowledge, and the annual cycle of LAI is constrained by the NDVI AVHRR temporal profiles
231 using a linear relationship between NDVI and LAI. The second version, ECOCLIMAP-II, contains
232 573 ecosystems across Europe based on more recent land-cover maps, and the annual LAI profiles
233 are derived from MODIS Collection 5 for the years 2002-2006 (Faroux et al. 2013).
234 The ECOCLIMAP LAI data set at the original $1/112^\circ$ resolution was aggregated at 0.5° spatial
235 sampling and averaged at monthly temporal sampling.

236 **3.1.2. MODIS climatology**

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

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

249 **3.1.3. GIMMS3g climatology**

250 The GIMMS3g LAI product derived from AVHRR data is available at 15-day temporal steps and
251 1/12° spatial resolution for the period July 1981 to December 2011. The principles used for the
252 generation of this LAI data set are based on the use of neural networks which were trained first with
253 GIMMS NDVI3g and MODIS LAI products for the overlapping period 2000-2009. The trained
254 neural network algorithm is then applied using the land-cover class, the latitude and longitude
255 coordinates, and the NDVI3g as the input data to generate the full temporal coverage of the
256 GIMMS3g LAI data set. Further details on the algorithm for GIMMS3g retrieval can be found in
257 Zhu et al. (2013).

258 We derived the GIMMS3g climatology as the interannual mean of GIMMS3g LAI time series at
259 15-day temporal step for the period 1999-2010. We aggregated the 1/12° products at 0.5° spatial
260 resolution and averaged at monthly temporal sampling for comparison purposes.

261 **3.1.4. MCD12Q2 product**

262 MCD12Q2 (Collection 5) (Ganguly et al. 2010; Zhang et al. 2003) provides global yearly
263 vegetation phenologies at 500 m from 2001 to 2010 MODIS time series. The MCD12Q2 algorithm
264 uses a series of piecewise logistic functions fitted over the annual cycle of EVI data (Zhang et al.
265 2003). Among the transition dates provided by the MCD12Q2 product, the parameter “onset of
266 greenness maximum” is used here for comparisons with the parameter “peak of growing season”
267 derived from the LAI climatologies , i.e. the date for which the climatology reach its maximum
268 value in the LAI annual cycle. The onset of maximal greenness conceptually corresponds to the
269 transition date at which the annual cycle of the vegetation reaches maturity. This date is thus
270 expected to be earlier than the date of maximum vegetation. Zhang et al (2006) compared the
271 MCD12Q2 parameter to *in-situ* measurements and found that it corresponded to the time at which
272 85-90% of the individual leaves reached their final size.

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

277 **3.2. Spatiotemporal consistency**

278 The GEOCLIM biophysical variables had highly consistent spatial and temporal patterns (Figure
279 5a), in agreement with the global distributions of biomes (Figure 4b). The seasonal patterns of
280 GEOCLIM LAI also reflected the expected regimes of vegetation at the global scale. Evergreen

281 broadleaf forests exhibited null seasonality (Figure 5d) in the tropical belt where LAI was near 5
282 throughout the year (Figure 5b). Deserts also had no seasonality (Figure 5d) where LAI was near
283 zero (Figure 5b). These results were expected given the forcing applied for evergreen broadleaf
284 forests and bare soils (cf. Section 2.3). As expected, deciduous broadleaf forests and crops had the
285 highest seasonalities (Figure 5d). The observed seasonality in needleleaf forests (Figure 5d) with a
286 $LAI \leq 4$ (Figure 5a) and means near 2 (Figure 5b) agreed with the observed seasonality of the
287 understory layer, which can reach a LAI of ~ 2 or more in summer but which is often near zero in
288 winter (Chen et al. 1997; Jiao et al. 2014; Masson et al. 2003).

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

300

[Figure 5]

301 **3.3. Comparison with ECOCLIMAP, MODIS and GIMMS3g climatologies**

302 The map of annual mean differences between GEOCLIM and the LAI climatologies derived from
303 AVHRR and MODIS data (Figure 6) shows LAI differences of ± 0.5 for 54%, 83% and 91% of the
304 land pixels as compared with ECOCLIMAP, MODIS and GIMMS3g, respectively. GEOCLIM

305 produced systematically lower values than ECOCLIMAP for the remaining 46% of pixels, with
306 larger differences for dense forests (northern boreal and tropical forests) but with significant
307 differences also for crops (e.g. USA and eastern Asia) (Figure 6a). These systematic negative bias
308 of GEOCLIM as compared to ECOCLIMAP was evident across latitudes and along the year (Figure
309 7). GEOCLIM produced also systematically lower values than MODIS (Figure 6b) over tropical
310 forests with differences ~ 0.5 along the year (Figure 7b) and over northern deciduous broadleaf
311 forest during the maximum growing leaf development (lower frequencies for the maximum values
312 in Figure 8c). On the contrary GEOCLIM produced slightly higher LAI values than GIMMS3g
313 (Figure 6c) over Amazon and Indonesian evergreen broadleaf forests and over boreal needle leaf
314 forests in Russia and USA during the winter time (Figure 7d).

315

[Figure 6]

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

329

[Figure 7]

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

350

[Figure 8]

351 **3.4. Assessment of global phenology**

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

358 The phenological metrics (Figure 10) were spatially consistent in the timing of the maximum of the
359 growing season as derived from the different data sets and particularly between GEOCLIM and
360 MODIS (GIMMS3g) with uncertainties of around 14 days in terms of RMSE, bias of less than 1
361 day, a correlation higher than 0.95 and a slope of the linear regression close to the unity (Table 1).
362 The phenology derived from ECOCLIMAP was also highly spatially consistent with GEOCLIM
363 (correlation about 0.9, slope close to the unity and bias about 6 days, Table1) but it diverged to
364 some degree (uncertainties of about one month in terms of RMSE), mostly in the Southern
365 Hemisphere (Figure 10) in regions with a limited seasonality (Figure 5d). As expected (Section
366 3.1.4), the phenological phase for the “onset of greenness maximum” retrieved in MCD12Q2
367 occurred earlier than the peak date in GEOCLIM (bias of 14 days, Table 1) due to differences in the
368 definitions of the phenological metrics. The phenological metrics derived from GEOCLIM
369 constitutes an intermediate solution across latitudes between MODIS, GIMMS3g, ECOCLIMAP
370 and MCD12Q2 for the date of maximum foliar development (Figure 10)

371 **[Figure 9]**

372 **[Figure 10]**

373 **[Table 1]**

374 **4. Discussion**

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

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

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

416 LAI values were systematically higher for ECOCLIMAP than for GEOCLIM, MODIS and
417 GIMMS3g. Boussetta et al (2013) reported similar higher LAI values for ECOCLIMAP than for
418 MODIS. Garrigues et al (2008) also reported large positive biases for ECOCLIMAP compared with
419 CYCLOPES, MODIS and GLOBCARBON and with ground measurements. The differences in the
420 temporal period, input data and sensors (VEGETATION for GEOCLIM and AVHRR and MODIS
421 for ECOCLIMAP, Section 3.1) can partially account for the significant discrepancies between
422 GEOCLIM and ECOCLIMAP although the relatively good agreement of GEOCLIM with MODIS
423 and GIMMS3g AVHRR derived LAI climatologies indicates that the major source of discrepancies

424 are related to the retrieval algorithms. The linear relationship between NDVI and LAI used to
425 retrieve the ECOCLIMAP product for pixels out of Europe (Masson et al. 2003) may have
426 introduced some overestimation because the LAI-NDVI relationship is exponential and saturates at
427 medium to high values (e.g. Myneni et al. (2002)). Since ECOCLIMAP assumes low spatial
428 variability within each class of land cover, it is limited to capture the LAI spatial variability as
429 compared to other LAI datasets (Garrigues et al. 2008). Nevertheless, identifying the source of the
430 differences between ECOCLIMAP and the other LAI datasets being analyzed would require further
431 attention and it is out of the scope of this paper.

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

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

445 **5. Conclusions**

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

448 GEOV1 products built from 1999 to 2010. Results showed GEOCLIM was temporally consistent
449 for the seasonality across biomes and latitudes. GEOCLIM showed a high agreement with MODIS
450 and AVHRR climatologies of LAI: differences within the Global Climate Observing System
451 requirements, i.e. ± 0.5 LAI, in more than 80% (90%) of GEOCLIM land pixels compared with
452 MODIS (GIMMS3g). ECOCLIMAP systematically produced higher LAI values. Further research
453 should focus on tropical and boreal latitudes where higher uncertainties exist in the LAI datasets.
454 The phenology of the timing of maximum foliar development derived from GEOCLIM constituted
455 an intermediate solution between those of GIMMS3g, MODIS, ECOCLIMAP and MCD12Q2:
456 differences within 14 days and spatial correlation > 0.9 at the global scale.
457 The GEOCLIM data set is continuous, both spatially and temporally, and it can be used for a wide
458 range of land-biosphere applications. It may contribute to a better characterisation of the seasonal
459 variability of vegetation in global land-surface models. It provides the baseline characteristics of the
460 seasonal cycle of LAI, FAPAR, and FCOVER for the identification of anomalies and trends in
461 global vegetation.

462 **Acknowledgements**

463 This research was partially supported by the European Earth observation programme Copernicus
464 (388533), *Pôle Thématique Surfaces Continentales THEIA* (CNES 140570) and the FP7 geoland2
465 (218795), GIOBIO (32-566) and LONGLOVE (32-594) projects. This research was also supported
466 by the Spanish Government grant CGL2013-48074-P, the Catalan Government grant SGR 2014-
467 274, and the European Research Council Synergy grant ERC-2013-SyG-610028 IMBALANCE-P.
468 Alexandre Verger is the recipient of a *Juan de la Cierva* postdoctoral fellowship from the Spanish
469 Ministry of Science and Innovation. The authors are indebted to Ranga Myneni, Taejin Park, the
470 Boston University, MODIS, CNES, VITO, and CNRM-GAME (Météo France, CNRS) teams for
471 providing access to the data.

472 **References**

- 473 Atzberger, C., Klisch, A., Mattiuzzi, M., & Vuolo, F. (2013). Phenological Metrics Derived over
474 the European Continent from NDVI3g Data and MODIS Time Series. *Remote Sensing*, 6, 257-284
- 475 Barbu, A.L., Calvet, J.C., Mahfouf, J.F., & Lafont, S. (2014). Integrating ASCAT surface soil
476 moisture and GEOV1 leaf area index into the SURFEX modelling platform: a land data assimilation
477 application over France. *Hydrol. Earth Syst. Sci.*, 18, 173-192
- 478 Baret , F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Weiss, M.,
479 Samain, O., Roujean, J.L., & Leroy, M. (2007). LAI, fAPAR and fCover CYCLOPES global
480 products derived from VEGETATION. Part 1: Principles of the algorithm. *Remote Sensing of*
481 *Enviroment*, 110, 275-286
- 482 Baret , F., Verger, A., Weiss, M., Makhmara, H., Lacaze, R., & Pacholczyk, P. (2012). Use of
483 GEOV1 climatology based on 11 years of consistent observations: identification of trends at the
484 global scale. In, *EGU General Assembly 2012: Geophysical Research Abstracts*, Vol. 14,
485 EGU2012-7912
- 486 Baret , F., Weiss, M., & Kandasamy, S. (2010). Climatology of LAI, FAPAR and FCOVER
487 products (V1). In (p. 30). Avignon (France): INRA-EMMAH
- 488 Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholczyk, P., & Smets, B. (2013).
489 GEOV1: LAI, FAPAR Essential Climate Variables and FCOVER global time series capitalizing
490 over existing products. Part1: Principles of development and production. *Remote Sensing of*
491 *Environment*, 137, 299-309
- 492 Beck, P.S.A., Atzberger, C., Hogda, K.A., Johansen, B., & Skidmore, A.K. (2006). Improved
493 monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI.
494 *Remote Sensing of Environment*, 100, 321

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

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

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

504 Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., & Weiss, M. (2013). GEOV1: LAI, FAPAR
505 Essential Climate Variables and FCOVER global time series capitalizing over existing products.
506 Part 2: Validation and intercomparison with reference products. *Remote Sensing of Environment*,
507 137, 310-329

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

569 McCallum, I., Wagner, W., Schullius, C., Shvidenko, A., Obersteiner, M., Fritz, S., & Nilsson, S.
570 (2009). Satellite-based terrestrial production efficiency modeling. *Carbon Balance and*
571 *Management*, 4:8

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

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

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

582 Roujean, J.L., Leroy, M., & Deschamps, P.Y. (1992). A bidirectional reflectance model of the
583 Earth's surface for the correction of remote sensing data. *Journal of Geophysical Research*, 97,
584 20455-20468

585 Samanta, A., Costa, M.H., Nunes, E.L., Vieira, S.A., Xu, L., & Myneni, R.B. (2011). Comment on
586 "Drought-Induced Reduction in Global Terrestrial Net Primary Production from 2000 Through
587 2009". *Science*, 333, 1093

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

590 Su, H., McCabe, M.F., Wood, E.F., Su, Z., & Prueger, J.H. (2005). Modeling Evapotranspiration
591 during SMACEX: Comparing Two Approaches for Local- and Regional-Scale Prediction. *Journal*
592 *of Hydrometeorology*, 6, 910-922

593 van den Hurk, B.J.J.M., Viterbo, P., & Los, S.O. (2003). Impact of leaf area index seasonality on
594 the annual land surface evaporation in a global circulation model. *Journal of Geophysical Research:*
595 *Atmospheres*, 108, 4191

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

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

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

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

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

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

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

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

618 North America estimated from remote sensing for 1982-2006 *Global Change Biology*, 15,
619 2335–2359

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

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

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

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

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

635 Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Reed, B.C., & Huete,
636 A. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84,
637 471-475

638 Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R., &
639 Myneni, R. (2013). Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of
640 Photosynthetically Active Radiation (FPAR)3g Derived from Global Inventory Modeling and
641 Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the Period
642 1981 to 2011. *Remote Sensing*, 5, 927-948

643 **WWW sites**

644 www1: GEOV1 Biophysical Products.

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

646 www2: ECOCLIMAP code and data

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

649 www3: MODIS Reprojection Tool

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

651

652

653

654

655

656

657

658

659

660

661

662

663

664 **List of Figure Captions**

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

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

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

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

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

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

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

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

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

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

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727 **Table 1.** The root mean square error (RMSE), bias, standard deviation (σ), correlation coefficient
728 (R), and slope of the regression line through the origin for comparisons between the dates of
729 maximum foliar development derived from GEOCLIM, ECOCLIMAP, and MCD12Q2 at a global
730 scale at 0.5°. The areas of bare soil and evergreen broadleaf forests (Figure 4a) with insufficient
731 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

732

733

734

735

736

737

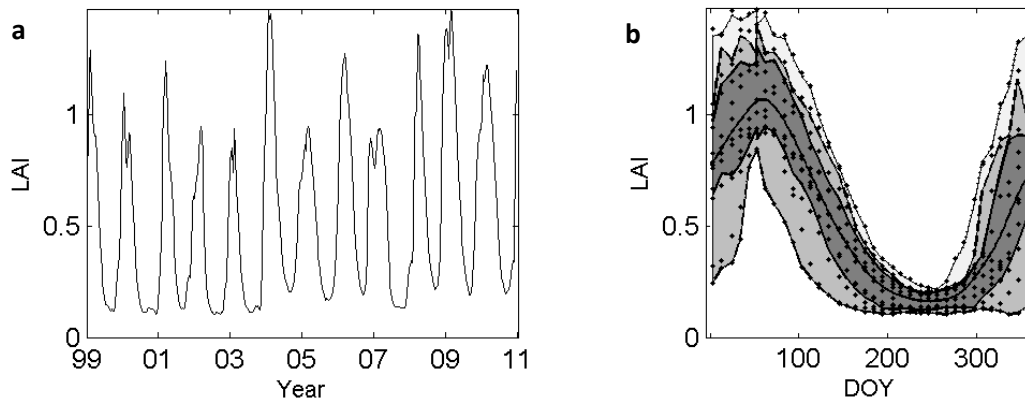
738

739

740

741

742



743

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

750

751

752

753

754

755

756

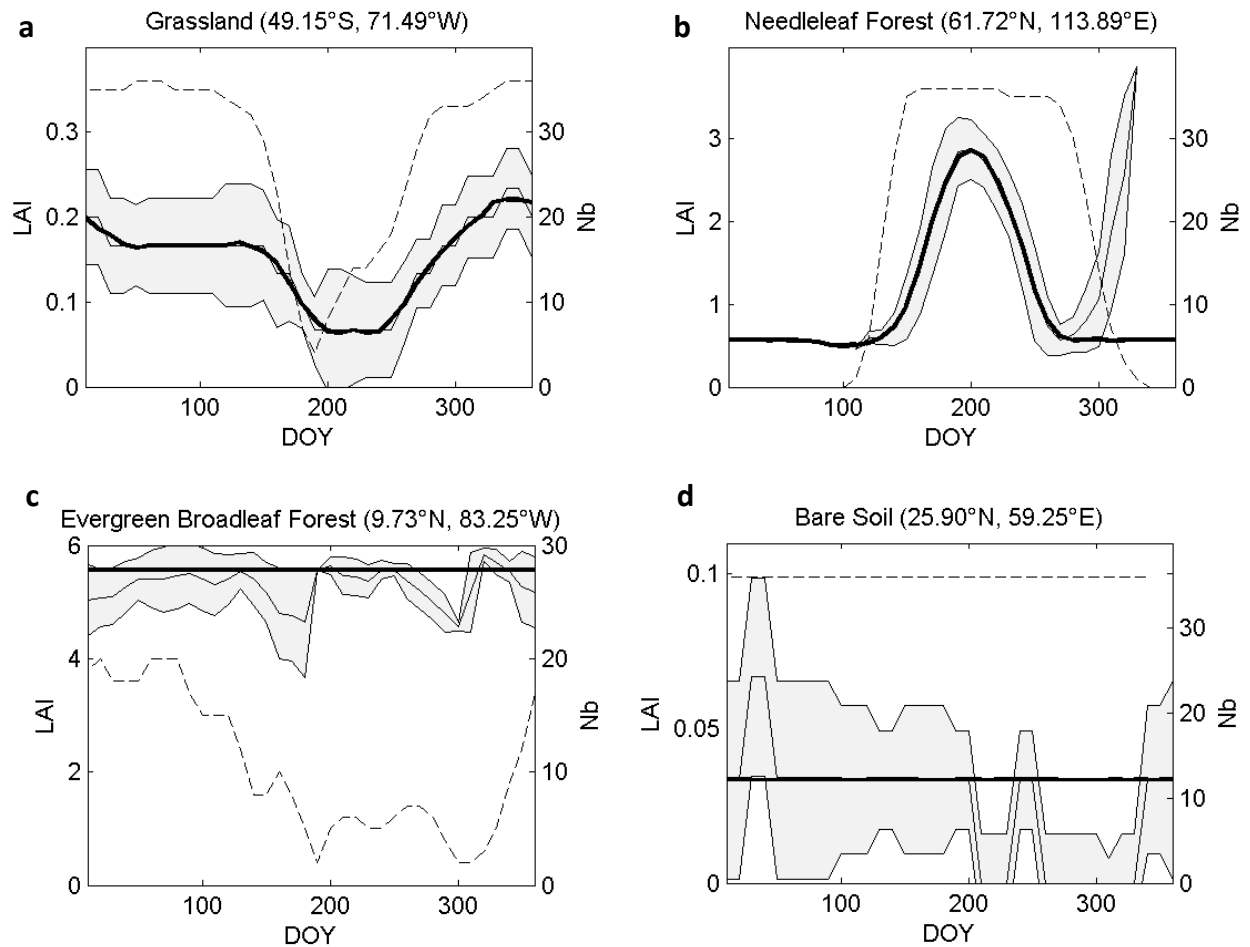
757

758

759

760

761



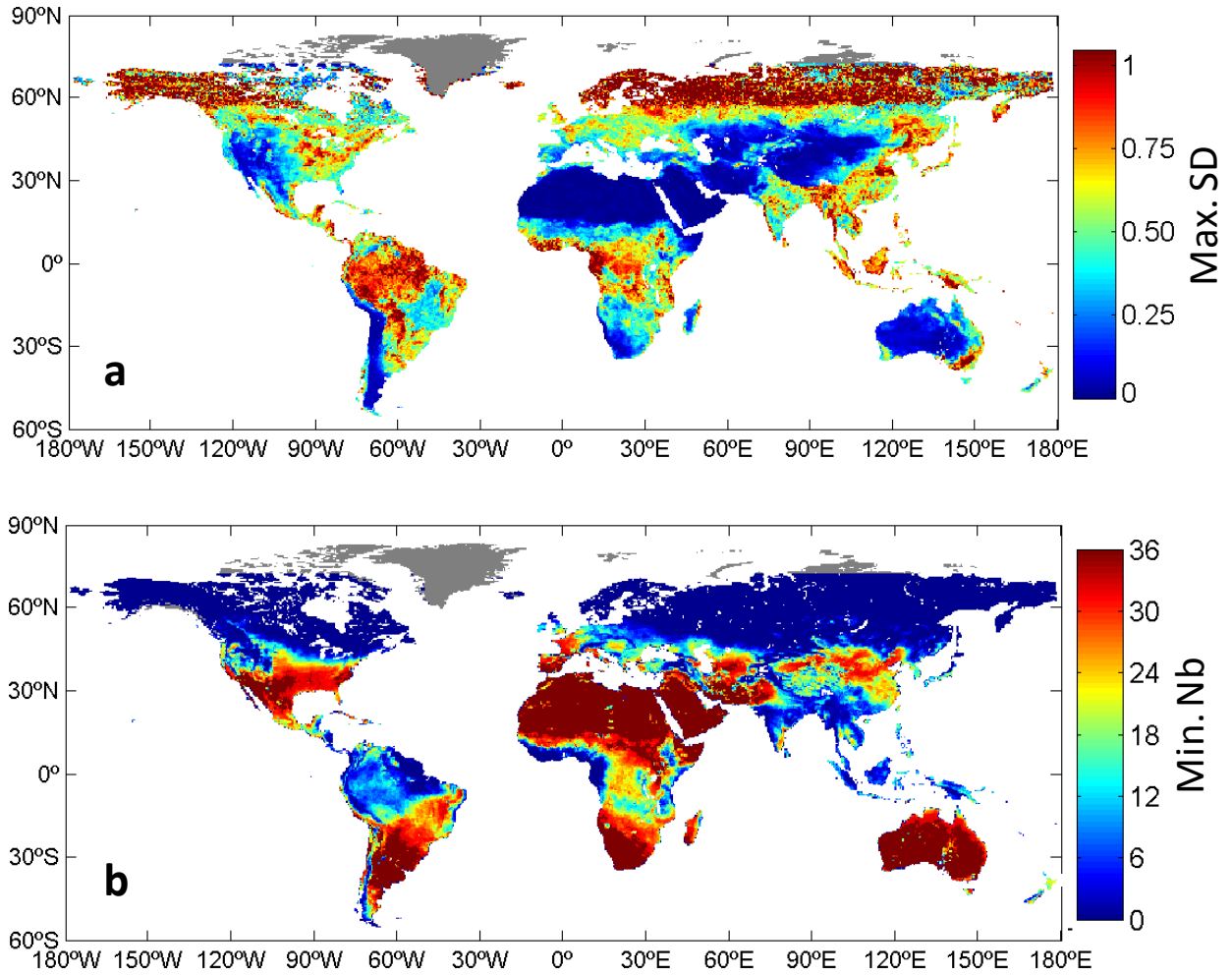
762
763

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.

769

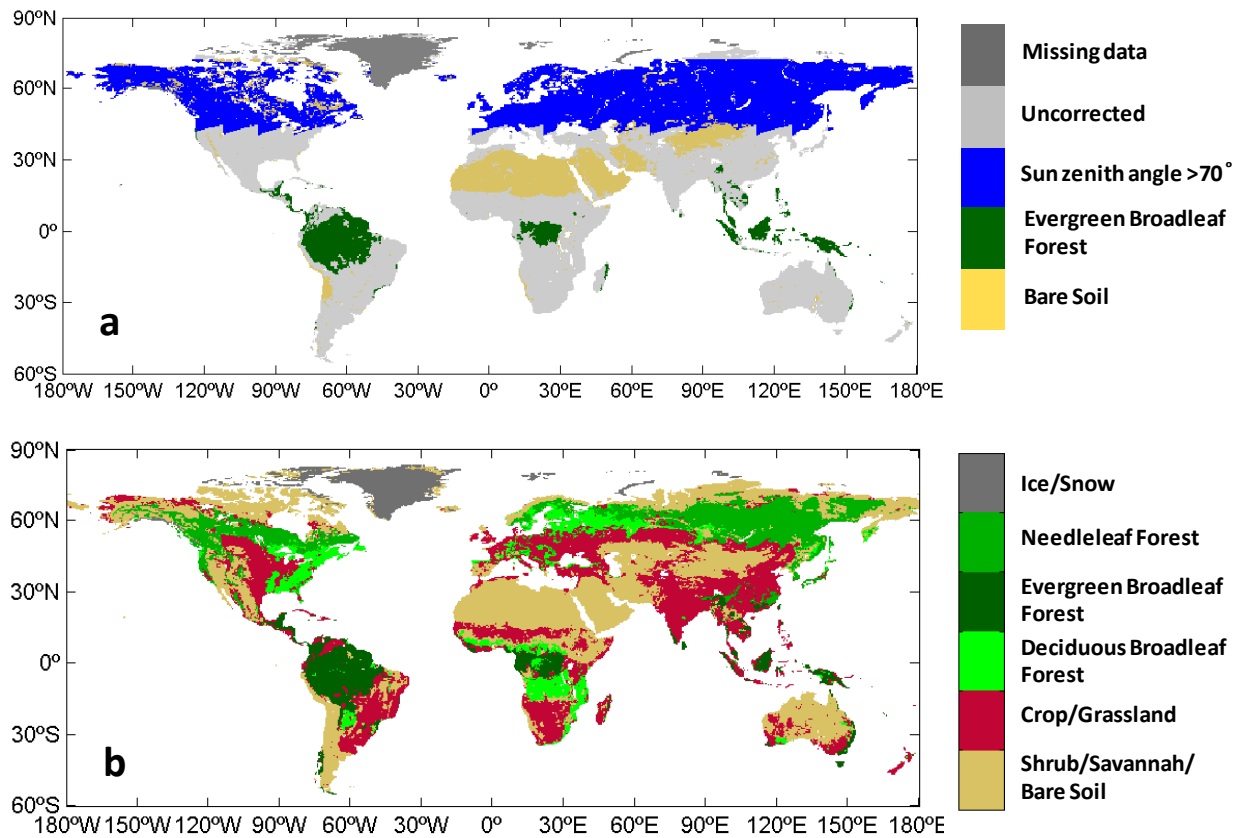
770

771



772

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



776

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

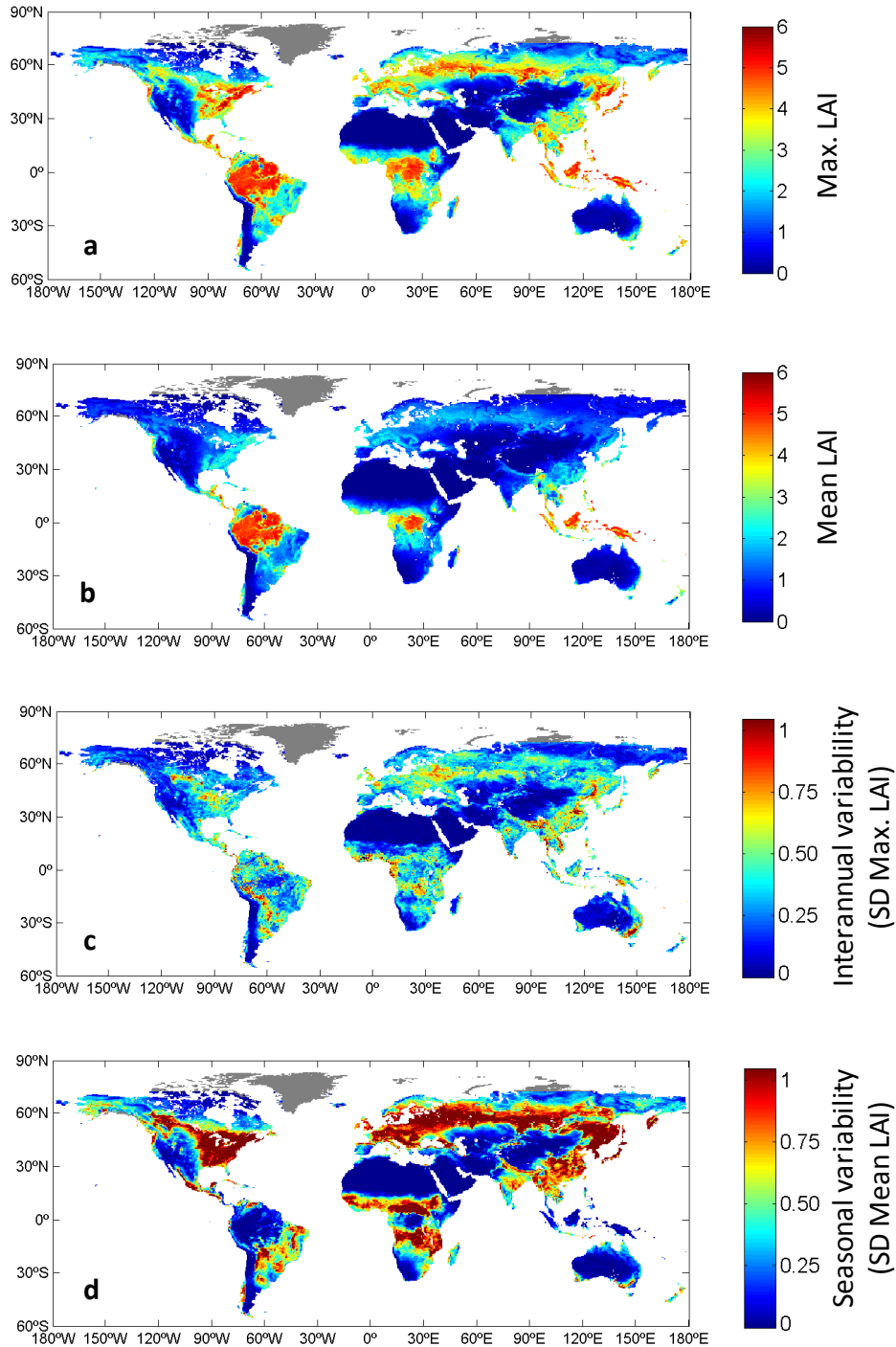
781

782

783

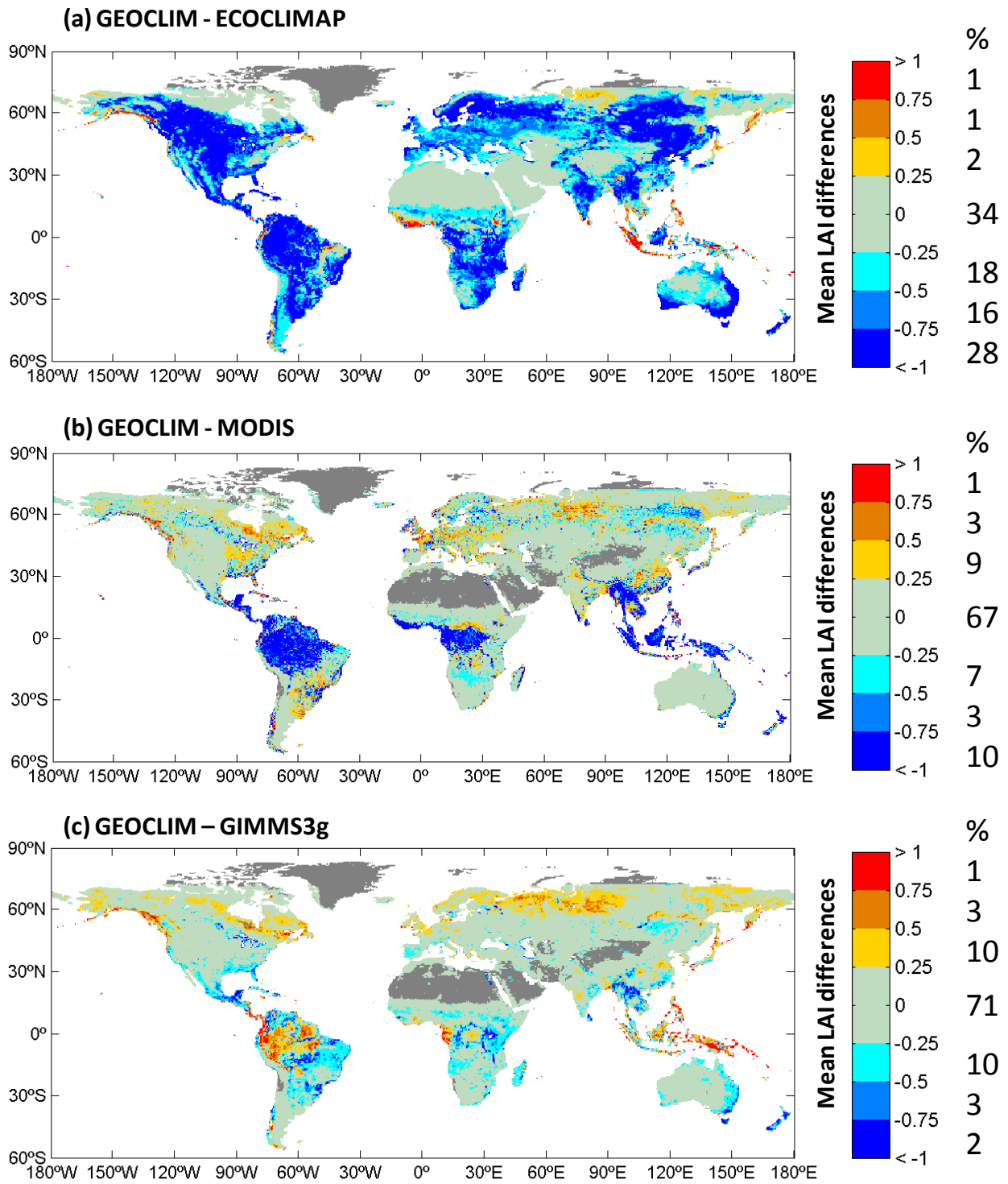
784

785



786
787

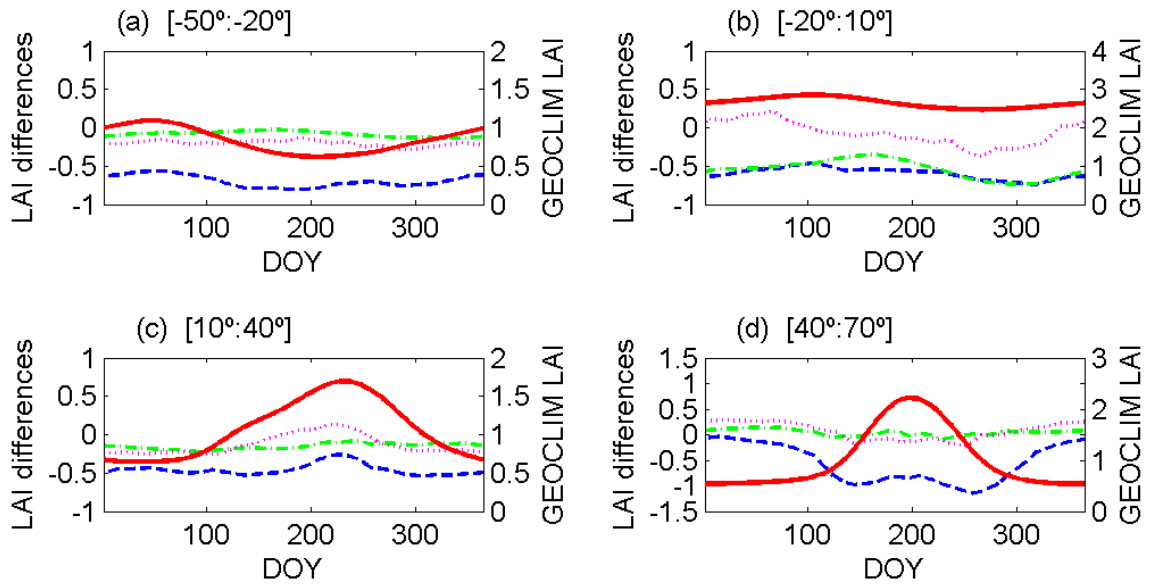
788 **Figure 5.** GEOCLIM global maps of (a) the maximum LAI at the peak of the growing season
789 (Max. LAI), (b) the mean annual LAI (Mean LAI), (c) the standard deviation of interannual LAI
790 values for the date of the peak (interannual variability) (SD Max. LAI), and (d) the standard
791 deviation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



793

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

797

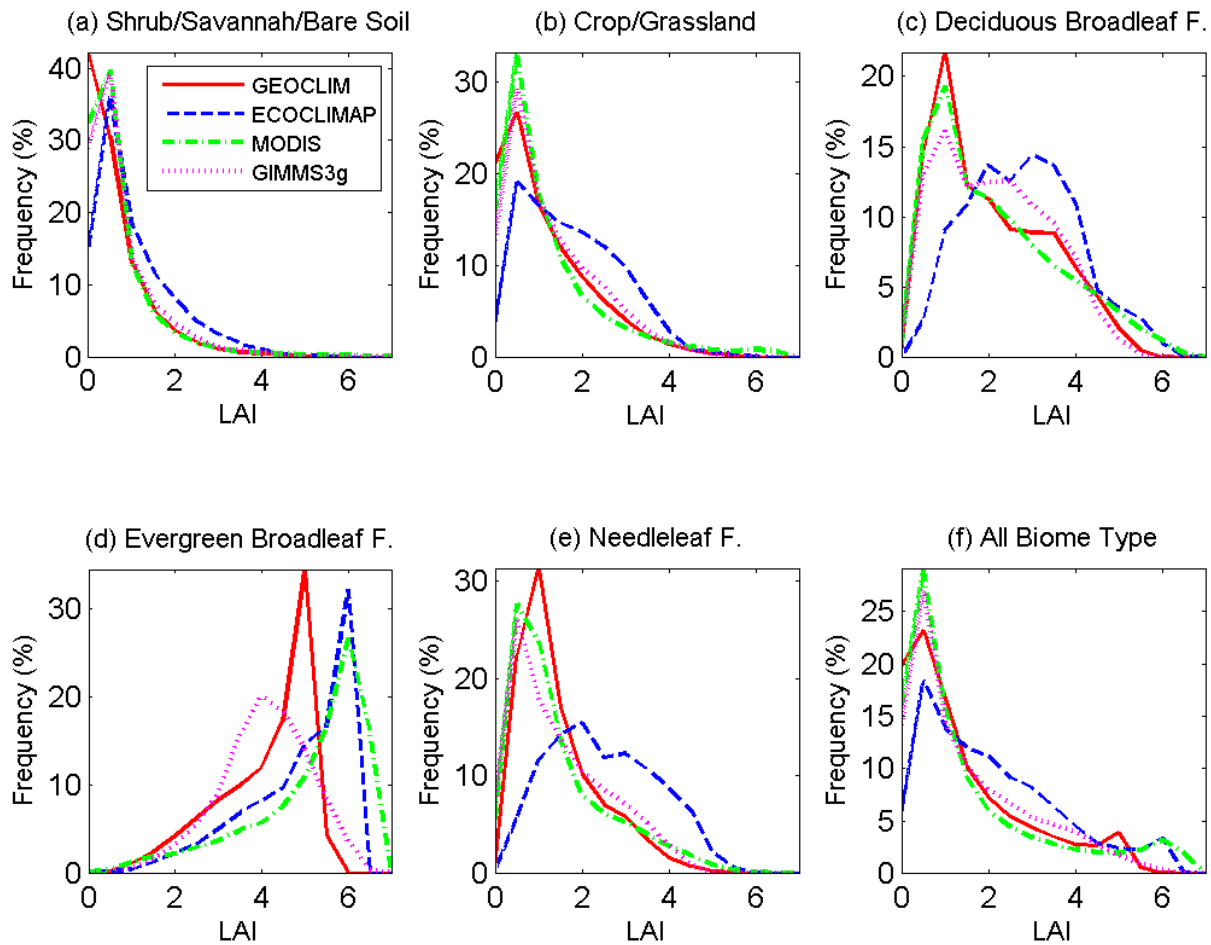


798

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

801

802



803

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

808

809

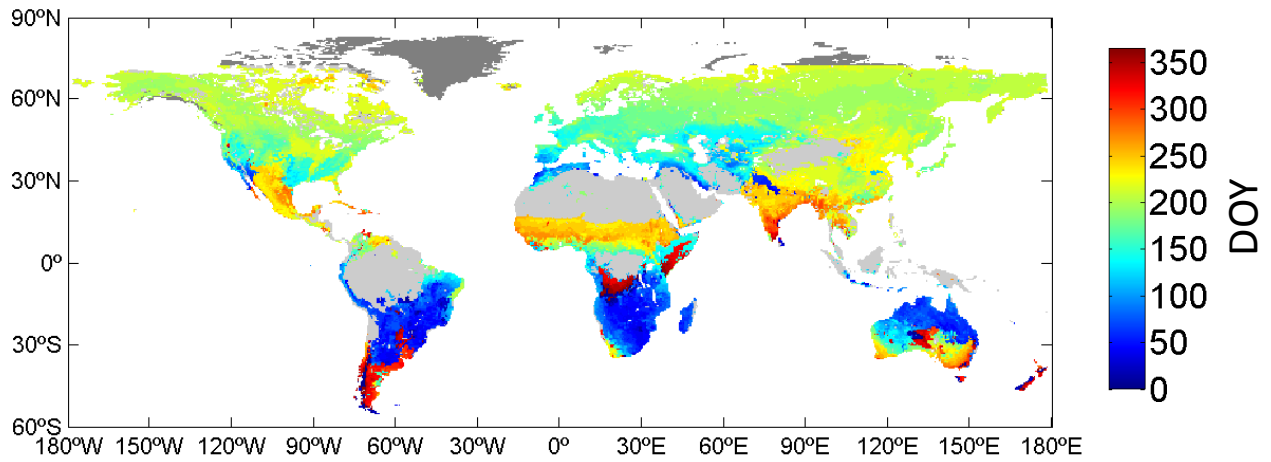
810

811

812

813

814



815

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

820

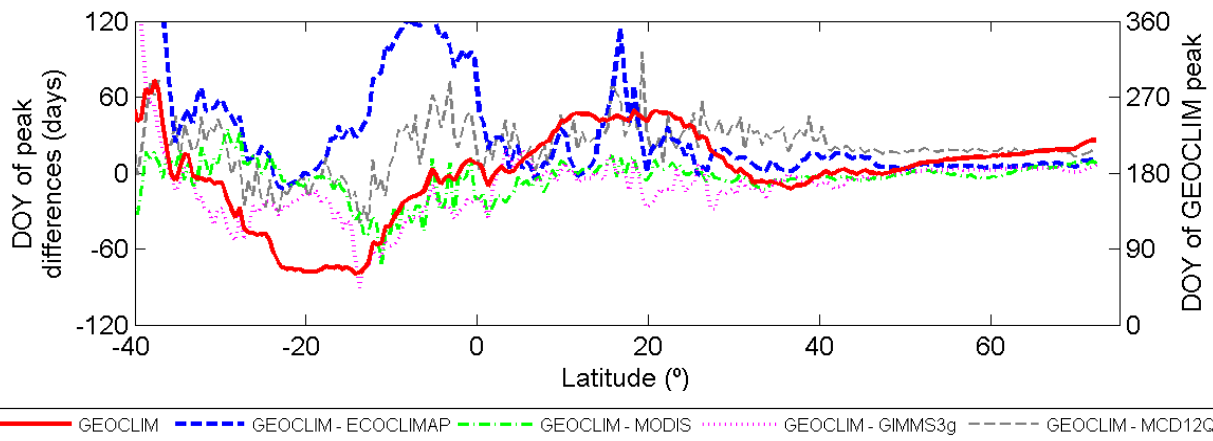
821

822

823

824

825



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

This is the author's version of a work that was accepted for publication in Remote sensing of environment (Elsevier). Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Remote sensing and environment, vol. 166 (set. 2015), 126-137. DOI 10.1016/j.rse.2015.05.027