OPEN ACCESS
Remote Sensing
ISSN 2072-4292
www.mdpi.com/journal/remotesensing

Supplementary Information

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

Zaichun Zhu [1,2,*,†], Jian Bi [1,†], Yaozhong Pan [2], Sangram Ganguly [3], Alessandro Anav [4], Liang Xu [1], Arindam Samanta [5], Shilong Piao [6,7], Ramakrishna R. Nemani [8] and Ranga B. Myneni [1]

[1] Department of Earth and Environment, Boston University, 685 Commonwealth Avenue, Boston, MA 02215, USA; E-Mails: bijian.bj@gmail.com (J.B.); bireme@gmail.com (L.X.); ranga.myneni@gmail.com (R.B.M.)
[2] College of Resources Science & Technology, State Key Laboratory of Earth Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China; E-Mail: pyz@bnu.edu.cn
[3] Bay Area Environmental Research Institute, NASA Ames Research Center, Moffett Field, CA 94035, USA; E-Mail: sangramganguly@gmail.com
[4] College of Engineering, Mathematics & Physical Sciences, Harrison Building, University of Exeter, North Park Road, Exeter EX4 4QF, UK; E-Mail: A.Anav@exeter.ac.uk
[5] Atmospheric and Environmental Research Inc., 131 Hartwell Avenue, Lexington, MA 02421, USA; E-Mail: arindam.sam@gmail.com
[6] Department of Ecology, Peking University, Beijing 100871, China; E-Mail: slpiao@pku.edu.cn
[7] Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100085, China
[8] NASA Advanced Supercomputing Division, NASA Ames Research Center, Moffett Field, CA 94035, USA; E-Mail: rama.nemani@nasa.gov

[†] These authors contributed equally to this work.

[*] Author to whom correspondence should be addressed; E-Mails: zzc@bu.edu or zhu.zaichun@gmail.com; Tel.: +1-617-353-8828; Fax: +1-617-353-8399.

This supplemental material supports the main manuscript with supplementary text and references (Sections S0–S8), figures (Figures S1–S10), and tables (Tables S1–S5).

S0. List of Abbreviations

ANN	Artificial Neural Network

AO	Arctic Oscillation
AVHRR	Advanced Very High Resolution Radiometers
BELMANIP	Benchmark Land Multisite Analysis and Intercomparison of Products
BNU	Beijing Normal University
BU	Boston University
CFs	Canonical Factors
CMIP5	Coupled Model Intercomparison Project 5
CRU	Climatic Research Unit
CYCLOPES	Carbon Cycle and Change in Land Observational Products from an Ensemble of Satellites
ENSO	EI Niño-Southern Oscillation
EOS	Earth Observing System
ESM	Earth System Models
FFNN	Feed-Forward Neural Network
FPAR	Fraction of Photosynthetically Active Radiation absorbed by vegetation
GIMMS	Global Inventory Modeling and Mapping Studies
IGBP	International Geosphere Biosphere Programme
JFM	January to March
LAI	Leaf Area Index
MODIS	Moderate Resolution Imaging Spectroradiometer
mTSF	modified Temporal Spatial Filter
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NDVI3g	Normalized Difference Vegetation Index-third generation
NEX	Earth Exchange
NOAA	National Oceanic and Atmospheric Administration
PCs	Principal Components
RMSE	Root Mean Square Error
SON-1	September to November of the preceding year
SPOT	Système Pour l'Observation de la Terre
TSF	Temporal Spatial Filter

S1. Field LAI Measurements

The field LAI measurements from different field campaigns were used for validation of the LAI3g product (Figure 2). The references related to description of site characteristics are provided in Table A4 of [40]. Detailed methodologies and documentation of field campaigns are available online: http://mercury.ornl.gov/ornldaac/ and http://lpvs.gsfc.nasa.gov/lai_in--tercomp.php.

S2. Earth System Model LAI Simulations

We used LAI simulated by 18 Earth System Models (Figure 7). These new simulations were performed in support of the IPCC Fifth Assessment Report and are referred to as Coupled Model Intercomparison Project phase-5 (CMIP5) [77]. We used monthly-mean model outputs for the last 24

years of the 20th century simulations (1982–2005) because of the overlap with the satellite-based LAI3g product. Information about models used in this study is given in Table S1.

S3. CYCLOPES LAI and FPAR Products

We used CYCLOPES LAI and FPAR data from 1999 to 2007 [66] in our study. The GIMMS and CYCLOPES products were composited to a monthly timescale. Monthly CYCLOPES products were spatially aggregated to 1/12 degree spatial resolution of GIMMS products using a bi-cubic resampling technique. To perform reasonable comparisons, the monthly data were averaged over the period 1999 to 2007. This also significantly reduced residual atmospheric corruptions effects. The comparative analysis reported in the manuscript is based on these values. It should be noted that only data from "L-valid" and "F-valid" pixels (definition in Section S6.1) was used in all the analyses.

CYLOPES products from a select Benchmark Land Multisite Analysis and Intercomparison of Products (BELMANIP) benchmark network of sites [67] were used in the local-scale comparison study. The land cover type of BELMANIP sites was defined using the ECOCLIMAP [S1], which classifies the vegetated surface into seven broad vegetation types.

S4. Climatic Research Unit (CRU) Climate Data

The CRU Time-Series (TS) 3.10.01 temperature and precipitation from the University of East Anglia, UK, for the period January 1982 to December 2009 were used in this study. The data sets provide monthly mean temperature and monthly total precipitation for the whole global at 0.5° by 0.5° spatial resolution. Detailed methodologies and documentation of CRU TS 3.10.01 are available online: http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_1256223773328276.

S5. Remote Sensing Data

S5.1. AVHRR GIMMS NDVI3g

The latest version, third generation, of the GIMMS NDVI data set (NDVI3g) generated from AVHRR sensors onboard a series of NOAA satellites (NOAA 7, 9, 11, 14, 16, 17 and 18) was used in this study. The NDVI3g data set has a spatial resolution of 1/12 degree. The maximum NDVI value during a 15 day period is used to represent each 15 day interval because atmospheric corruption of measured radiances decreases the magnitude of NDVI. This compositing scheme results in two maximum-value NDVI composites per month. The entire available NDVI3g record—July 1981 to December 2011—was used in this study. Only positive NDVI values were used to eliminate snow and other contaminated data (e.g., pixels with large inland water bodies) from all analyses.

S5.2. MODIS Land Cover Map

The MODIS land cover classification product (MCD12C1) identifies 17 IGBP land covers [47]. They include 11 classes of natural vegetation, 3 classes of developed and mosaicked land, and 3 classes of non-vegetated land. The latest version of the MODIS land cover map at 5600m resolution was used in our analyses.

S5.3. MODIS BNU LAI

The MODIS BNU LAI data set was created by post-processing the standard Collection 5 Terra MODIS LAI product (MOD15A2) [47]. An integrated two-step method was developed to generate the improved LAI data sets. First, a modified temporal spatial filter (mTSF), based on a TSF method developed by [S2,S3], was used to gap-fill the LAI product and screen out low-quality data using quality control flags and fill-value information. Second, the TIMESAT Savitzky-Golay filter [S4] was applied to generate the improved MODIS BNU LAI data set [47].

S5.4. MODIS BU FPAR

The MODIS BU FPAR data set was developed by post-processing the standard Collection 5 Terra MODIS FPAR product (MOD15A2) through a two-step process [48]: (1) A 1×1 km^2 8-day FPAR pixel was considered valid when (a) data is of good quality or (b) Clouds are absent. (2) As the 8-day LAI aerosol flag fails to distinguish between average and high aerosol loadings nor reports use of climatological aerosol values, valid 8-day values were averaged to produce 16-day FPAR values whose validity is determined using the 1×1 km^2 16-day VI quality flags corresponding to cloud and aerosol flags: (a) VI data is produced, (b) VI usefulness is between 0 and 11, and (c) Clouds are absent. Valid 1×1 km^2 16-day FPAR values were averaged to obtain monthly FPAR values. Finally, valid 1×1 km^2 monthly values were aggregated to 8×8 km^2 spatial resolution.

S6. Data Preprocessing

AVHRR GIMMS NDVI3g, Terra MODIS BNU LAI and Terra MODIS BU FPAR are the three key input data sets for generating GIMMS LAI3g and FPAR3g data sets. The temporal coverage of the three input data sets are July 1981 to December 2011, February 2000 to December 2010 and February 2000 to December 2009, respectively. The common period of these data sets is therefore February 2000 to December 2009. The following preprocessing of the three key input data sets during the common period was implemented.

The native temporal frequency of AVHRR GIMMS NDVI3g, MODIS BNU LAI and MODIS BU FPAR are 15-day, 8-day and monthly, respectively. Data sets with different temporal frequencies are difficult to utilize in a Feed Forward Neural Network (FFNN) model. Therefore, we composited AVHRR GIMMS NDVI3g and MODIS BNU LAI to a monthly timescale. This has the added advantage of further reducing any residual atmospheric corruption in the data sets.

All the remote sensing data sets (Table S2) are in the same projection system (Geographic Lat/Lon, WGS84). However, the spatial resolutions of the data sets are different. It is critical to resample the data to have the same spatial resolution for each of the key input data sets because the FFNN model is based on a pixel-level training process. Therefore, we resampled the data sets listed in Table S2 to the same spatial resolution as the NDVI3g data set (2,160 rows × 4,320 columns with a pixel size of 0.083 degree). The “nearest neighbor” algorithm was used for resampling the MODIS land cover map, while the bi-cubic resampling was used for all the other data sets.

The following Sections describe in detail the data preprocessing steps. The data were first composited and then resampled to prevent errors due to resampling being transferred to the compositing process.

S6.1. Definition of Valid Pixels

Typical NDVI values range between −0.2 and 0.1 for snow, inland water bodies, deserts, and exposed soils, and increase from about 0.1 to over 0.9 for increasing amounts of vegetation [44]. Thus, we define the valid range for NDVI in our study as 0.1 to 1. Pixels with NDVI value in this range are defined as valid NDVI pixels. Similarly, we define the valid range for LAI as 0.05 to 7 and for FPAR as 0.05 to 1. Only pixels that have valid NDVI, LAI, and FPAR values were used in this study. The following groups of valid pixels are relevant to our study: (a) pixels with valid NDVI values (N-valid pixels), (b) pixels with valid LAI values (L-valid pixels), (c) pixels with valid FPAR values (F-valid pixels), (d) pixels with valid NDVI values and valid LAI values (NL-valid pixels) and (e) pixels with valid NDVI values and valid FPAR values (NF-valid pixels). Obviously, there will be other valid combinations, e.g. NLF-valid pixels, i.e., pixels with valid NDVI values, valid LAI values and valid FPAR values and LF-valid pixels, i.e., pixels with valid LAI values and valid FPAR values. However, these combinations are not relevant because they are not used in our study. Type (a) pixels were used in the process of generating GIMMS LAI3g and FPAR3g. Type (b) and (c) pixels were used in preprocessing of MODIS BNU LAI and MODIS BU FPAR, respectively. Type (d) and (e) pixels were used in the process of training the FFNN models.

S6.2. Creation of Multi-Year Average Monthly AVHRR GIMMS NDVI3g

We preprocessed the NDVI3g in three steps: (a) generated monthly data by averaging the 15-day composites for each month and (b) further averaged these over the same months during the common period (February 2000 to December 2009). No resampling was needed as 1/12 degree is the benchmark spatial resolution of all the data sets. It should be noted that only "N-valid pixels" (Section S6.1) were used in the preprocessing.

S6.3. Creation of Multi-Year Average Monthly MODIS BNU LAI

A four-step compositing method was used to create monthly MODIS BNU LAI: (a) composite the 8-day data to 15-day using an accurate date-weighted compositing method, (b) composite the 15-day data to monthly values by averaging the two 15-day composites in each month, (c) evaluate multi-year average monthly MODIS BNU LAI data set by averaging the monthly composites during the common period, February 2000 to December 2009, and (d) resample the resulting 1km data to 1/12 degree. It should be noted that only "L-valid pixels" (Section S6.1) were used in the four steps. The following is a more detailed explanation of step one.

Figure S1 is an illustration of how we calculated the weight for each 8-day composite. We separated the date line into three parts using four points. "B" and "D" are the start and end days of the half month, "A" is the date seven days before "B", "C" is the date seven days before "D". The native temporal resolution of MODIS BNU LAI is 8 days. The starting dates of MODIS BNU LAI in a year are:

$$Dates = 1:8:361 \quad \text{(S1)}$$

For non-leap years, we obtain the bi-weekly break points as in Table S3. For leap years, we obtain the bi-weekly break points as in Table S4. Data with starting dates before "A" or later than "D" are not

taken into account in estimating the LAI for the 15-day period in question. Different weights are assigned to data with starting dates (Equation (S1)) in the range of [A, B), [B, C], and (C, D] separately:

$$\begin{cases} w(i) = 8 - [B - d(i)] & d(i) \in [A, B) \\ w(i) = 8 & d(i) \in [B, C] \\ w(i) = D - d(i) + 1 & d(i) \in (C, D] \end{cases} \quad \text{(S2)}$$

Here $w(i)$ and $d(i)$ are the assigned weight and date of 8-day composite i, respectively. Assuming that there are n_c composites in the range of [A, D], we use the following method to create the LAI composite, pixel-by-pixel, for the 15-day period:

$$LAI_{15-day}(c) = \sum_{i=1}^{n_c} LAI_{8-day}(i) \cdot w(i) \Big/ \sum_{i=1}^{n_c} w(i) \quad \text{(S3)}$$

Here $\mathrm{LAI}_{15-day}(c)$ is the LAI value of the pixel for the *cth* 15-day composite, $LAI_{8-day}(i)$ is the LAI value of the pixel of 8-day composite i, and $w(i)$ is the corresponding weight (Equation (S2)), n_c is the number of 8-day composites that fall in the date range of *cth* 15-day composite. Equation (S3) was used to create the 15-day MODIS BNU LAI composites during the period February 2000 to December 2009.

S6.4. Creation of multi-year average monthly MODIS BU FPAR

The only required preprocessing is multi-year averaging and resampling as the native temporal frequency of MODIS BU FPAR is monthly. First, the data are averaged over the corresponding months during the common period February 2000 to December 2009. The resulting data are then resampled to the common spatial resolution (1/12 degree). Only "F-valid pixels" (Section S6.1) are taken into account.

S7. Feed Forward Back-Propagation Neural Network (FFNN)

S7.1. Preparation of Training Data

A four-step data processing was performed to create more training data for each month. The following description uses LAI as an example—the processing for FPAR is identical.

(1) Set a window size as M × N: M is the number of rows; N is the number of columns. The total pixel number in the window is M × N.

(2) Create a table with five columns: NDVI, land cover, latitude, longitude and LAI. The table will be used to record the entries for each window in Step (1).

(3) For each pixel in the window:

a. Tag vegetated (land covers shown in Table S5) "NL-valid pixels" (Section S6.1) as "valid".
b. Count the total number of pixels for each land cover.
c. Set a threshold *t*. If the number of valid pixels of a particular land cover is greater than $M \times N \times t$, calculate the mean NDVI value and mean LAI value.
d. Calculate the central latitude and longitude for the window.
e. Record these variables in the table created in Step (2) above.

(4) Move the window across the entire global land surface and repeat the calculations described in Step (3).

Here is an example for Step (3). Assume that there are four land covers in a window but only three of which are vegetated land covers: class A, class B and class C (Figure S2). We denote the vegetated "NL-valid pixels" (Section S6.1) as valid. We then count the number of valid pixels for each land cover class. If we set the threshold value of t as 0.2, then only Class A and Class B will be taken into account. Then, we calculate their mean NDVI and mean LAI. Further we calculate the central latitude and central longitude of this window.

We create a table for each month of the year using the process described above. Thus, twelve tables (LAI tables) were created to train the FFNN models for generating LAI3g product. Similarly, another twelve tables (FPAR tables) were created to train the FFNN models for generating the FPAR3g product. In our study, window size was set to 10 × 10 pixels and the threshold t was set to 0.1.

S7.2. Training the FFNN Model

Artificial neural networks (ANN) have been used in a diverse range of research fields [S5–S10]. ANNs have the ability to reveal unknown relationships underlying the data and require no initial assumptions. A neural network consists of: (1) the transfer model of a single neuron, (2) the topological connections between these simple neurons, and (3) the weight or strengths of each connection between the neurons.

A widely used neural network named Feed Forward Neural Network (FFNN) generally consists of three layers: input layer, hidden layer and output layer. The neurons are connected from one layer to the next but not within the same layer. Signals always flow from the input layer to the hidden layer and from the hidden layer to the output layer through the unidirectional connections. FFNN can approximate any continuous function to an arbitrary precision [S11]. However, a problem called over-fitting might occur during neural network training process. However, there is little or no chance of over-fitting if the number of training data is much larger than the number of parameters in the neural network, which is the case in our study.

In this study, we built a FFNN model with three layers: input layer, hidden layer and output layer (Figure S3). The numbers of neurons of input layer and output layer are four and one. The four neurons in the input layer transfer the signals from the input vectors (GIMMS NDVI3g, land cover, window-center latitude and window-center longitude) to the hidden layer. The one neuron in the output layer receives the signals sent by the hidden layer and predicts LAI3g (or, FPAR3g). Unlike the input layer and the output layer, we cannot arbitrarily determine the number of neurons in the hidden layer. Generally, larger numbers of neurons in the hidden layer give the network more flexibility. However, too many neurons in the hidden layer requires more computational resources and may cause the problem of over-fitting. Through iteration, we decided to create a hidden layer with eleven neurons.

To prevent over-fitting and test the performance of the FFNN, the training data set was split into three sets: 70% as training data, 15% as validation data and 15% as test data. The network was trained with training data until its performance began to decrease on the validation data, which means that generalization has peaked. The test data provided a completely independent test of network generalization. The performance function to evaluate the network training process is mean square error

which is the average squared error between the network outputs and the target outputs [S12]. Ten networks with different initial values were trained independently. The network providing the best performance was selected as the final FFNN model that was used for generating GIMMS LAI3g and FPAR3g data sets.

Following the above training process, we trained 12 FFNN models for generating LAI3g using the tables prepared in Section S7.1. Similarly, we also trained 12 FFNN models for generating FPAR3g data set.

S8. Analysis of the Impact of Mount Pinatubo Eruption and NOAA 11 Orbital Drift on Identifying Dominant Modes of Interannual Variability

The impact of Mount Pinatubo eruption in mid-1991 and significant orbit loss of NOAA 11 is clearly visible in the time series in Figure 1, especially in the tropics and in the forested regions of the globe. This section intends to analyze the impact of Mount Pinatubo eruption and orbit loss of NOAA 11 on identifying dominant modes of interannual variability in Section 3.4. There are two ways to further investigate the problem: (1) performing the CCA analysis with the time series that affected by the Mount Pinatubo eruption and NOAA 11 orbital drift (1989 to 1994), and (2) performing the CCA analysis with the full time series except for the period of Mount Pinatubo eruption and NOAA 11 orbital drift. However, the first method may provide unreliable results because the six-year time series (1989 to 1994) is too short to perform statistical analysis. Thus, we used the second method to investigate the potential effects of the disturbances. The LAI3g and FPAR3g for the period that NOAA 11 satellite provides the GIMMS NDVI3g data and Mount Pinatubo erupted (January1989 to December 1994) were truncated. We combined the data sets in two discontinuous periods (January 1982 to December 1988 and January 1995 to December 2009) to obtain a new time series that are not affected by Mount Pinatubo eruption and satellite orbital drift.

The Mount Pinatubo eruption and orbital drift mainly affected the tropical region between 1989 and 1994 (Figure 1). We can hardly detect any strong suspicious coincidence with the orbital drift of the NOAA 11 satellite in other regions. Thus, we decided to do this analysis for the latitude band −40°S to 40°N. From Section 3.4 in the Manuscript we know the 1st CCA factor from our products and precipitation is strongly correlated with NINO3 index. Also, we can see there were several noticeable ENSO events (the shadings) during the excluded period. It is very likely that we may get lower correlation coefficients if we drop this period. However, if the new coefficients are greater than the coefficients in Figure 6(c) and Figure S9(c), we can conclude that the Mount Pinatubo eruption and orbital drift of NOAA 11 did influence the CCA analysis in Section 3.4 and the Mount Pinatubo eruption and orbital drift effects cannot be well explained by the ENSO events during that period. Thus, we decided to do the CCA analysis of the truncated LAI3g/FPAR3g and precipitation data records and calculate the correlation between the 1st CCA factor and NINO3 index.

We performed the CCA analysis between the truncated LAI3g and FPAR3g and precipitation, and evaluated the correlation between the 1st CCA factor and NINO3 index in the latitude band −40°S to 40°N. The results are shown in Figure S10. Comparing Figure S10(a) and Figure 6(c), the correlation coefficient between precipitation CF1 and NINO3 (SON-1) increased slightly (0.53 *vs*. 0.50). The correlation coefficient between LAI CF1 and NINO3 (SON-1) similarly increased slightly (0.52 *vs*.

0.48). Comparing Figure S10(b) and Figure S9(c), the correlation coefficient between precipitation CF1 and NINO3 (SON-1) increased slightly (0.51 *vs*. 0.44). The correlation coefficient between FPAR CF1 and NINO3 (SON-1) increased more (0.57 *vs*. 0.45).

Thus, we may conclude that the Mount Pinatubo eruption and orbital drift of NOAA 11 does influence the CCA analysis in Section 3.4 and that the orbital drift effects cannot be well explained by the ENSO events during that period. However, the magnitude of these effects remains unknown.

S10. Figures S1–S10

Figure S1. Method to calculate the weight while compositing 8-day MODIS BNU LAI data to a bi-monthly LAI composite as described in Section S6.3. B and D are the start day and end day of the month, A and C are 7 days lagging B and D.

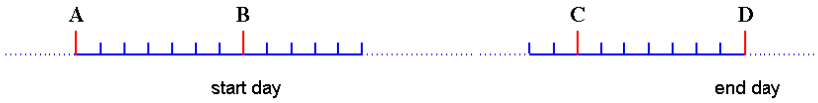


Figure S2. Schematic illustration of how to calculate mean values of NDVI, LAI, and FPAR in a sub-image. Assume that there are four land covers in the window: Class A, Class B and Class C are IGBP land covers considered in this study, Class D is non-vegetated. Pixels with valid data are shown as "O" and pixels with invalid data are shown as "×".

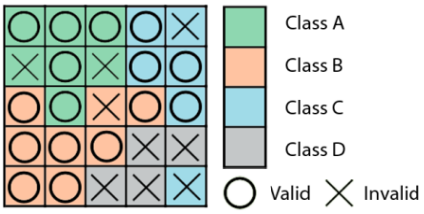


Figure S3. Structure of the Feed Forward Neural Network (FFNN) used in this study. The FFNN consists of a four-neuron input layer, an eleven-neuron hidden layer and a one-neuron output layer. GIMMS NDVI3g, MODIS IGBP land covers, pixel latitude and pixel longitude are input parameters. The output of the FFNN is the predicted LAI (or FPAR) value.

Figure S4. Comparison of monthly FPAR values from CYCLOPES and FPAR3g data sets for four broad vegetation classes (forests (**a**), other woody vegetation (**b**), herbaceous vegetation (**c**) and cropland/natural vegetation mosaics) for the period 1999 to 2007. These classes are groups of IGBP land cover types as per Table S5.

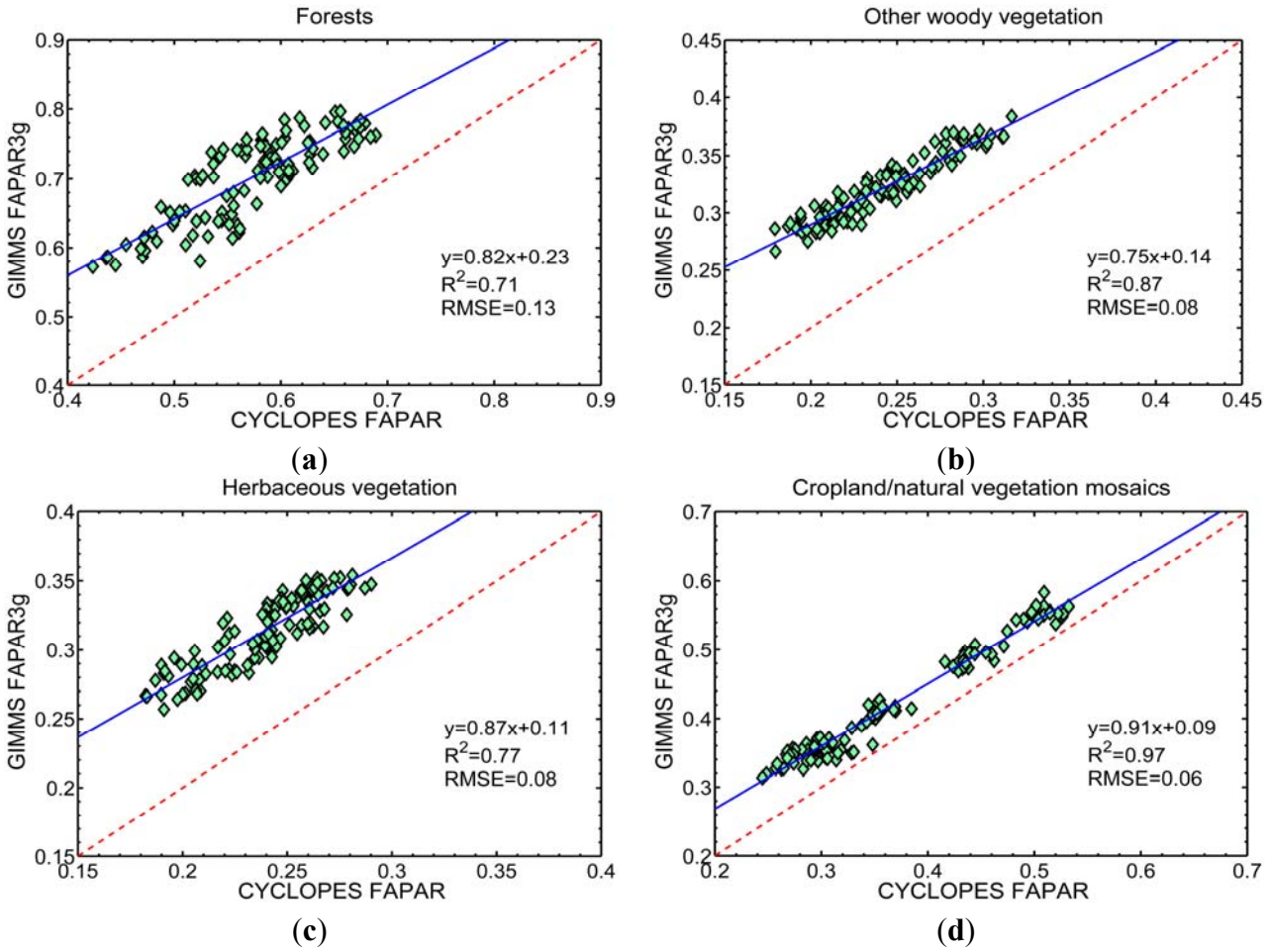


Figure 5. Comparison of monthly LAI values from CYCLOPES and LAI3g data sets for four broad vegetation classes (forests (**a**), other woody vegetation (**b**), herbaceous vegetation (**c**) and cropland/natural vegetation mosaics). These classes are groups of IGBP land cover types as per Table S5. The plotted data are monthly averages for the years 1999 to 2007.

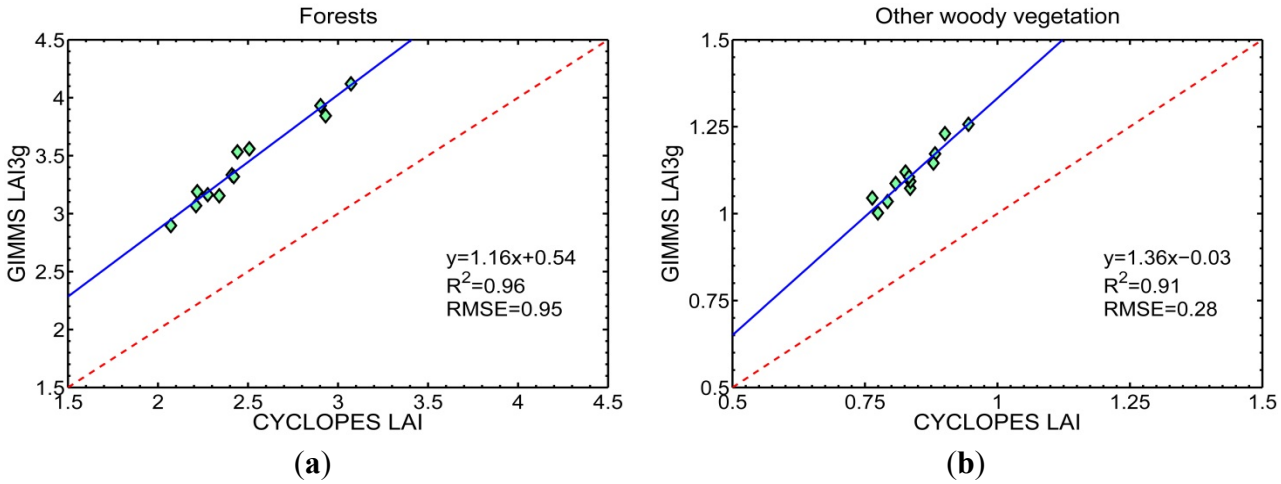


Figure S5. *Cont.*

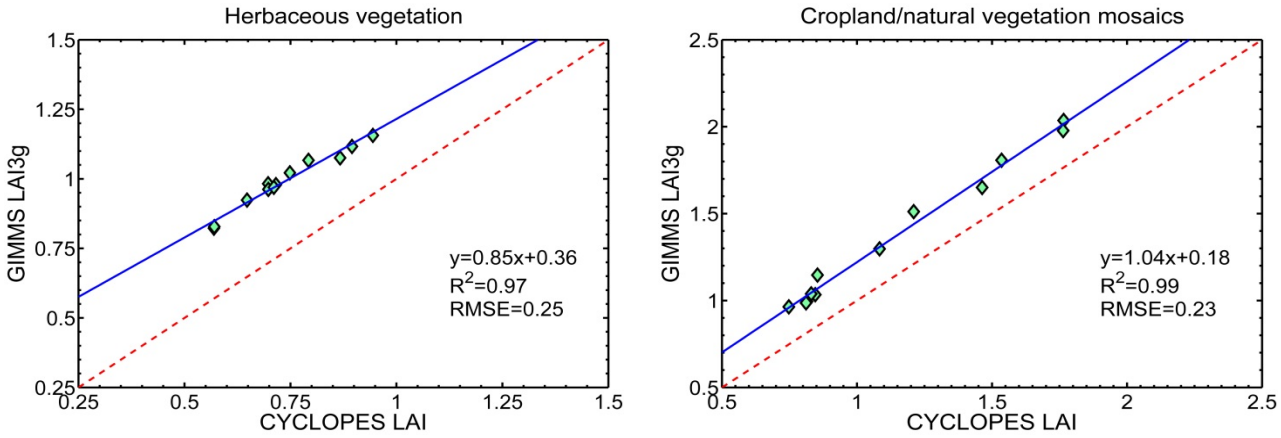


Figure S6. Comparison of monthly FPAR values from CYCLOPES and FPAR3g data sets for four broad vegetation classes (forests (**a**), other woody vegetation (**b**), herbaceous vegetation (**c**) and cropland/natural vegetation mosaics). These classes are groups of IGBP land cover types as per Table S5. The plotted data are monthly averages for the years 1999 to 2007.

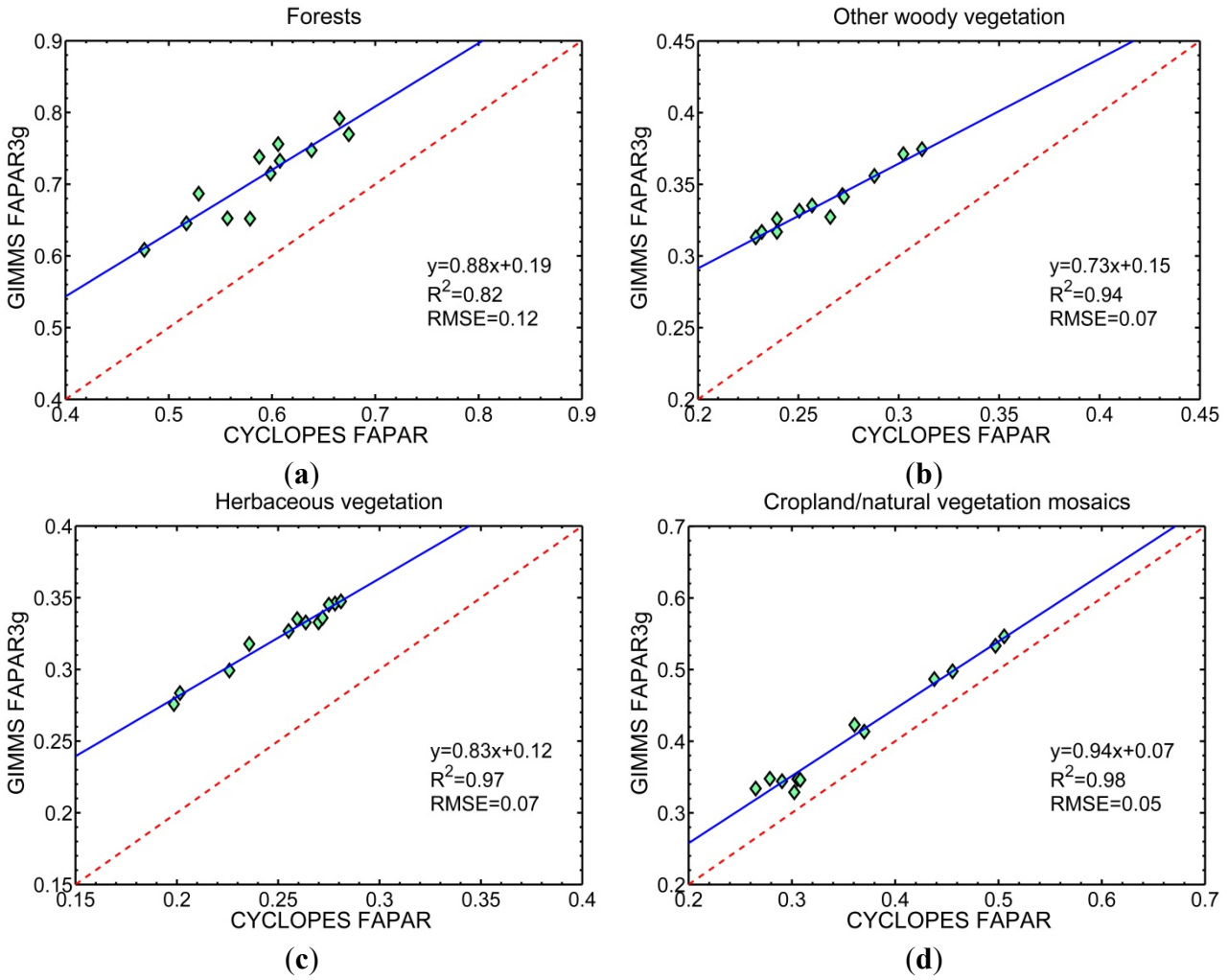


Figure S7. Density scatter plots of monthly FPAR3g and CYCLOPES FPAR for 323 BELMANIP sites for the time period from 1999 to 2007. The plots show correlation between the two products for four broad groups of vegetation which are grouping of the IGBP land covers (Table S5). The black dash line is the 1:1 line. The solid black lines are regression lines derived from the scatter plot.

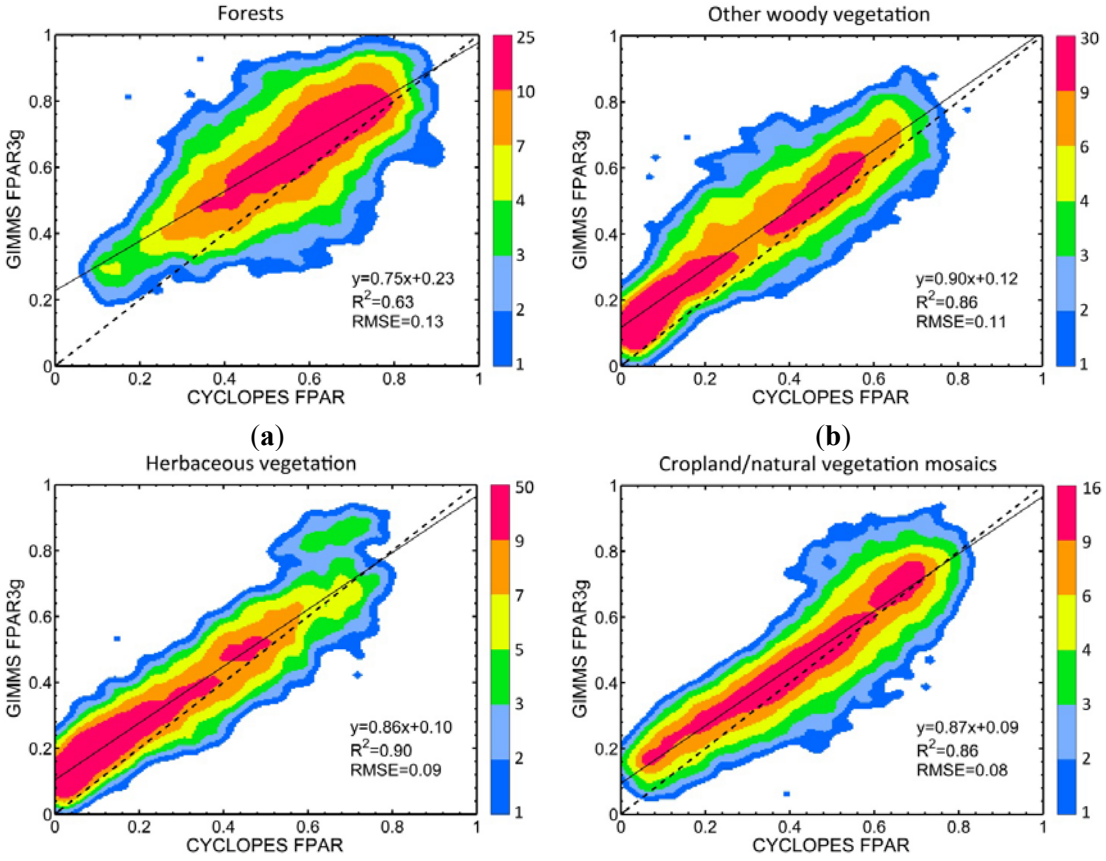


Figure S8. Statistical evaluation of FPAR3g with temperature in the northern latitudes and precipitation in the tropical regions. (**a**) Statistical analyses between approximate growing season (May to September) averages of FPAR3g and surface temperature in the northern latitudes (50°N–90°N) for the overlapping period of the two data sets (1982 to 2009). The inset in (a) shows temporal variations of standardized anomalies of growing season averages of FPAR3g and temperature. (**b**) Correlation between annual mean FPAR3g and annual total precipitation in the tropical latitudes (23°S–23°N).

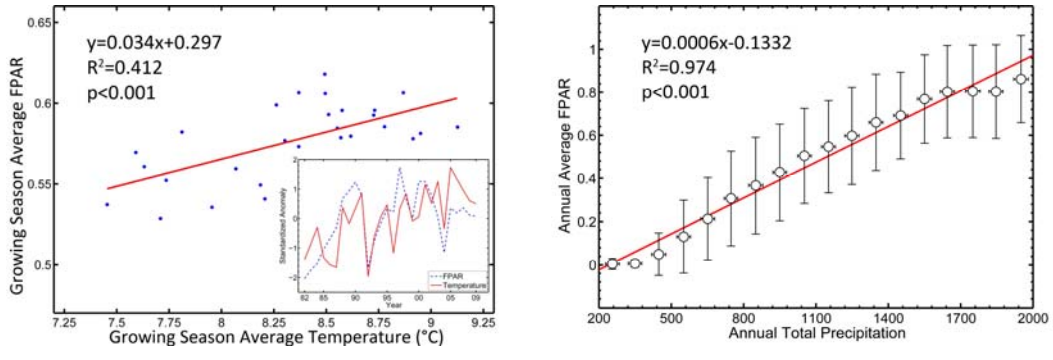


Figure S9. Correlations between the standardized time series of the first and second canonical factors (CF1 and CF2) of land surface temperature, precipitation and FPAR3g with NINO3 and AO indices in the northern (10°N to 90°N) and tropical/extra-tropical regions (40°S to 40°N) for the period 1982 to 2009. The standardized September through November average NINO3 index time series of the preceding year and the January through March average AO index are shown in these plots as black dash lines.

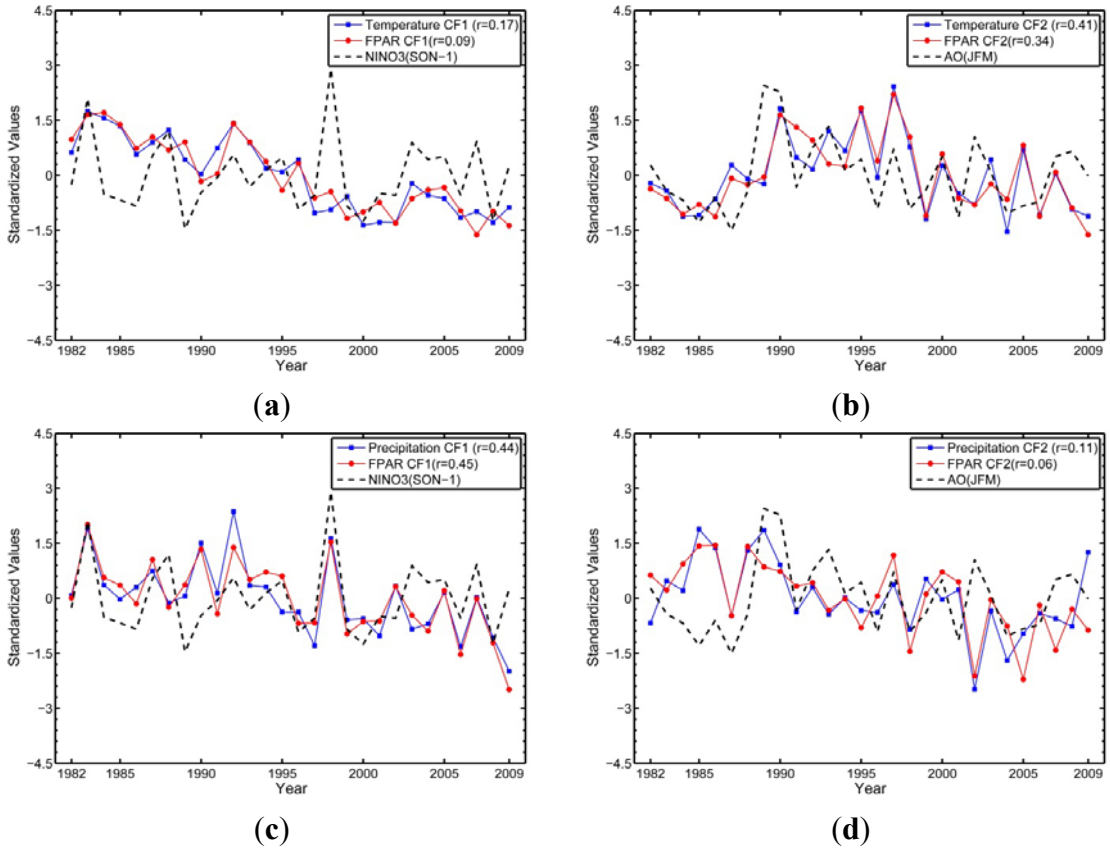
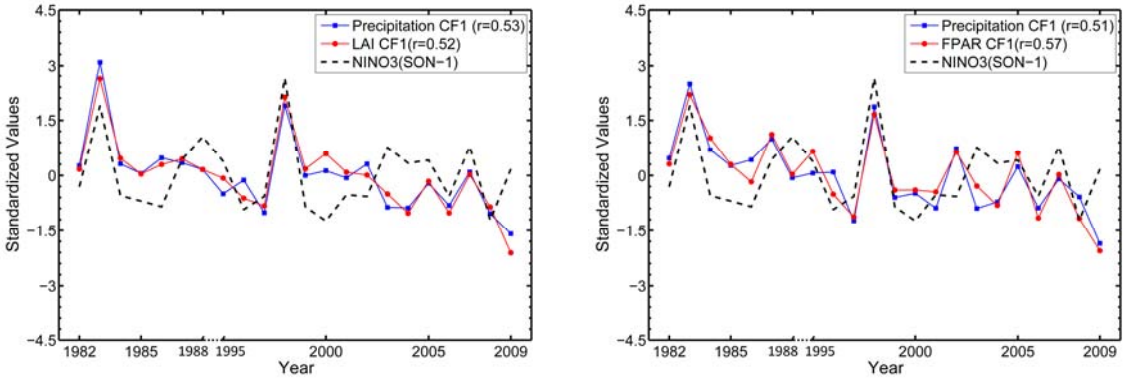


Figure S10. Correlations between the standardized time series of the first canonical factors (CF1) of precipitation and (**a**) LAI3g, (**b**) FPAR3g and NINO3 index in the tropical/extra-tropical regions (40°S to 40°N) for the discontinuous periods 1982 to 1988 and 1995 to 2009. The standardized September through November average NINO3 index time series of the preceding year is shown in these plots as black dash lines.



S11. Tables S1-S5

Table S1. CMIP5 ESM simulations of LAI used in this article (Figure 7).

MODELS	Source	Spatial Resolution (lon × lat)
BCC-CSM1	Beijing Climate Center, China Meteorological Administration, China	2.813° × 2.813°
CanESM2	Canadian Centre for Climate Modelling and Analysis, Canada	2.813°× 2.813°
CCSM4	National Center for Atmospheric Research, United States	1.250° × 0.938°
GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory, United States	2.500° × 2.000°
GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory, United States	2.500° × 2.000°
HadGEM2-CC	Met Office Hadley Centre, UK	1.875° × 1.241°
HadGEM2-ES	Met Office Hadley Centre, UK	1.875° × 1.241°
INMCM4	Institute for Numerical Mathematics, Russia	2.000° × 1.500°
IPSL-CM5A-LR	Institut Pierre Simon Laplace, France	3.750° × 1.875°
IPSL-CM5A-MR	Institut Pierre Simon Laplace, France	2.500° × 1.259°
IPSL-CM5B-LR	Institut Pierre Simon Laplace, France	3.750° × 1.875°
MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, Japan; Atmosphere and Ocean Research Institute, Japan; National Institute for Environmental Studies, Japan	2.813° × 2.813°
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Japan; Atmosphere and Ocean Research Institute, Japan; National Institute for Environmental Studies, Japan	2.813° × 2.813°
MPI-ESM-LR	Max Planck Institute for Meteorology, Germany	1.875°× 1.875°
MPI-ESM-P	Max Planck Institute for Meteorology, Germany	1.875° × 1.875°
MPI-ESM-MR	Max Planck Institute for Meteorology, Germany	1.875° × 1.875°
NorESM1-M	Norwegian Climate Centre, Norway	2.500° × 1.875°
NorESM1-ME	Norwegian Climate Centre, Norway	2.500° × 1.875°

Table S2. Remote sensing data sets used for generating the GIMMS LAI3g and FPAR3g data sets.

Data Name	AVHRR GIMMS NDVI3g	MODIS BNU LAI	MODIS BU FPAR	MODIS Land Cover
Area	Global	Global	Global	Global
Projection	Lat/Lon	Lat/Lon	Lat/Lon	Lat/Lon
Spatial Resolution	0.083 degree	1 km	0.072 degree	0.05 degree
Temporal resolution	15-day	8-day	monthly	yearly
Temporal Coverage	Jul 1981–Dec 2011	Feb 2000–Dec 2009	Feb 2000–Dec 2010	2007
Data Type	Int16	Uint8	Unint8	Uint8
Data Format	HDF	NETCDF	Binary	HDF

Table S3. Bi-weekly break points for non-leap years. We separate each month into two parts: (1) first half of the month, from the 1st to 15th of a month; (2) second half of the month, from the 16th to the end of that month. All the break points are presented as Julian days.

Month	First Half Month		Second Half Month	
	Start Day	End Day	Start Day	End Day
January	1	15	16	31
February	32	46	47	59
March	60	74	75	90
April	91	105	106	120
May	121	135	136	151
June	152	166	167	181
July	182	196	197	212
August	213	227	228	243
September	244	258	259	273
October	274	288	289	304
November	305	319	320	334
December	335	349	350	365

Table S4. Bi-week break points for leap years. We separate each month into two parts: (1) first half of the month, from the 1st to 15th of a month; (2) second half of the month, from the 16th to the end of that month. All the break points are presented as Julian days.

Month	First Half Month		Second Half Month	
	Start Day	End Day	Start Day	End Day
January	1	15	16	31
February	32	46	47	60
March	61	75	76	91
April	92	106	107	121
May	122	136	137	152
June	153	167	168	182
July	183	197	198	213
August	214	228	229	244
September	245	259	260	274
October	275	289	290	305
November	306	320	321	335
December	336	350	351	366

Table S5. Cross-walk table from IGBP vegetation classes [50] to 4 broad vegetation classes used in some of our analysis.

IGBP Class	Vegetation Names	Dominant Classes
1	Evergreen needleleaf forests	Forests
2	Evergreen broadleaf forests	
3	Deciduous needleleaf forests	
4	Deciduous broadleaf forests	
5	Mixed forests	
6	Closed shrublands	Other woody vegetation
7	Open shrublands	
8	Woody savannas	

Table S5. *Cont.*

IGBP Class	Vegetation Names	Dominant Classes
9	Savannas	Herbaceous vegetation
10	Grasslands	
12	Croplands	Cropland/natural vegetation mosaics
14	Cropland/natural vegetation mosaics	

References

S1. Masson, V.; Champeaux, J.-L.; Chauvin, F.; Meriguet, C.; Lacaze, R. A global database of land surface parameters at 1-km resolution in meteorological and climate models. *J. Climate* **2003**, *16*, 1261–1282.

S2. Fang, H.; Liang, S.; Kim, H.-Y.; Townshend, J.R.; Schaaf, C.L.; Strahler, A.H.; Dickinson, R.E. Developing a spatially continuous 1 km surface albedo data set over North America from Terra MODIS products. *J. Geophys. Res.* **2007**, *112*, D20206.

S3. Fang, H.; Liang, S.; Townshend, J.R.; Dickinson, R.E. Spatially and temporally continuous LAI data sets based on an integrated filtering method: Examples from North America. *Remote Sens. Environ.* **2008**, *112*, 75–93.

S4. Jönsson, P.; Eklundh, L. TIMESAT—A program for analyzing time-series of satellite sensor data. *Comput. Geosci.* **2004**, *30*, 833–845.

S5. Heermann, P.D.; Khazenie, N. Classification of multispectral remote sensing data using a back-propagation neural network. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 81–88.

S6. Hall, L.O.; Bensaid, A.M.; Clarke, L.P.; Velthuizen, R.P.; Silbiger, M.S.; Bezdek, J.C. A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain. *IEEE Trans. Neural Netw.* **1992**, *3*, 672–682.

S7. Gopal, S.; Woodcock, C. Remote sensing of forest change using artificial neural networks. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 398–404.

S8. Chang, D.-H.; Islam, S. Estimation of soil physical properties using remote sensing and artificial neural network. *Remote Sens. Environ.* **2000**, *74*, 534–544.

S9. Yuan, H.; van Der Wiele, C.F.; Khorram, S. An automated artificial neural network system for land use/land cover classification from Landsat TM imagery. *Remote Sens.* **2009**, *1*, 243–265.

S10. Veronez, M.R.; Florêncio de Souza, S.; Matsuoka, M.T.; Reinhardt, A.; Macedônio da Silva, R. Regional mapping of the geoid using GNSS (GPS) measurements and an artificial neural network. *Remote Sens.* **2011**, *3*, 668–683.

S11. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control Singnal* **1992**, *5*, 455–455.

S12. Beale, M.H.; Hagan, M.T.; Demuth, H.B. *Neural Network Toolbox User's Guide*, R2012b ed.; The Math Works: Natick, MA, USA, 2012; pp. 2–15.