Received: 31 January 2019 | Revised: 11 December 2019 | Accepted: 20 December 2019

DOI: 10.1002/gdj3.87

DATA PAPER

Geoscience Data Journal | RMetS | WILEY

Leaf phenology amplitude derived from MODIS NDVI and EVI: Maps of leaf phenology synchrony for Meso- and South America

France F. Gerard[1] | **Charles T. George**[1] | **Garry Hayman**[1] | **Cecilia Chavana-Bryant**[2,3] | **Graham P. Weedon**[4]

[1]Climate System Group, UK Centre for Ecology and Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, UK

[2]Climate & Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

[3]Department of Environmental Science, Policy and Management, UC Berkeley, Berkeley, CA, USA

[4]JCHMR, Met Office, Maclean Building, Crowmarsh Gifford, Wallingford, UK

Correspondence
France F. Gerard, UK CEH Wallingford, Maclean Building, Crowmarsh Gifford, Wallingford, OX10 8BB, UK.
Email: ffg@ceh.ac.uk

Funding information
This article was funded by NERC national capability funding and the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 283093 - The Role of Biodiversity In climate change mitigatioN (ROBIN).

Abstract

The leaf phenology (i.e. the seasonality of leaf amount and leaf demography) of ecosystems can be characterized through the use of Earth observation data using a variety of different approaches. The most common approach is to derive time series of vegetation indices (VIs) which are related to the temporal evolution of FPAR, LAI and GPP or alternatively used to derive phenology metrics that quantify the growing season. The product presented here shows a map of average 'amplitude' (i.e. maximum minus minimum) of annual cycles observed in MODIS-derived NDVI and EVI from 2000 to 2013 for Meso- and South America. It is a robust determination of the amplitude of annual cycles of vegetation greenness derived from a Lomb–Scargle spectral analysis of unevenly spaced data. VI time series pre-processing was used to eliminate measurement outliers, and the outputs of the spectral analysis were screened for statistically significant annual signals. Amplitude maps provide an indication of net ecosystem phenology since the satellite observations integrate the greenness variations across the plant individuals within each pixel. The average amplitude values can be interpreted as indicating the degree to which the leaf life cycles of individual plants and species are synchronized. Areas without statistically significant annual variations in greenness may still consist of individuals that show a well-defined annual leaf phenology. In such cases, the timing of the phenology events will vary strongly within the year between individuals. Alternatively, such areas may consist mainly of plants with leaf turnover strategies that maintain a constant canopy

Dataset

Identifier: https://doi.org/10.5285/dae416b4-3762-45bd-ae14-c554883d482c

Creator: George, C. T., Hayman, G., Weedon, G. P., and Gerard, F. F., CEH

Title: CEH Leaf Phenology Amplitude for Meso-and South America

Publisher: EIDC

Publication year: 2018

Resource type: GEOTIFF

Version: 2.0

of leaves of different ages. Comparison with in situ observations confirms our interpretation of the average amplitude measure. VI amplitude interpreted as leaf life cycle synchrony can support model evaluation by informing on the likely leaf turn over rates and seasonal variation in ecosystem leaf age distribution.

KEYWORDS

leaf phenology, Lomb–Scargle spectral analysis, MODIS time series, synchrony, vegetation indices

1 | INTRODUCTION

The characterization of the leaf phenology of ecosystems and the plant functional types (PFTs) within them is important for the identification of spatial phenological patterns and their drivers. This information is essential to improve dynamic global vegetation models through better representation and parameterization of PFTs. This is particularly relevant in the tropical and subtropical regions where observed leaf phenological behaviours range from (a) rapid shedding of short-lived, leaves early in the dry season in deciduous species; (b) irregular exchange in brevi-deciduous species to (c) constant exchange of long-lived, leaves in evergreen species, and where both intraspecific synchrony and asynchrony have been reported (van Schaik *et al.*, 1993; Wright and Schaik, 1994; Reich, 1995; Reich *et al.*, 2004).

A variety of different Earth observation (EO) approaches have been used to characterize the leaf phenology of ecosystems and their PFTs. A common approach is to derive time series of vegetation indices (normalized difference vegetation index, NDVI and enhanced vegetation index, EVI) which can be related to the temporal evolution of the fraction of absorbed photosynthetically active radiation (fAPAR), leaf area index (LAI) and gross primary productivity (GPP) (Justice *et al.*, 1998; Fensholt *et al.*, 2004; Vina and Gitelson, 2005; Yan *et al.*, 2008; Zhu *et al.*, 2013). EO-derived land cover maps often include 'evergreen' and 'deciduous' classes (GlobCover, Arino *et al.*, 2012, e.g. MODIS land cover, Friedl *et al.*, 2010). However, these classes are limited to a qualitative description of generic phenological behaviours. Another approach is to characterize leaf phenology through combining EO-derived phenology metrics, such as start (leaf out), end (leaf fall) and length of growing season, and the magnitude of greening (i.e. the range or maximum minus minimum, commonly referred to in the phenology literature as 'amplitude') from time series of vegetation indices (Ivits *et al.*, 2013; Eastman *et al.*, 2013; Buitenwerf *et al.*, 2015). Such metrics will often show high levels of uncertainty (White *et al.*, 2009), particularly in the tropics, if derived from EO data where the incidence of cloud is high and/or the growing season is less well-defined.

Here we have implemented an approach that deals with the noisy time series and missing observations in vegetation indices, prevalent in these regions, and that robustly tests for the presence of statistically significant annual leaf phenology signals. We present maps of average ranges (described as average 'amplitudes') of vegetation indices for Meso- and South America, representing the synchrony of vegetation greening within 500 m^2, 1 km^2 or 0.5 $degree^2$ pixels or grid boxes. Where the uncertainty is too high, no data are provided. These maps can be used as reference phenology maps by the DGVM modelling community, which is currently concerned with the development of improved PFTs (e.g. Arora and Boer, 2005; De Weirdt *et al.*, 2012; Zhao *et al.*, 2013; Sakschewski *et al.*, 2016) in the tropical and subtropical regions, and to inform on the likely seasonal variation in leaf age distribution to help improve model GPP estimates (Wu *et al.*, 2016; Albert *et al.*, 2018).

2 | METHODS

The approach is based on a method we developed as part of (Bradley *et al.*, 2011) with improvements introduced at different stages of the data processing. Important differences in the method, which is outlined below, are (a) the use of 14 years of vegetation index (VI) data compared with 6 years in (Bradley *et al.*, 2011) which resulted in an much better estimate of the amplitude; (b) the application of a stringent quality control procedure to select the most reliable VI observations; and (c) the use of the Lomb–Scargle transform (Press *et al.*, 1992), thus avoiding the need for time series gap filling that can bias the detection of annual cycles.

Quality-assured EVI and NDVI time series were derived from the 14-year-long (2000–2013) 16-day time series of MODIS vegetation indices (500 and 1,000 m EVI; and 500 and 1,000 m NDVI). These input time series were acquired from LPDAAC and are referred to as MOD13A1 (500 m^2) and MOD13A2 (1,000 m^2).

Both NDVI and EVI are an indirect measure of vegetation growth (i.e. increase in LAI) and activity (i.e. photosynthesis). NDVI is a ratio of the reflectance from the red and near infrared (NIR) wavebands and so reduces noise caused by sources present in both bands, such as illumination differences and atmospheric attenuation. However, NDVI is nonlinear and saturates at high vegetation biomass levels while at low vegetation levels it is sensitive to variations in canopy

background reflectance (e.g. soil, litter) (Huete, 1988). The MODIS EVI product was designed to reduce these problems and also correct for atmospheric effects (Huete *et al.*, 2002). NDVI is more sensitive to leaf chlorophyll concentrations through the red band (620–670 nm) while EVI is more sensitive to canopy structure, including LAI, through the NIR band (840–875 nm) (Gao *et al.*, 2000).

The NDVI equation is:

$$\mathrm{NDVI} = (\rho_{\mathrm{NIR}} - \rho_{\mathrm{red}})/(\rho_{\mathrm{NIR}} + \rho_{\mathrm{red}})$$

with values ranging between −1 and 1.

The EVI equation is:

$$\mathrm{EVI} = G\frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{red}}}{\rho_{\mathrm{NIR}} + C_1 * \rho_{\mathrm{red}} - C_2 * \rho_{\mathrm{blue}} + L}$$

where ρ are atmospherically corrected reflectance values; G is the gain factor; L is the canopy background adjustment; and C_1, C_2 are the coefficients of the aerosol resistance term, which uses the blue band reflectance to correct for aerosol influences on the reflectance in the red band. The coefficients adopted in EVI MODIS are $L = 1$, $C_1 = 6$, $C_2 = 7.5$ and $G = 2.5$ (Huete *et al.*, 1994) resulting in values that range between −2.5 and 1.25. Both vegetation indices are normally rescaled to a practical range of 0–1.

Data quality control of NDVI and EVI time series was performed in a series of steps. Firstly, the vegetation index data were screened to remove outliers as follows. Values below 0.0 and data not flagged as 'good quality' according to the MODIS VI Quality dataset were eliminated. Any values of more than three standard deviations above or below the mean of the remaining values were considered as outliers and also removed. Next, to ensure relatively homogeneous observations, pixels with time series that did not consist of 100 observations or more (out of a possible maximum of 320 over the years 2000–2013 inclusive) or that did not have at least one observation in every calendar month were also removed. Finally, pixels with mean NDVI or EVI values of <0.01 were removed from the spectral analysis as these cannot present significant variations in vegetation greenness (e.g. pixels in the Atacama Desert). Despite these stringent restrictions, sufficient pixels were retained for analysis even in cloud-prone areas at the 500 and 1,000 m^2 resolutions (Figure 1).

Pre-processing of the quality-assured time series involved standard linear detrending by least squares regression (Weedon, 2003). The first and last five per cent of the detrended time series were then cosine tapered to supress 'periodogram leakage' (Priestley, 1981; Weedon, 2003).

The previous time series analysis of MODIS NDVI and EVI covering March 2000 to December 2006, that was reported in Bradley *et al.* (2011), used limited 'gap filling' as this would not have significantly influenced the relationship (coherency and phase) between paired variables (the focus of that work). However, gap filling could potentially bias the detection of annual cycles in the vegetation indices. Hence, the data used in this study were transformed to the frequency domain using the Lomb–Scargle algorithm of Press *et al.* (1992) instead of via a standard discrete Fourier Transform (including FFT). The transformed data were used to generate periodogram estimates of power – that is values of the average variance or, equivalently, the squared average deviation – or squared average amplitude sensu *stricto* – across a range of frequencies. Three applications of a discrete Hanning spectral window were used to reduce the variability of the periodograms and provide power spectral estimates with eight degrees of freedom (Priestley, 1981; Weedon *et al.*, 2015).

For regularly spaced data, the Lomb–Scargle algorithm yields identical results to a standard discrete Fourier Transform. However, for data that are irregularly spaced in time, such as the time series data used here, the specific distribution of values by time can lead to biased (over-estimated) Lomb–Scargle spectral estimates at high frequencies – necessitating a bias correction procedure (Schulz and Mudelsee, 2002). Typically, the Lomb–Scargle algorithm is used with the spacing of spectral estimates (the Rayleigh Frequency, *RF*) determined by the length of the dataset (i.e. $RF = 1/T$, where T is the total duration of the data). However, since outlier removal can lead to differences in the total length of the data, here the *RF* was adjusted in the spectrum for each pixel so that there was estimation of the power at the frequency of precisely 1/(1 year). This adjustment of *RF* prevented underestimation of the power at the annual scale that would otherwise occur due to 'scalloping loss' or 'picket fencing' (Ifeachor and Jervis, 1993).

The spectral background was modelled as resulting from a first-order autoregressive process and located using the robust median smoothing method of Mann and Lees (1996). One-sided confidence levels were obtained using standard chi-squared distribution according to the degrees of freedom of the spectral estimates (Priestley, 1981; Weedon, 2003). In order for a pixel to be judged to have statistically significant annual variations in greenness, the spectral peak at the annual scale must exceed the 95% confidence level above the spectral background. No corrections were made to the confidence levels for the effects of multiple testing (i.e. for signal detection at multiple frequencies) because this study is concerned specifically with testing peak significance at the annual scale alone. Given the degrees of freedom of our spectral estimates and the 95% confidence level threshold for significance, the uncertainty in resulting amplitudes ranges between (larger than) −41.6% and (less than) +39.2%. Figure 2 shows example time series and their power spectra which were assigned a low, medium or high average range 'amplitude' value.

In time series analysis, the term 'amplitude' relates to the deviation from the mean (or linear regression) level. Spectral estimates derived from the periodogram directly indicate the

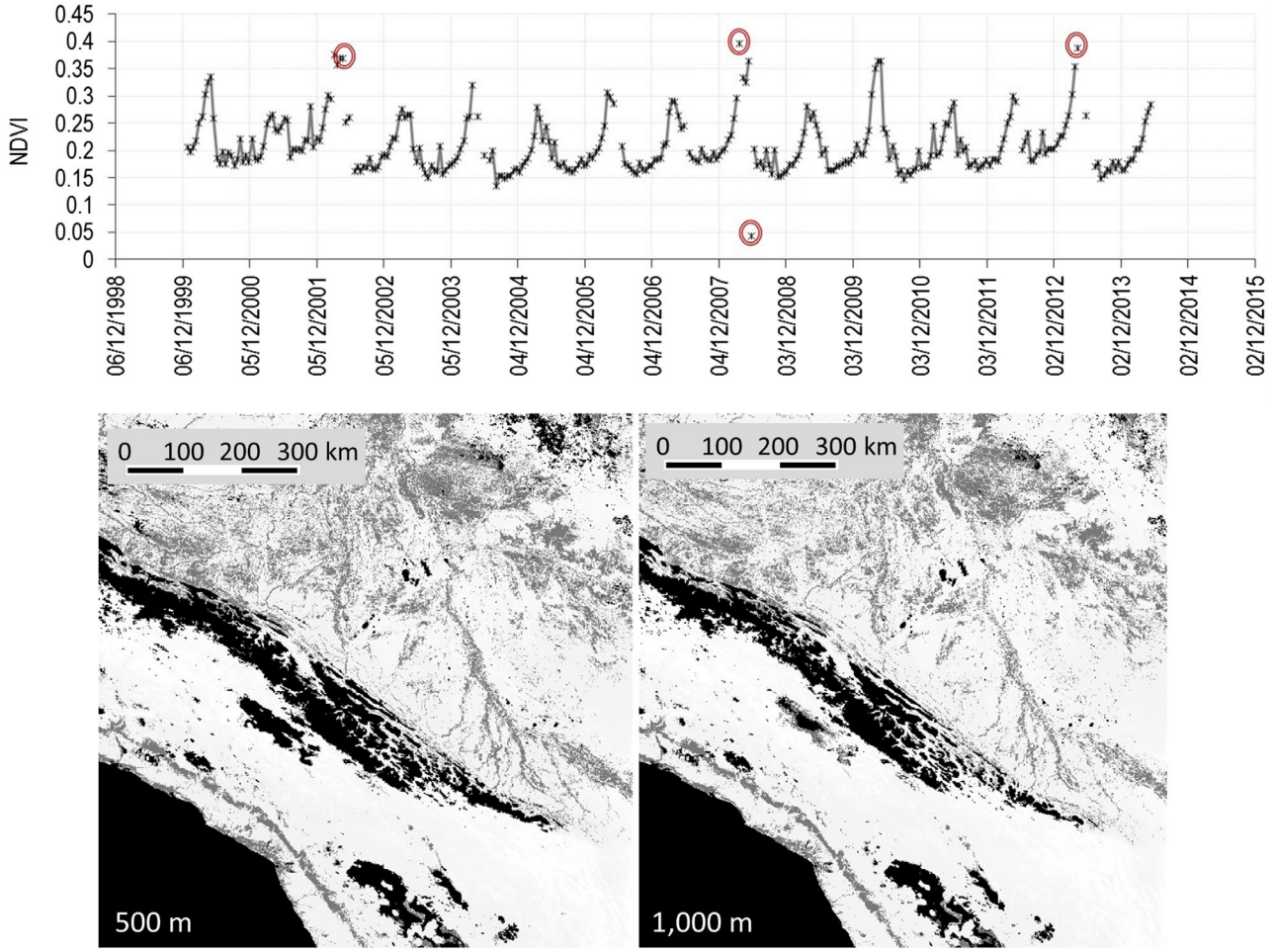

FIGURE 1 Top: Example of outliers (inside red circles) that were removed from a 16 day MODIS NDVI time series; Bottom: a MODIS tile (500 m pixels - left, 1,000 m pixels – right), including parts of Peru and Bolivia, illustrating the impact of screening for time series with insufficient good quality observations; the black areas represent pixels with insufficient data, which mainly include lakes (e.g. Titicaca) and the Andes mountain range

power (i.e. variance) at each frequency. The square root of the periodogram estimates gives the average amplitude at each frequency. Note that other methods of spectral estimation (e.g. the Blackman–Tukey method) yield the spectral density so that the variance is proportional to the area under the spectrum.

In the phenological literature, the maximum minus the minimum of a sinusoidal time series is often described as the 'amplitude' rather than as wave height or range. In order to produce maps of annual-scale variability that can be readily interpreted, the spectra were rescaled so the map values of 'average amplitude' show the average annual range (maximum minus minimum) of vegetation index values where the spectra have statistically significant annual cycles. The re-scaling of the power spectra accounts for the amplitude squaring, effects of data tapering, spectral smoothing and conversion of average deviation (average amplitude sensu *stricto*) to average range (peak to trough or maximum minus minimum) yielding the final mapped 'average amplitude'.

The reliability of the spectral results is affected by the number of data points in a time series. So in addition to the significance test, the number of good quality observations used in the analysis for each pixel is provided to allow for a further quality screening by the user. The steps described in Figure 3 delivered maps showing the average range ('amplitude') of the annual cycle observed for each pixel. Pixels with too few high quality data ('no data' flag) are distinguished from pixels with no significant annual cycles in greenness ('no significant annual cycle' flag). The values are presented as (1) absolute amplitude values of NDVI or EVI and (2) normalized amplitude values of NDVI or EVI. The normalized amplitude is the absolute amplitude value divided by the average of the 14 year VI time series. The other values found in the data are:

0 = mean VI < 0.01
−1 = 'no significant annual cycle';
−2 = 'no data' or pixels with insufficient observations;

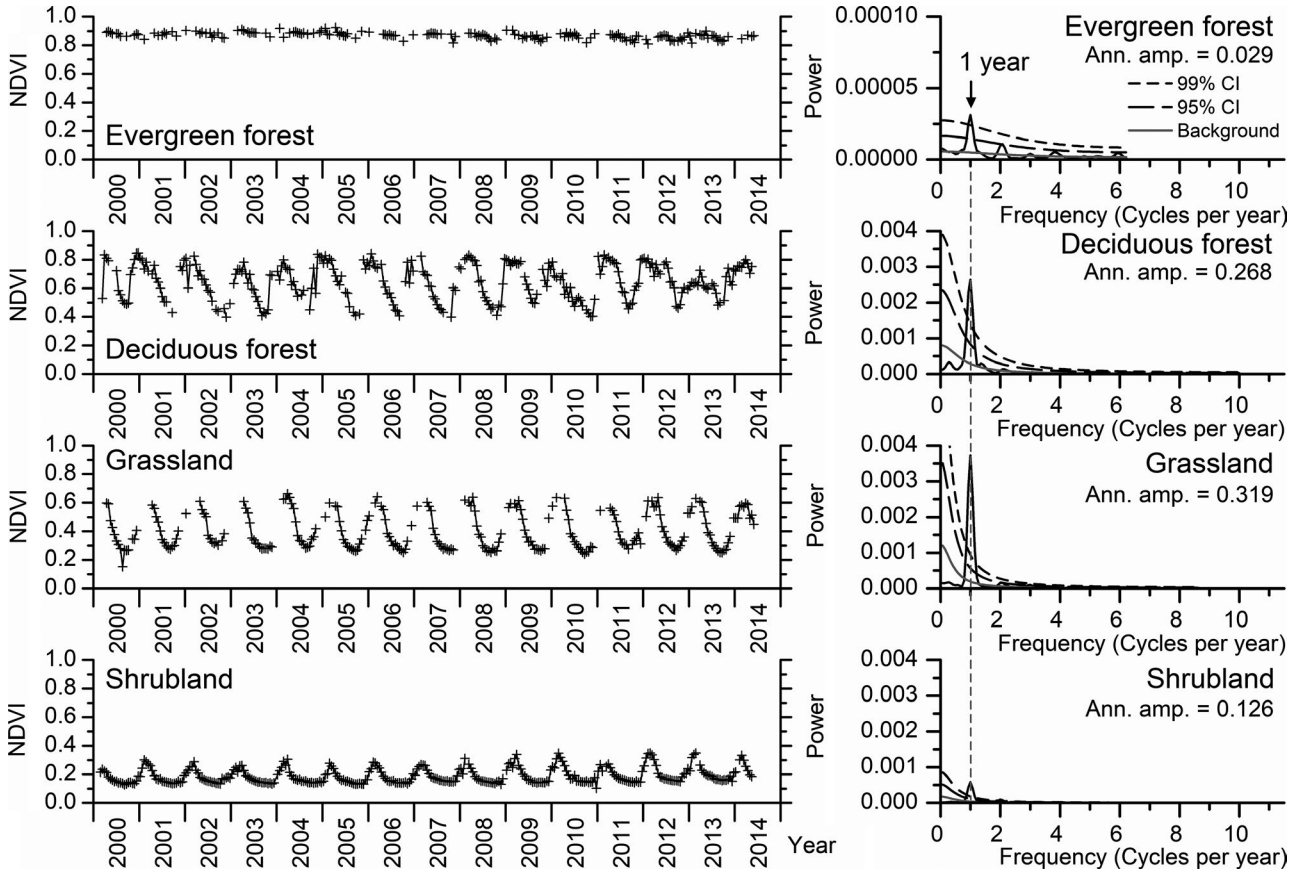

FIGURE 2 Example of 16 day MODIS NDVI time series for individual pixels representing the IGBP covers for evergreen and deciduous forests, grassland and shrubland (left), and the Lamb–Scargle transform power spectra derived from these time series (right)

−3 = Sea or Open water;

A quality layer with the % number of good quality observations for each pixel used in the spectral analysis is provided for each leaf phenology synchrony map. For each vegetation index (NDVI, EVI)m maps with the following spatial resolutions are available: 500m, 1km and a version re-gridded to 0.5 degrees.

Figure 4 shows the amplitude maps derived from NDVI and EVI for the whole of Meso- and South America and highlights the level of spatial detail achieved for two areas shown close-up (Figure S1 shows the normalized amplitude maps).

3 | INTERPRETING THE DATASET

The surface area contributing to a MODIS VI value is always larger than the nominal pixel size of the VI value (Tan *et al.*, 2006). Also, the NDVI and EVI time series present a temporal signal that results from the spatial and temporal integration of the leaf phenology of vegetation contained within the area contributing to a pixel. Hence, any phenology measures derived from such a signal should be considered at the plant community level.

The EO community has previously introduced vegetation phenology metrics (e.g. start, end and length of growing season) and productivity metrics (e.g. Maximum VI, seasonal amplitude) (e.g. Jonsson and Eklundh, 2004) that are derived from time series of VIs for each growing season. However, the dataset presented here is more aligned with the plant leaf phenological patterns characterized by ecologists through measures that capture the timing of life cycle events, such as frequency, regularity, amplitude, synchrony and duration (Newstrom *et al.*, 1994). The significance testing of the spectral peaks at the annual scale and confirms the existence or absence of a regular annual cycle in vegetation greenness. This aligns with the combined concepts of frequency and regularity described by Newstrom *et al.* (1994). The average amplitude values denote the degree of synchrony of changing net leaf greenness of the plant population (Newstrom *et al.*, 1994) across the whole pixel. Changes in land cover during the study period, which led to a sudden change in leaf phenology patterns, will influence the average amplitude values. However, testing for the potential influences of such land cover changes was beyond the scope of this investigation.

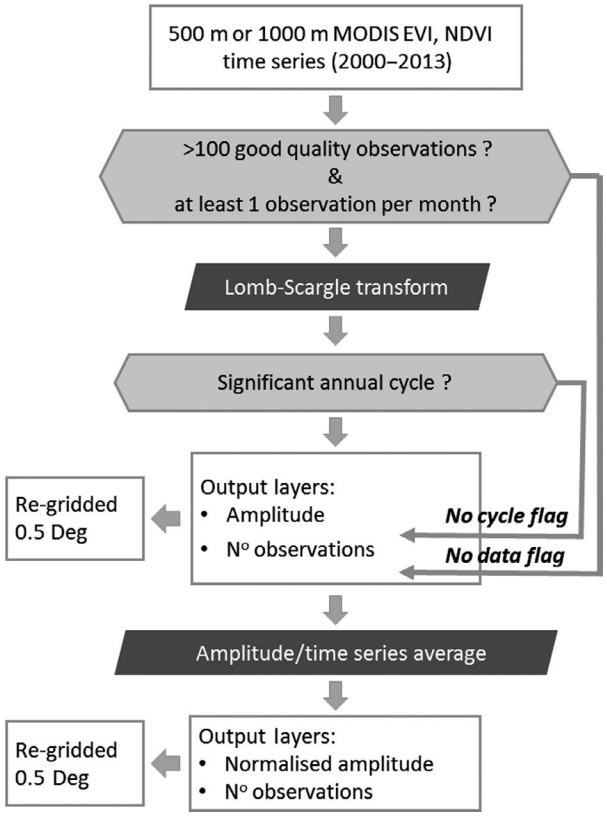

FIGURE 3 Schematic showing the main processing steps involved in deriving the absolute and normalised amplitude values from the 16 day MODIS VI time series

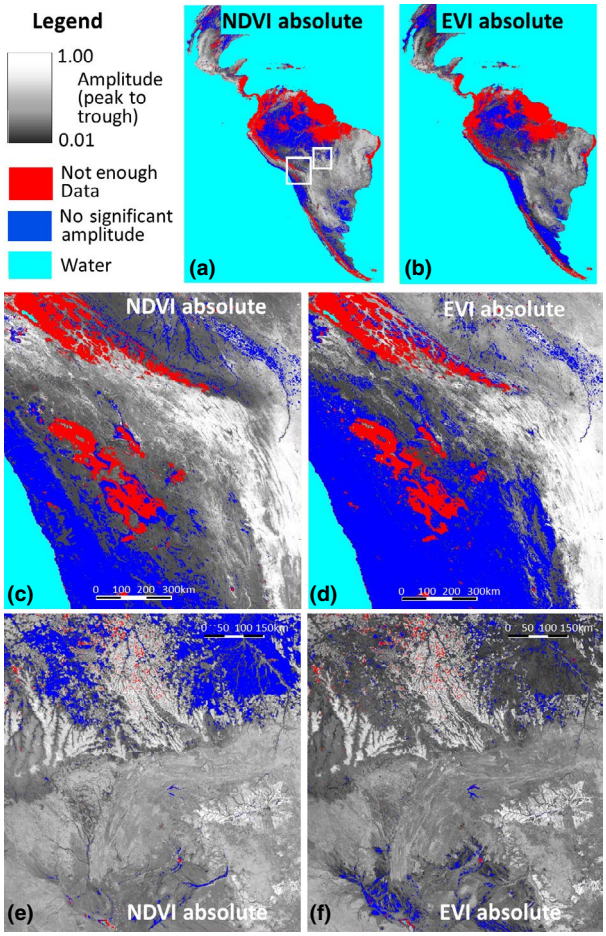

FIGURE 4 The amplitude maps derived from NDVI (a, c, e) and EVI (b, d, f) for the whole of Meso- and South America (a, b) and two smaller close-ups centred on the Salar de Uyuni desert, Bolivia (c, d) and on the state of Mato Grosso, Brazil (e, f)

3.1 | The amplitude values derived from NDVI and EVI time series

Figure 5 shows the relationship between the EVI- and NDVI-derived amplitude values for the main IGBP cover types found in Meso- and South America. As expected, the EVI-based amplitude values were larger than their corresponding NDVI-based amplitudes in areas of high biomass, because the EVI is designed to be more sensitive in these areas. Recent work from Morton *et al* (2014) has highlighted that in the tropics, the MODIS VIs are affected by bi-directional reflectance effects. In other words, the seasonal changes in solar zenith angle introduce an underlying artificial increase/decrease in EVI/NDVI values between June and October. The magnitude of this effect is greater for EVI than for NDVI. However, the average amplitude (i.e. average range) values found for the IGBP 'evergreen broadleaf forest' are generally larger than the magnitude of the BRDF effect shown in Morton *et al.* (2014) (i.e. a median of 0.04 and 0.07 versus the uncorrected June-October differences of ~0.02 and ~0.065 reported by Morton for NVDI and EVI, respectively) which is in line with Jian *et al.* (2015). Moreover, large proportions of the tropical region, where the impact of the BRDF effect is expected to be greatest, are mapped as having no significant annual cycle. This gives us confidence that the VI products are correctly identifying areas which exhibit real phenological cycles. It is important to note that the NDVI product is more conservative than the EVI product, showing a larger proportion of 500 m or 1 km pixels with no significant cycle (see Figure 4a and e).

3.2 | The absolute and normalized amplitude values

The absolute amplitude value ranges between 0 and 1, and can be interpreted as the combined effect of synchrony and vegetation cover. In other words, the two extremes could be interpreted as:

- Areas with a high vegetation cover and annual variations in greenness that is well synchronized across individuals and vegetation species, yielding a high value.
- Areas with small annual greenness variations that could indicate a high vegetation cover that is poorly synchronized or with a very low vegetation cover but that is well synchronized. Both cases would yield a low value.

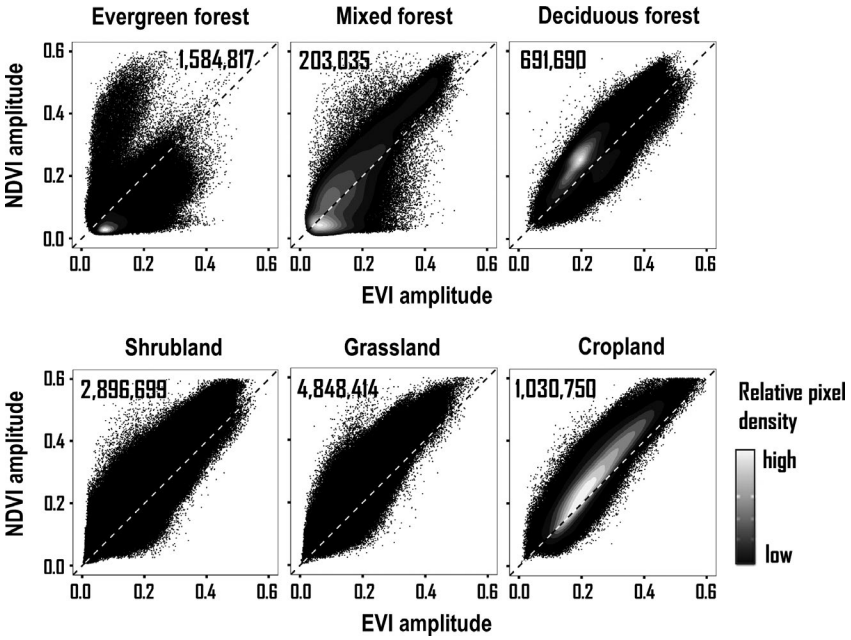

FIGURE 5 Scatterplots between the amplitudes of NDVI and EVI for the main IGBP land cover classes found in Meso- and South America, where shrubland includes IGBP classes 'closed Shrubland', 'Open Shrubland' and 'Woody Savanna'; and grassland includes IGBP classes ' Savanna' and ' Grassland'. The amplitudes shown are for a random subsample of points taken to represent each IGBP class. The total number of points sampled varies per cover class and is given in the respective plot. The black to white point density gradient is relative

It is critical that the user understands that the absolute amplitude maps relate to annual-scale VI variability. The variability might or might not be linked to the average greenness. For example, areas with low (e.g. arid grassland with low leaf area index, LAI) or high vegetation cover (subtropical grassland with high LAI), both showing a regularly timed growing season, could display a low and high amplitude value, respectively. In order to judge the average greenness, the maps include the average NDVI and EVI.

The normalized amplitude value could in theory range between 0 and 100; however, in practice the range is between 0 and ~2. It can be interpreted as the level of net synchrony between the leaf life cycles of the vegetation observed within pixels. In other words, given a within-pixel variation in environmental characteristics and climate, and the diversity of plant species in terms of their different leaf life cycle strategies (leaf phenological behaviours), a particular pixel will be between two extremes:

- Well synchronized, where the leaf life cycle of whole plant communities are driven by the same environmental and climatological factors, yielding a high amplitude value.
- Poorly synchronized where the life cycles of individuals within a community are likely to be driven by different environmental and climatological factors, yielding a low amplitude value.

Note that areas with low (e.g. arid grassland) and high vegetation cover (tropical forest) showing a regularly timed and synchronized growing season could potentially show similar normalized amplitude values.

In both cases (absolute and normalized), combining the amplitude with land cover information will help make the distinction between possible options.

4 | EVALUATING THE AMPLITUDE VALUE

We used two approaches to evaluate the amplitude values. The first involved comparing the full extent of the average amplitude data grids with the independent IGBP land cover map (Loveland and Belward, 1997; Friedl *et al.*, 2010). This was done with histograms showing the distribution of the amplitude values found within each of the IGBP cover classes. The second approach compared published in situ site observations that describe canopy leaf life cycles through leaf flushing and falling, LAI and litter fall or other observations, with the corresponding amplitude values of these sites.

4.1 | Comparison with independent IGBP land cover map

Figure 6 shows for NDVI the distribution (i.e. histograms) of the 500 m^2 average amplitude values found within the land cover classes of the 500 m IGBP global land cover product – MCD12Q1 (Figure S2 in Supplement shows similar histogram plots for EVI).

The histogram plots for the land cover classes categorized explicitly as 'evergreen' and 'deciduous' confirm our interpretation of the amplitude values (i.e. a measure of synchronized behaviour of the vegetation found within a pixel with respect to their leaf phenology). We find a large proportion of pixels with

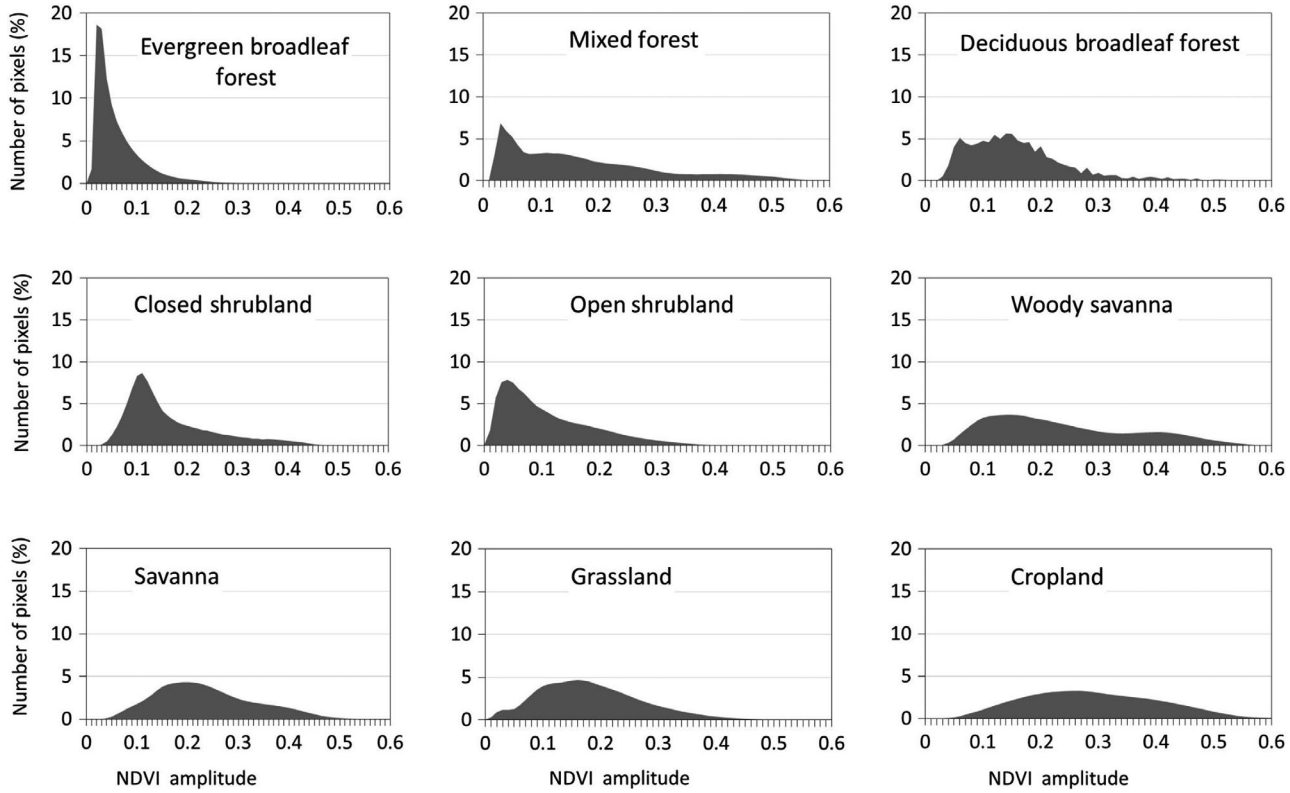

FIGURE 6 Histograms of the 500m NDVI amplitude values found within each of the main IGBP land cover classes. The histograms are expressed in per cent of total pixel number

TABLE 1 Proportion of pixels with 'no data' and 'no significant peak' and median of amplitudes and normalized amplitudes found within the prevailing IGBP land cover classes

IGBP land cover class	Total No pixels	Pixels with 'no significant peak' (%)		Amplitude (median)		Normalized Amplitude (median)	
		NDVI	EVI	NDVI	EVI	NDVI	EVI
Evergreen Broadleaf Forest	32,524,979	63	43	0.04	0.07	0.06	0.14
Mixed Forest	1,520,450	14	12	0.14	0.15	0.20	0.42
Deciduous Broadleaf Forest	2,878,948	1	1	0.25	0.20	0.37	0.59
Closed Shrubland	287,361	11	12	0.13	0.09	0.28	0.40
Open Shrubland	12,274,625	21	30	0.09	0.07	0.37	0.45
Woody Savanna	6,724,949	5	5	0.21	0.19	0.31	0.51
Savanna	16,306,392	3	4	0.22	0.18	0.37	0.51
Grassland	9,372,956	26	19	0.17	0.12	0.40	0.45
Cropland	5,179,719	17	11	0.28	0.23	0.20	0.40
Permanent Wetlands	1,411,861	34	40	0.14	0.12	0.54	0.71

no annual cycle (Table 1) or with annual cycles showing very low average amplitudes within the IGBP cover class defined as 'evergreen broadleaf forest'. By contrast, we find a large proportion of pixels with significant annual NDVI cycles (Table 1) showing large average amplitudes in the cover class defined as 'deciduous broadleaf forest'. The 'mixed forest' class shows

TABLE 2 Description of sites identified in literature used to evaluate the VI-derived amplitude (published field observations are provided in Supplement)

Site No	References	Location: Lat, Long	Site description	Leaf phenology Level of synchrony observed in situ	No plots, Length obs.	Observations related to leaf phenology
1	Williams-Linera (1997)	19°29′30″N; 96°57′55″W[a]	Lower montane wet forest, Mexico	Deciduous and evergreen trees species show synchronized leaf flushing. Only deciduous species show synchronized leaf shedding.	1 site, 5 years	Leaf shedding and flushing for individual trees
2	Maass *et al.* (1995)	19°30′N; 105°03′W	Seasonally dry tropical forest, Mexico	Strongly synchronized seasonal LAI changes and leaf shedding.	3 sites, 2 years	LAI through light transmission method, leaf litter
3	Rivera *et al.* (2002)	15°55′31″S; 47°51′14″W	Tropical dry forest, Brazil	A combination of brevi-deciduous and evergreen tree species showing a mixed picture of synchronized photoperiod induced leaf flushing and more variable dry season induced leaf flushing.	1 site, 3 years	Timing of leaf shedding and flushing for individual trees
4	https://fluxnet.ornl.gov/site/90	21°38′13″S; 47°47′25″W	Sugar cane, Brazil		na	na
5	Morellato *et al.* (2000)	24°13′29″S; 48°05′03″W[b]	Atlantic Forest, Brazil	Forest maintains its evergreen appearance throughout the year with high proportion of trees showing synchronized shedding and flushing leaves during the 6 month rainy season.	4 sites, 1.5 years	Leaf shedding and flushing for individual trees
6	Wallace and Painter (2002)	13°36′S; 60°55′W	Southern Amazonian tropical forests, Noell Kempff Park, Bolivia	A mosaic of 5 distinct habitat types. 3 habitats show no seasonal variation in foliage abundance. 2 habitats show a seasonal signal, one of which is very strong. There is synchrony between the 2 seasonal sites.	5 sites, 1 year	% crown covered by leaves for individual trees to produce a site foliage index
7	Justiniano and Fredericksen (2000)	16°13′S; 61°50′W	Dry forest, Bolivia	49% of trees become leafless at some point during the dry season. Most canopy tree species (65%) are deciduous, while most sub-canopy tree species (85%) are evergreen.	2 sites, 2 years	Presence of leaf for individual trees, timing of leaf fall for majority of trees
8	RAINFOR Tam-06; Tam-09 Malhi *et al.* (2002) Chavanna *et al.* (in prep)	12°50′24″S, 69°17′59″W; 12°49′48″S, 69°16′48″W	Lowland tropical rainforest, Peru	Canopy trees show a very slightly synchronized temporal evolution in leave age and leaf volume which is in sink with the dry season.	2 plots, 9 months	Leaf demography for 1m branches of individual canopy trees.
9	Murphy and Lugo (1986)	17°58′26″N; 66°51′34″W	Deciduous subtropical dry forest, Puerto Rico	A site level 50% reduction of LAI between the dry and wet periods	1 plot, 2.5 years	LAI at 3 monthly intervals

(Continues)

TABLE 2 (Continued)

Site No	References	Location: Lat, Long	Site description	Leaf phenology Level of synchrony observed in situ	No plots, Length obs.	Observations related to leaf phenology
10	Rivera *et al.* (2002)	23°44′60″S; 64°48′18″W[c]	Tropical dry forest, Argentina	A combination of deciduous, brevi-deciduous and evergreen species showing synchronized photoperiod induced leaf flushing among many trees and some trees showing rain induced leaf flushing	1 site, 3 years	Timing of leaf shedding and flushing for individual trees
11	Hamilton *et al.* (2002)	4.5°N, 61.5°W (NW); 2.25°N, 59°W (SE)	Wetland, Roraima, flooding synchronous with local wet season, Brazil and Guyana		Area, 15 years	Time series of satellite passive microwave
12	Hamilton *et al.* (2002)	10°N,72°W (NW); 4°N,66°W (SE)	Wetland, Llanos del Orinoco, flooding synchronous with local wet season, Venezuela and Colombia		Area, 15 years	Time series of satellite passive microwave
13	Hamilton *et al.* (2002)	9°S, 52°W (NW); 15°S, 49.5°W (SE)	Wetland, Bananal, flooding synchronous with local wet season, Brazil		Area, 15 years	Time series of satellite passive microwave
14	Hamilton *et al.* (2002)	16°S, 58°W (NW); 21°S, 55°W (SE)	Wetland, Pantanal, flooding retained during dry season, showing dramatic contrast in dry season with surrounding area, Brazil		Area, 15 years	Time series of satellite passive microwave
15	Hamilton *et al.* (2002)	12°S, 68°W (NW); 16°S, 61°W (SE)	Wetland, Llanos de Moxos, flooding retained during dry season, Bolivia		Area, 15 years	Time series of satellite passive microwave

[a]Original cited location (19°30′N; 96°57′W) identified a 1 km pixel containing a mixed landscape of forest, arable and urban cover.

[b]Original cited location (24°14′08″S, 48°04′42″W) identified a 1 km pixel containing a mixed landscape of forest, arable and urban cover.

[c]Original cited location (23°45′52″S; 64°48′0.6″W) identified a 1 km pixel containing a mixed landscape of forest, arable and urban cover.

TABLE 3 Amplitudes and normalized amplitudes for sites identified in literature and listed in Table 2 (Published field observations are provided in the Supplement)

		Amplitude		Normalized amplitude	
Site N°	Short site description	NDVI	EVI	NDVI	EVI
1	Lower montane wet forest, Mexico	0.056	0.153	0.068	0.302
2	Seasonally dry tropical forest, Mexico	0.433	0.385	0.653	0.980
3	Tropical dry forest, Brazil	0.210	0.160	0.350	0.549
4	Sugar cane, Brazil	0.348	0.274	0.640	0.828
5	Atlantic Forest, Brazil	0.051	0.172	0.059	0.368
6	Tropical forests, Bolivia	0.039	0.104	0.044	0.183
7	Dry forest, Bolivia	0.160	0.197	0.201	0.436
8	Lowland tropical rainforest, Peru	0.037; 0.041	0.108; 0.086	0.043; 0.047	0.210; 0.166
9	Deciduous subtropical dry forest, Puerto Rico	0.118	0.110	0.151	0.233
10	Tropical dry forest, Argentina	0.261	0.351	0.325	0.780
	Wetlands	**Median of amplitudes found within wetland area**			
11	Roraima	0.229	0.277	0.088	0.048
12	Llanos del Orinoco	0.197	0.305	0.082	0.059
13	Bananal	0.190	0.237	0.093	0.051
14	Pantanal	0.162	0.289	0.107	0.076
15	Llanos de Moxos	0.112	0.238	0.076	0.071

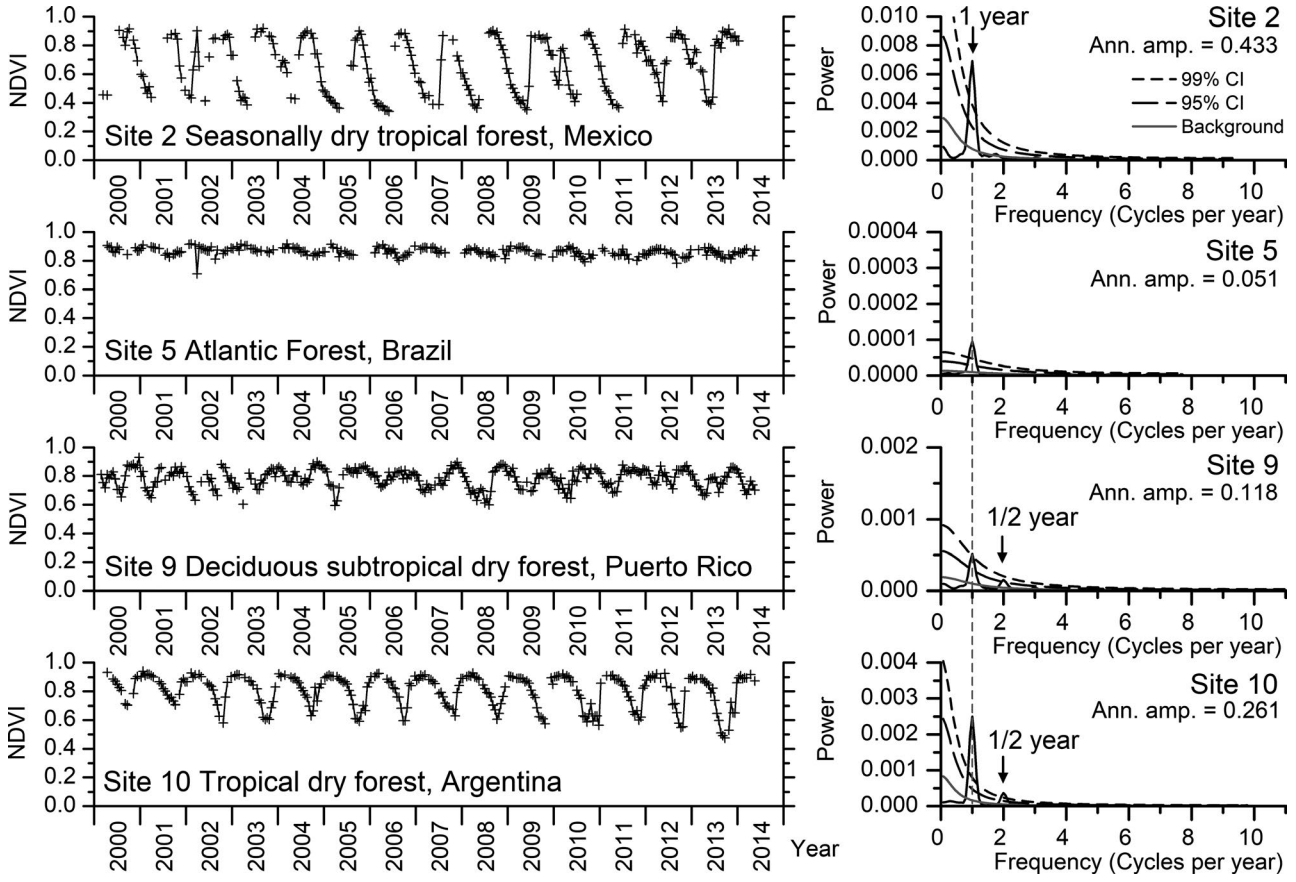

FIGURE 7 Examples of 16 day MODIS NDVI time series for 4 of the sites identified in the literature (left) and their corresponding Lamb–Scargle transform power spectra (right). The spectrum for site 9 (deciduous subtropical dry forest) shows a half-annual spectral peak associated with half-annual cycles visible in the time series, while the spectrum for site 10 (tropical dry forest) shows a half-annual spectral peak due to the second harmonic of the annual cycle in the time series, showing non-sinusoidal annual cycles with no half-annual oscillations

TABLE 4 List of available data

Data description	Pixel resolution
Absolute amplitude based on EVI (2nd layer: % number of good quality observations used)	500 m
Absolute amplitude based on EVI (2nd layer: % number of good quality observations used)	1 km
Absolute amplitude based on EVI (2nd layer: % number of good quality observations used)	0.5 Deg
Absolute amplitude based on NDVI (2nd layer: % number of good quality observations used)	500 m
Absolute amplitude based on NDVI (2nd layer: % number of good quality observations used)	1 km
Absolute amplitude based on NDVI (2nd layer: % number of good quality observations used)	0.5 Deg
Normalized amplitude based on EVI (2nd layer: % number of good quality observations used)	500 m
Normalized amplitude based on EVI (2nd layer: % number of good quality observations used)	1 km
Normalized amplitude based on EVI (2nd layer: % number of good quality observations used)	0.5 Deg
Normalized amplitude based on NDVI (2nd layer: % number of good quality observations used)	500 m
Normalized amplitude based on NDVI (2nd layer: % number of good quality observations used)	1 km
Normalized amplitude based on NDVI (2nd layer: % number of good quality observations used)	0.5 Deg

a spread-out histogram, but with a higher proportion of pixels with low average amplitudes, confirming that this class indeed represents a mixture of deciduous and evergreen phenological behaviours. The histograms of the shrubland, savanna, grassland and cropland classes indicate a mixture of phenological behaviour for all these classes.

4.2 | Comparisons with in situ observations

Table 2 lists and describes 15 in situ sites that were found in the literature and used to evaluate the amplitude values. Figure S3 in supplement shows the location of these sites, Table 3 lists their NDVI and EVI amplitudes (absolute and normalized) and Figure 7 the NDVI time series and resulting Lomb–Scargle power spectra for four of these sites.

The sites vary in location (from Mexico to Argentina) and vegetation type (i.e. evergreen forests, deciduous forests, savanna, wetlands and sugar cane). The oldest cited work is from 1986, the most recent from 2008. In most cases, the data were collected from nature reserves, so we can assume that the vegetation will not have changed much over time in terms of its overall composition and, thus, phenological behaviour. Google Earth was also used to ensure that the sites were not located in heterogeneous areas or areas heavily impacted by humans.

Nine of the 15 sites are sites with suitable field data most of which are presented as a time series plot. The published plots are shown in the supplement (Figures S4-S12). The Sugar cane site – an active Fluxnet site but with no site based phenology observations – was included as an example cropland site with a large amplitude. Although the type and quality of the reported field data varies substantially, the amplitudes derived for the sites, from both NDVI and EVI, generally match the level of synchronicity described. Sites (5, 6 and 8) with species or trees exhibiting unsynchronized leaf flushing and leaf fall have the lowest amplitudes (NDVI amplitude: 0.05, 0.04 and 0.04). Sites (1, 3, 7, 9 and 10) which show a mixture of synchronized behaviour with unsynchronized behaviour or those with an 'evergreen' layer show amplitudes in the middle range (NDVI amplitude: 0.21, 0.16, 0.12 and 0.26). Finally, sites which are fully synchronized (2 and 4), show the highest amplitudes (NDVI: 0.43 and 0.34). The one exception is site 1 which has a low amplitude (NDVI: 0.06) but which has field observations showing a synchrony in leaf flushing (deciduous and evergreen species) and partial synchrony in leaf fall (deciduous species only). Site 1 contains a mixture of deciduous and evergreen species, the latter of which are retaining leaves throughout the year. This could explain why the field data suggest there is a highly significant annual cycle, but the map shows a small amplitude. Figure 7 illustrates four of these published sites plus their NDVI power spectra – showing the highly significant annual-scale variations even for sites with low variability (e.g. site 5).

The remaining five sites are the major wetland areas of South America. Their location and inundation patterns are described by Hamilton *et al.* (2002). All five wetland sites show a medium amplitude. This is in contrast with the larger amplitudes of their non-inundated surroundings which are much more affected by the local wet and dry seasons.

5 | DATASET LOCATION AND FORMAT

The data (listed in Table 4) are freely available for download from the EIDC data centre (http://eidc.ceh.ac.uk/ https://doi.org/10.5285/dae416b4-3762-45bd-ae14-c554883d482c) as GeoTiffs.

UL 40N, 143.59W, LR 60S, 60W.

George, C. T., Hayman, G., Weedon, G. P., Gerard, F. F. (2018). Leaf phenology synchrony for Meso- and South America. NERC Environmental Information Data Centre. https://doi.org/10.5285/dae416b4-3762-45bd-ae14-c554883d482c

The code used to process the data is available from https://github.com/GarryHayman/Time_Series.

6 | DATASET USE AND REUSE

We have shown that the magnitude of the amplitude value could be interpreted as the degree at which the leaf flushing and leaf senescence of the vegetation in a particular area is synchronized. There is an increasing impetus to introduce plant diversity and plasticity into global dynamic vegetation models and allowing for variations in leaf life cycle strategies is an important aspect of this. The introduction of ecosystem leaf phenology synchronicity represented in the form of a normalized amplitude map with tested significance can be used to evaluate the spatial distribution of net ecosystem phenology patterns resulting from the range of leaf life cycle strategies or leaf turn over rates introduced into the DGVMs. To facilitate this, the map could be combined with a land cover (e.g. IGBP) and woody cover map (e.g. Hansen *et al.*, 2013) to locate aseasonal (~ evergreen) and seasonal (~ deciduous) forests in the tropical and subtropical regions, or help better understand the leaf life cycle dynamics present in the woody-grassland matrix of the savannah. The product can also be used to inform on the likely seasonal variation in leaf age distribution and so help explain current observed mismatches between flux tower and model GPP estimates (Wu *et al.*, 2016; Albert *et al.*, 2018). The high spatial resolution grid layers could also provide valuable continental scale information for biodiversity and ecology studies that require a baseline map of community level leaf phenology behaviour.

ACKNOWLEDGEMENTS

The research presented in this paper was funded by NERC national capability funding and the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 283093 – The Role of Biodiversity In climate change mitigatioN (ROBIN).

ORCID

France F. Gerard https://orcid.org/0000-0001-5916-5174
Graham P. Weedon https://orcid.org/0000-0003-1262-9984

REFERENCES

Albert, L.P., Wu, J., Prohaska, N., de Camargo, P.B., Huxman, T.E., Tribuzy, E.S., *et al*. (2018) Age-dependent leaf physiology and consequences for crown-scale carbon uptake during the dry season in an Amazon evergreen forest. *New Phytologist*, 219, 870–884.

Arino, O., Ramos, P., Jose, J., Kalogirou, V., Bontemps, S., Defourny, P., *et al*. (2012). *Global Land Cover Map for 2009 (GlobCover 2009)*. European Space Agency (ESA) & Université catholique de Louvain (UCL).

Arora, V.K. and Boer, G.J. (2005) A parameterization of leaf phenology for the terrestrial ecosystem component of climate models. *Global Change Biology*, 11, 39–59.

Bradley, A.V., Gerard, F.F., Barbier, N., Weedon, G.P., Anderson, L.O., Huntingford, C., *et al*. (2011) Relationships between phenology, radiation and precipitation in the Amazon region. *Global Change Biology*, 17, 2245–2260.

Buitenwerf, R., Rose, L. and Higgins, S.I. (2015) Three decades of multi-dimensional change in global leaf phenology. *Nature Climate Change*, 5, 364–368.

Eastman, J.R., Sangermano, F., Machado, E.A., Rogan, J. and Anyamba, A. (2013) Global trends in seasonality of normalized difference vegetation index (NDVI), 1982–2011. *Remote Sensing*, 5, 4799–4818.

Fensholt, R., Sandholt, I. and Rasmussen, M.S. (2004) Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements. *Remote Sensing of Environment*, 91, 490–507.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., *et al*. (2010) MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114, 168–182.

Gao, X., Huete, A.R., Ni, W.G. and Miura, T. (2000) Optical-biophysical relationships of vegetation spectra without background contamination. *Remote Sensing of Environment*, 74, 609–620.

Hamilton, S.K., Sippel, S.J. and Melack, J.M. (2002) Comparison of inundation patterns among major South American floodplains. *Journal of Geophysical Research: Atmospheres*, 107, LBA 5-1-LBA 5-14.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., *et al*. (2013) High-resolution global maps of 21st-century forest cover change. *Science*, 342, 850–853.

Huete, A.R. (1988) A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25, 295–309.

Huete, A., Justice, C. and Liu, H. (1994) Development of vegetation and soil indices for MODIS-EOS. *Remote Sensing of Environment*, 49, 224–234.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X. and Ferreira, L.G. (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83, 195–213.

Ifeachor, E.C. and Jervis, B.W. (1993) *Digital Signal Processing*. Wokingham, England: Addison-Wesley.

Ivits, E., Cherlet, M., Horion, S. and Fensholt, R. (2013) Global biogeographical pattern of ecosystem functional types derived from earth observation data. *Remote Sensing*, 5, 3305–3330.

Jian, B., Yuri, K., Sungho, C., Taejin, P., Jonathan, B., Philippe, C., *et al*. (2015) Sunlight mediated seasonality in canopy structure and photosynthetic activity of Amazonian rainforests. *Environmental Research Letters*, 10, 064014.

Jonsson, P. and Eklundh, L. (2004) TIMESAT - a program for analyzing time-series of satellite sensor data. *Computers & Geosciences*, 30, 833–845.

Justice, C.O., Vermote, E., Townshend, J.R.G., Defries, R., Roy, D.P., Hall, D.K., *et al*. (1998) The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, 36, 1228–1249.

Justiniano, M.J. and Fredericksen, T.S. (2000) Phenology of tree species in Bolivian dry forests. *Biotropica*, 32, 276–281.

Loveland, T.R. and Belward, A.S. (1997) The IGBP-DIS global 1 km land cover data set, DISCover: First results. *International Journal of Remote Sensing*, 18, 3289–3295.

Maass, J., Vose, J.M., Swank, W.T. and Martínez-Yrízar, A. (1995) Seasonal changes of leaf area index (LAI) in a tropical deciduous forest in west Mexico. *Forest Ecology and Management*, 74, 171–180.

Malhi, Y., Phillips, O.L., Lloyd, J., Baker, T., Wright, J., Almeida, S., *et al*. (2002) An international network to monitor the structure, composition and dynamics of Amazonian forests (RAINFOR). *Journal of Vegetation Science*, 13, 439–450.

Mann, M.E. and Lees, J.M. (1996) Robust estimation of background noise and signal detection in climatic time series. *Climatic Change*, 33, 409–445.

Morellato, L.P.C., Talora, D.C., Takahasi, A., Bencke, C.C., Romera, E.C. and Zipparro, V.B. (2000) Phenology of Atlantic rain forest trees: A comparative study. *Biotropica*, 32, 811–823.

Morton, D.C., Nagol, J., Carabajal, C.C., Rosette, J., Palace, M., Cook, B.D., *et al*. (2014) Amazon forests maintain consistent canopy structure and greenness during the dry season. *Nature*, 506, 221–224.

Murphy, P.G. and Lugo, A.E. (1986) Structure and biomass of a sub-tropical dry forest in Puerto Rico. *Biotropica*, 18, 89–96.

Newstrom, L.E., Frankie, G.W. and Baker, H.G. (1994) A new classification for plant phenology based on flowering patterns in lowland tropical rain-forest trees at La-Selva. *Costa Rica Biotropica*, 26, 141–159.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992). *Numerical recipes in Fortran: The art of scientific computing*. New York: Cambridge University Press.

Priestley, M.B. (1981) *Spectral analysis and time series*. London: Academic Press.

Reich, P.B. (1995) Phenology of tropical forests: patterns, causes, and consequences. *Canadian Journal of Botany*, 73, 164–174.

Reich, P.B., Uhl, C., Walters, M.B., Prugh, L. and Ellsworth, D.S. (2004) Leaf demography and phenology in Amazonian rain forest: a census of 40 000 leaves of 23 tree species. *Ecological Monographs*, 74, 3–23.

Rivera, G., Elliott, S., Caldas, L.S., Nicolossi, G., Coradin, V.T. and Borchert, R. (2002) Increasing day-length induces spring flushing of tropical dry forest trees in the absence of rain. *Trees*, 16, 445–456.

Sakschewski, B., von Bloh, W., Boit, A., Poorter, L., Pena-Claros, M., Heinke, J., *et al*. (2016) Resilience of Amazon forests emerges from plant trait diversity. *Nature Climate Change*, 6, 1032–1036.

van Schaik, C.P., Terborgh, J.W. and Wright, S.J. (1993) The Phenology of Tropical Forests: Adaptive Significance and Consequences for Primary Consumers. *Annual Review of Ecology and Systematics*, 24, 353–377.

Schulz, M. and Mudelsee, M. (2002) REDFIT: estimating red-noise spectra directly from unevenly spaced paleoclimatic time series. *Computers & Geosciences*, 28, 421–426.

Tan, B., Woodcock, C.E., Hu, J., Zhang, P., Ozdogan, M., Huang, D., *et al*. (2006) The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions. *Remote Sensing of Environment*, 105, 98–114.

Vina, A. and Gitelson, A.A.. (2005). New developments in the remote estimation of the fraction of absorbed photosynthetically active radiation in crops. *Geophysical Research Letters*, 32, 1–4.

Wallace, R.B. and Painter, R.L.E. (2002) Phenological patterns in a southern Amazonian tropical forest: implications for sustainable management. *Forest Ecology and Management*, 160, 19–33.

Weedon, G.P. (2003) *Time-Series Analysis and Cyclostratigraphy. Examining Stratigraphic Records of Environmental Cycles*. Cambridge: Cambridge University Press.

Weedon, G.P., Prudhomme, C., Crooks, S., Ellis, R.J., Folwell, S.S. and Best, M.J. (2015) Evaluating the performance of hydrological models via cross-spectral analysis: Case study of the Thames Basin, United Kingdom. *Journal of Hydrometeorology*, 16, 214–231.

de Weirdt, M., Verbeeck, H., Maignan, F., Peylin, P., Poulter, B., Bonal, D., *et al*. (2012) Seasonal leaf dynamics for tropical evergreen forests in a process-based global ecosystem model. *Geoscientific Model Development*, 5, 1091–1108.

White, M.A., de Beurs, K.M., Didan, K., Inouye, D.W., Richardson, A.D., Jensen, O.P., *et al*. (2009) Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. *Global Change Biology*, 15, 2335–2359.

Williams-Linera, G. (1997) Phenology of deciduous and broad-leaved-evergreen tree species in a Mexican tropical lower montane forest. *Global Ecology and Biogeography Letters*, 6, 115–127.

Wright, S.J. and Schaik, C.P.V. (1994) Light and the phenology of tropical trees. *The American Naturalist*, 143, 192–199.

Wu, J., Albert, L.P., Lopes, A.P., Restrepo-Coupe, N., Hayek, M., Wiedemann, K.T., *et al*. (2016) Leaf development and demography explain photosynthetic seasonality in Amazon evergreen forests. *Science*, 351, 972–976.

Yan, Y., Zhao, B.I.N., Chen, J., Guo, H., Gu, Y., Wu, Q., *et al*. (2008) Closing the carbon budget of estuarine wetlands with tower-based measurements and MODIS time series. *Global Change Biology*, 14, 1690–1702.

Zhao, M., Peng, C., Xiang, W., Deng, X., Tian, D., Zhou, X., *et al*. (2013) Plant phenological modeling and its application in global climate change research: overview and future challenges. *Environmental Reviews*, 21, 1–14.

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

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Gerard FF, George CT, Hayman G, Chavana-Bryant C, Weedon GP. Leaf phenology amplitude derived from MODIS NDVI and EVI: Maps of leaf phenology synchrony for Meso- and South America. *Geosci Data J*. 2020;00:1–14. https://doi.org/10.1002/gdj3.87