Global Assessment of Vegetation Index & Phenology Lab (VIP) and Global Inventory Modeling and Mapping Studies (GIMMS) Version 3 Products

Michael Marshall[a]†, Erick Okuto[a], Yanghui Kang[b], Erick Opiyo[a], Muhammed Ahmed[a]

a: Climate Research Unit, World Agroforestry Centre, United Nations Ave, Gigiri, P.O. Box 30677, Nairobi, 00100, Kenya
b: Center for Sustainability and the Global Environment, University of Wisconsin-Madison, 1710 University Ave, Madison, WI, 53726, USA
†: corresponding author (Tel: +254207224244, Fax: +254207224001, Email: m.marshall@cgiar.org)

Abstract

Earth observation based long-term global vegetation index products are used by scientists from a wide range of disciplines concerned with global change. Inter-comparison studies are commonly performed to keep the user community informed on the consistency and accuracy of such records as they evolve. In this study, we compared two new records: 1) Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index Version 3 (NDVI3g) and 2) Vegetation Index & Phenology Lab (VIP) Version 3 NDVI (NDVI3v) and Enhanced Vegetation Index 2 (EVI3v). We evaluated the two records via three experiments that addressed the primary use of such records in global change research: 1) ~~prediction of the~~ Leaf Area Index (LAI) ~~used in light-use efficiency modeling~~; 2) ~~estimation of~~ vegetation climatology ~~in Soil-Vegetation-Atmosphere Transfer models~~; and 3) trend analysis of the magnitude and ~~phenology~~ timing of vegetation productivity. ~~Experiment one, unlike~~Unlike previous ~~inter-comparison~~ global studies, ~~was performed with~~ a unique Landsat 30 m spatial resolution and *in situ* LAI database for major crop types on five continents was used to evaluate the performance of not only NDVI3g and NDVI3v, but EVI3v as well. The performance of NDVI3v and EVI3v was worse than NDVI3g using the *in situ* data, which was attributed to the fusion of GIMMS and MODIS data in the VIP record. EVI3v has potential to contribute biophysical information beyond NDVI3g and NDVI3v to global change studies, but we caution its use due to the poor performance of EVI3v in this study. Overall, ~~the two records showed a high level of agreement~~

Formatted: Font: Italic

~~both in direction and magnitude on a monthly basis, though VIP values were higher and more variable and showed lower correlations and higher error with *in situ* LAI. The~~ the records were most consistent at northern latitudes during the primary growing season and southern latitudes and the tropics throughout much of the year, while the records were less consistent at northern latitudes during green-up and senescence and in the great deserts of the world throughout much of the year. These patterns led to general agreement (disagreement) between trends in the magnitude (timing) of NDVI over the study period. Bias in inter-calibration of the VIP record at northernmost latitudes was suspected to contribute most to these discrepancies. ~~The two records were also highly consistent in terms of trend direction/magnitude, showing a 30+ year increase (decrease) in NDVI over much of the globe (tropical rainforests). The two records were less consistent in terms of timing due to the poor correlation of the records during start and end of growing season.~~

Formatted: Indent: First line: 1.27 cm

Key words: Normalized Difference Vegetation Index (NDVI); Leaf Area Index; Enhanced Vegetation Index (EVI); remote sensing; agro-ecosystems

1.0 Introduction

The Normalized Difference Vegetation Index (NDVI) (Rouse, 1974) is defined as (ρ_{NIR} – ρ_{RED})/(ρ_{NIR} + ρ_{RED}), where ρ_{NIR} and ρ_{RED} are surface reflectance in the Near Infrared (NIR: 0.725–1.10 µm) and visible red (0.58–0.68 µm), respectively. As plants become more photoactive, they absorb more visible red light due to the chlorophyll content of leaves and stems, and scatter more in the Near Infrared due to the alignment of cell walls (Tucker et al., 1994). This relationship, detected by remote sensing instruments at the canopy scale, has the effect of making the index increase (decrease) as the density of the canopy increases (decreases) (Tucker, 1979). As such, NDVI has been used widely in global change research with Earth observation remote sensing for three general purposes: 1) the estimation of canopy properties related to light-use efficiency, such as the Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation intercepted by the canopy (F_{PAR}) (e.g. Zhu et al. (2013)); 2) representation of vegetation climatology in Soil-Vegetation-Atmosphere Transfer models (e.g. O'ishi and Abe-Ouchi (2009)); and 3) detection of trends in vegetation (e.g. de Jong et al. (2011)) and phenology (e.g. de Jong et al. (2012)). Several agro-ecosystem modeling applications fall into these categories, including: agro-climate forecasting (Funk and Brown, 2006); drought monitoring (Karnieli et al., 2006); and crop yield estimation (Xin et al., 2013). Although NDVI is widely used, it is sensitive to atmospheric effects, soil background, and saturates at high LAI. The Enhanced Vegetation Index (EVI) was introduced to overcome these limitations, as it includes a visible blue band to reduce atmospheric effects, calibration terms to reduce the effects of soil background, and does not saturate as severely as NDVI at high LAI (Huete et al., 2002). EVI has also been used in a wide array of global change studies, but post 2000, when the Moderate-Resolution Imaging Spectroradiometric (MODIS) satellite sensor began retrieving visible blue reflectance (see Huete et al. (2010) for a review).

The Advanced Very High Resolution Radiometer (AVHRR) is the most commonly used sensor for long-term (i.e. pre-MODIS) global change studies, because it began retrieving visible red and NIR reflectance needed to estimate NDVI from ~~in~~ 1981 ~~and thus facilitates 30+ year time series analyses of NDVI~~ (Brown et al., 2006). The AVHRR sensor has been on board eight National Oceanic and Atmospheric Administration (NOAA) satellites: 7 (1981-1985), 9 (1985-1988 and 1994-1995 descending), 11 (1988-1994), 14 (1995-2000), 16 (2000-2003), 17 (2003-2009), 18 (2005-present), and 19 (2009-present). Reflectance data collected from the earlier AVHRR sensors (7, 9, 11, and 14) were difficult to process and synthesize, because they lacked onboard calibration; the NIR channel was sensitive to water, sun glint, glaciers at high latitudes, and clouds; and of orbital drift (Rao and Chen, 1995, 1996). These issues were rectified with the launch of the AVHRR sensors onboard NOAA 16, 17, 18, and 19, but have resulted in radiometric and spectral inconsistencies across sensors that can significantly bias global change analyses (van Leeuwen et al., 2006). Various methods have been developed to make these data continuous and consistent through time, but take different approaches and are frequently updated, necessitating new accuracy assessments to inform the user community as they evolve.

The Global Inventory Modeling and Mapping Studies (GIMMS: Tucker et al. (1994) and Vegetation Index & Phenology Lab (VIP: Didan (2014)) AVHRR products are actively used and frequently updated, but represent fundamentally different approaches to synthesis. The NOAA Global Vegetation Index (Jiang et al., 2010) is a category onto itself, ~~but~~ since it is stationary and therefore not appropriate for change detection. Both GIMMS and VIP are aggregated to a 15-day time step from daily data and are calibrated with higher spatial resolution sensors in the period that overlaps NOAA 7, 9, 11, and 14 and NOAA 16, 17, 18, and 19. However before aggregation, the former undergoes minor radiometric and spectral correction, while the later undergoes rigorous atmospheric correction. Perhaps most importantly, GIMMS is developed solely from AVHRR, while VIP is a blend of the AVHRR 1981-1999 Long-Term Data Record (Nagol et al., 2009; Pedelty et al., 2007) and MODIS 2000-present. Finally, the VIP product includes EVI2 (Jiang et al., 2008), which is a red-NIR version of EVI that has not been widely evaluated and can potentially provide additional biophysical information and improve the accuracy of long-term global change analyses (Rocha and Shaver, 2009). Given these differences, studies have been performed at the global (Beck et al., 2011) and regional (Scheftic et al., 2014) scale to assess the performance of older product versions~~, while o~~. Only one recent study compared the ~~latest~~ product versions analyzed in this study globally, but only for the consistency of trends (Tian et al., 2015). There remains~~is~~ no general consensus on which product is superior; however, GIMMS NDVI tends to ~~perform more consistently temporally than VIP NDVI, making it~~ be more appropriate than VIP NDVI~~appropriate~~ for trend analysis, because the combination of poor orbital drift correction and blending between LTDR and MODIS potentially contributes to large interseasonal variations in VIP NDVI~~, while~~. VIP NDVI,on the other hand, may be more appropriate for estimating phenology (start of season, length of season, and timing of peak NDVI) and other applications that require absolute NDVI values.   In each case, the performance of EVI2 was not evaluated nor was *in situ* data used for intercomparison.

The aim of this study was to perform a global assessment of the latest version of GIMMS and VIP over a 30-year period (January 1982 to December 2011) in order to aid the user (global change) community in interpreting results that involve these data. In doing so, we helped resolve the superiority of one product over another. The assessment was performed with three experiments that address the three major themes of global change research that involve Earth observation remote sensing ~~previously introduced~~. Unlike other intercomparison studies, we evaluated EVI2 and used an agro-ecosystem database comprised of relatively high spatial resolution Landsat and *in situ* LAI sample pairs to assess the performance of each product for agro-ecosystems in absolute terms. In addition, unlike other studies, the trend analysis was performed not only on the magnitude of change across the globe on an annual basis, but the change in the timing of NDVI according to the unique phenology in each hemisphere.

2.0 Data, processing, and analytical methods

2.1 Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index Version 3 (NDVI3g)

The GIMMS vegetation index record evaluated is version three, which is labelled as NDVI3g for the remainder of the paper. Full details on the product version can be found in Pinzon and Tucker (2014). The new product includes a series of updates since the original GIMMS NDVI and second generation NDVIg (Tucker et al., 2005) products. Like NDVIg, it is a non-stationary NDVI series at 15-day intervals and 1/12° (~8km at the equator) resolution; corrected for orbital drift, Rayleigh scattering, and radiometric and spectral inconsistencies over deserts; and takes an empirical (Bayesian) approach to normalize overlapping AVHRR periods with another higher resolution sensor that overlaps the two periods. In addition, daily NDVI data are scaled to 15-day composites using a Maximum Value Compositing (MVC) algorithm (Holben, 1986), which reduces further inconsistencies in the daily data. The most unique development in NDVI3g is the use of Sea-viewing Wide Field-of-view Sensor (SeaWiFS) for intercalibration instead of the System Pour I'Observation de la Terre (SPOT) sensors. This is intended to reduce significant bias in NDVI at extreme northern latitudes that has been observed in SPOT imagery (Guay et al., 2014).

2.2 Vegetation Index & Phenology Lab Version 3 Normalized Difference Vegetation Index (NDVI3v) and Enhanced Vegetation Index 2 (EVI3v)

The VIP vegetation index record evaluated is also in its third version, which is labelled as NDVI3v and EVI3v for NDVI and EVI2 data, respectively, for the remainder of the paper. Further information on the product version can be found in Didan (2014). Like previous versions, it is a non-stationary series at 15-day intervals and 1/20° (~5km at the equator) resolution; corrected using radiometric, drift, and cloud screening procedures recommended in El Saleous et al. (2000), and an atmospheric algorithm that reduces the effects of Rayleigh scattering, ozone, aerosols, and water vapor (Vermote et al., 1997); and takes an empirical (linear regression by land cover type) approach for intercalibration. Unlike GIMMS, SPOT is used for intercalibration and daily data are aggregated to 15-day composites using the Constrained View angle - Maximum Value Composite (CV-MVC) approach (Cihlar et al., 1997). Unlike MVC, CV-MVC does not give preference to off-nadir values that may be higher than “true” (at-nadir) values. Version three includes one notable improvement over version two, namely the correction of NDVI and EVI2 for sparsely vegetated areas pre-MODIS era (Scheftic et al., 2014). EVI2 is derived from the following equation and responds similarly to EVI (Jiang et al., 2008):

$$E = 2.5\frac{\rho_N - \rho_R}{\rho_N + 2.4\rho_R + 1} \qquad \textbf{(1)}$$

The VIP product contained persistent data gaps due to cloud cover and other noise ~~data~~ and was at a higher spatial resolution than the GIMMS product, so additional steps were taken to process it before the assessment. A MODIS filtering algorithm described in Xiao et al. (2003), Fensholt et al. (2006), and adapted for the tropics in Opiyo et al. (2013) was used to fill ~~some~~ data flagged as less than ideal ~~gaps~~. Data gaps due to cloud cover and poor data quality were not gap-filled. The algorithm was considered a compromise between preserving the actual data as much as possible and filling missing data so that a reasonable comparison could be made. ~~Statistical smoothing could have been used to fill the remaining data gaps, but was not used, because it would have risked comparing GIMMS data to a smoother and not actual VIP data~~. **Figure 1** shows the percentage of missing data filled by the filtering algorithm. On a monthly basis, less than 20% of the data was filled for the majority of pixels. Notable exceptions were primarily in the mid and extreme latitudes during wintertime. The most severe case was in south Asia during the monsoon (June – September) where more than 50% of the pixels were filled by the filtering algorithm. After the filter was applied, NDVI3g was resampled to NDVI3v/EVI3v resolution using the gdalwarp utility (http://www.gdal.org/gdalwarp.html) with default parameters. Missing values were then made consistent across ~~the datasets~~GIMMS and VIP, so that the summary statistics (experiment two below) and trends (experiment three below) were captured only for the 15-day values that the two products shared. The datasets were then resampled back to the native NDVI3g spatial resolution for the evaluation. ~~These steps were taken to produce more reliable statistics and trends.~~

2.3 First experiment: evaluation of NDVI3g, NDVI3v, and EVI3v with biophysical data

NDVI and EVI are most commonly used in global change studies to capture F_{PAR}, which drives canopy and light interactions in SVATs and other process-based models that estimate plant productivity and evapotranspiration (Glenn et al., 2008). Monsi and Saeki (1953) found that light attenuation in the canopy followed Beer's Law (Beer, 1852). This means that for a random canopy with a spherical leaf angle distribution, LAI, the second most commonly derived biophysical parameter from NDVI and EVI, can be approximated from F_{PAR} using the following equation (Norman et al., 1995):

$$\mathbf{L} = \frac{-\ln(1 - \mathbf{F}_{\mathbf{P}})}{\mathbf{k}} \qquad \textbf{(2)}$$

Where k is an extinction coefficient and LAI is the Leaf Area Index (m^2 m^{-2}). Given the importance of NDVI and EVI in estimating F_{PAR} and LAI, standard regression techniques were used to measure the relative ability of NDVI3g, NDVI3v, and EVI3v to capture *in situ* LAI variability. It is difficult to compare these records to *in situ* LAI directly, because the NDVI/ EVI - LAI relationship is typically scale dependent or non-linear (Friedl et al., 1995; Gao et al., 2000; Hall et al., 1992; Huete et al., 2005). Therefore F_{PAR} derived from Landsat Thematic Mapper/The Enhanced Thematic Mapper Plus (TM/ETM+) 30 m resolution surface reflectance data was used intermediately to downscale NDVI3g, NDVI3v, and EVI3v to 30 m resolution to facilitate the comparison.

2.3.1 Landsat Thematic Mapper/The Enhanced Thematic Mapper Plus (TM/ETM+) and *in situ* Leaf Area Index (LAI)

The Landsat TM/ETM+ surface reflectance and *in situ* LAI data was extracted from a database that was developed to determine the ability of Landsat-based NDVI, EVI2, and other vegetation indices to predict LAI for field crops around the world. Results of the analysis, along with a full description of the database can be found in Kang et al. (2015). **Figure 2** shows the distribution of the Landsat-LAI sample pairs in the database. It includes nine major global field crops (barley, cotton, maize, pasture, potato, rice, soybean, sugar beet, and wheat) and several less common fields crops classified as "other" for purposes of this analysis. The ~~*i*~~*In situ* LAI was determined using ground-based optical (LAI 2000, AccuPar, and hemispherical) and destructive techniques and compiled from a number of sources. These include: AmeriFlux (http://ameriflux.ornl.gov/) and AsiaFlux (http://asiaflux.net/) regional flux networks; experimental and validation projects (e.g. Marshall and Thenkabail (2015)); the VALidation of

European Remote sensing Instruments project (Baret et al., 2014); the Australian Airborne Cal/val Experiments for SMOS project (Peischl et al., 2012); as well as data retrieved from peer-reviewed journals. For each LAI record in the database, Landsat TM/ETM+ radiance was extracted from the United States Geological Survey archive within a ±15-day window encompassing the date of *in situ* measurement and converted to surface reflectance with the Landsat Ecosystem Disturbance Adaptive Processing System (Masek et al., 2006). NDVI and EVI2 were computed using the equations above. In rare cases where more than one LAI observation fell in a single Landsat pixel, the LAI values were averaged, so that each *in situ* entry corresponded to a unique Landsat NDVI/EVI2 value. After averaging, the dataset consisted of 2086 LAI-Landsat pairs, which was subsequently reduced to 1459 measurements after further quality control measures described in Kang et al. (2015) were taken to remove inconsistent samples.

2.3.2 Downscaling long-term records with the Fraction Photosynthetically Active Radiation intercepted by the canopy (F_{PAR}) and evaluation with *in situ* Leaf Area Index (LAI) data

Downscaling was performed by converting AVHRR and Landsat vegetation indices to F_{PAR}. Unlike the NDVI/ EVI - LAI relationship, the NDVI/EVI - F_{PAR} relationship is quasi scale invariant (Asrar et al., 1992; Friedl et al., 1995; Gutman and Ignatov, 1998; Myneni et al., 2002; Sellers, 1985), meaning a coarse resolution F_{PAR} pixel is approximately equal to the average of overlapping higher resolution F_{PAR} pixels. ~~Hwang et al. (2011), for example, used the quality of scale invariance between NDVI and F_{PAR} to downscale MODIS (1 km and 250 m spatial resolution) NDVI to Landsat spatial resolution NDVI. Since they had access to multiple MODIS and Landsat pixels through time and the linear relationship is land cover dependent, MODIS was downscaled by multiplying each pixel by a ratio of Landsat to MODIS F_{PAR}.~~ In this study, on a per pixel basis, most of the *in situ* LAI was retrieved only once, so using a ratio-based approach was not feasible. Therefore, the AVHRR vegetation indices were downscaled to 30 m spatial resolution by regressing (linearly) Landsat F_{PAR} and NDVI3g, NDIV3v, and EVI3v F_{PAR}. In order to reduce the impact of land cover dependence, the models were developed for each crop.

The Fraction of Photosynthetically Active Radiation intercepted by the canopy was computed using the ratio method first proposed in Gutman and Ignatov (1998):

$$F_P = \frac{V \; - V_m}{V_m \; - V_m} \qquad \textbf{(4)}$$

Where VI_{min} is the vegetation index (NDVI or EVI2) for bare soil (LAI $\rightarrow$ 0), and VI_{max} is the vegetation index (NDVI or EVI2) for dense vegetation (LAI $\rightarrow \infty$). VI_{min} and VI_{max} for NDVI and EVI2 were set to 0.05 and 0.95 (Fisher et al., 2008; Mu et al., 2007). These limits are sometimes considered dependent on the spatial and temporal resolution and land cover type (Zeng et al., 2000). The limits proved arbitrary for downscaling purposes however, and using the range 0.05 to 0.95 guaranteed that fractions ranged from zero to one.

Once NDVI3g, NDIV3v, and EVI3v F_{PAR} were downscaled to corresponding Landsat data, their performance was evaluated by regressing them (linearly) with the *in situ* LAI data. Since the relationship between F_{PAR} and LAI is logarithmic, as shown in **Equation 2**, standardized residual plots (not shown) were made and linear transformations were performed to verify that the assumptions of normality were met. In most cases, transformations were not required. The performance of the final model selected in each case was characterized by the coefficient of determination (R^2), significance tests, and root-mean-square error (RMSE).

Of the original 1459 Landsat – LAI data pairs, only 242 were used for the final analysis. The majority of the data loss was due to considerable overlap of LAI data in space and time, because they were collected without remote sensing applications in mind: 1) LAI values that were captured by the same coarse resolution pixels were averaged along with Landsat NDVI/EVI2 and 2) due to the presence of missing values in the long-term records, LAI and Landsat NDVI/EVI2 were averaged on a 15-day basis. These reductions led to small sample sizes for each crop. The sample sizes for cotton and rice were so small that they were omitted to avoid over-fitting. In order to increase the sample size on a per-crop basis, two aggregations based on the presumed similarity of crop spectral/canopy characteristics were made: 1) barley and wheat (winter and spring varieties) were classified as wheat and 2) garlic, onion, potato, and sugar beet were classified as tuber.

Formatted: Font: Not Bold

2.4 Second experiment: comparison of NDVI3g and NDVI3v climatology used to parametrize SVAT models

SVAT models traditionally were stand-alone and used to simulate the interaction of incoming solar radiation with the canopy driven by F_{PAR} and ~~other biological and chemical canopies~~ biogeochemical processes for a single location, but are becoming increasingly coupled to regional and global scale climate models and run over regularized grids, given the importance of vegetation feedbacks on the atmosphere (Quillet et al., 2010). ~~With the exception of newer SVATs that include a dynamic vegetation component (see Scheiter et al. (2013) for a review), the vast majority of SVATs assume vegetation varies throughout the year without interannual variations.~~ A common dataset used to parameterize the F_{PAR}~~vegetation~~ component of SVATs is the 0.15° resolution monthly climatology of ~~F_{PAR} derived from~~ AVHRR NDVI (Gutman and Ignatov, 1998). Given the importance of the F_{PAR} climatology~~NDVI~~ ~~in representing vegetation~~ in SVATs, long-term summary statistics for NDVI3g and NDVI3v were computed as part of the assessment. EVI3v was not included in this ~~phase of the analysis~~experiment, because it does not have a GIMMS counterpart to compare it to, has different and well-documented statistical properties than NDVI, and it is derived from the same visible red and NIR channels and underwent the same corrections as NDVI3v making its comparison redundant. The summary statistics were computed from the 15-day data, but the results are presented here on a monthly basis to reflect the NDVI climatology used in SVATs. The summary statistics included: mean, standard deviation, coefficient of determination (R^2) from linear regression, and slope from linear regression. The mean and standard deviation statistics are most critical for understanding the differences in NDVI climatology, while R^2 and slope indicate the strength, magnitude, and direction of the correlation between the two datasets. All summary statistics are presented with significance (p) ≤ 0.05. Non-linear correlation statistics were also computed, but were not included, because they showed similar spatial patterns as the linear statistics.

2.5 Third experiment: comparison of NDVI3g and NDVI3v trends in magnitude and timing (phenology)

Changes in the magnitude and timing (phenology) of plant productivity are important for understanding how ecosystems respond to climate change (Nemani et al., 2003). In North America, for example, trend analysis of these changes has revealed that global warming is driving an increase in plant productivity and a lengthening of the growing season (i.e. earlier green-up in the spring and later senesce in autumn) (Barichivich et al., 2013). The characterization of the magnitude and phenology of productivity over a year is typically estimated with empirical methods that include NDVI and other bioclimatic predictors such as temperature and relative humidity (e.g. Brown et al. (2012)). In order to avoid confounding the assessment of GIMMS and VIP with other variables, harmonic regression (Eastman et al., 2009; Jakubauskas et al., 2001) was performed on the vegetation index records to measure the magnitude and timing of NDVI on an annual basis. As with experiment two, EVI3v was not evaluated in this experiment. A trend analysis was then performed on the regression parameters to compare NDVI3g and NDVI3v as surrogates for the change in magnitude and timing of plant productivity over time.

The primary parameters of harmonic regression are the amplitude (in this case the difference between peak and mean NDVI) and phase (in this case timing of NDVI peaks and troughs). Amplitude and phase are computed by fitting a series of sinusoidal functions to the time series (Eq. 3). The harmonic regression was performed on a monthly basis for each year. Monthly values were determined by taking the maximum NDVI of the two 15-day composites per month.

$$\mathbf{N}_{t} = \mathbf{N}_{0} + \sum_{i-1}^{i} \mathbf{A}_{i}\,\mathbf{c}\left(\frac{2}{\mathbf{N}}\right) + \mathbf{B}_{i}\,\mathbf{s}\left(\frac{2}{\mathbf{N}}\right) \qquad \textbf{(3)}$$

Where $NDVI_t$ is the predicted Normalized Difference Vegetation Index at month (t), $NDVI_0$ is the annual monthly mean, i is the number of harmonics up to the jth harmonic, N is the number of samples (months) in the year, and A and B are coefficients used to compute the amplitude and phase. The regression was performed for the first harmonic, which represents the primary growing season, because multimodal systems (harmonics > 1) are uncommon and capturing them risks over-fitting.

The change in amplitude and phase over time was quantified using the Theil-Sen technique (Gilbert, 1987). The Theil-Sen technique takes the median non-parametric slope over all possible pairwise slopes through time. Unlike linear regression, it does not require normality or homoscedasticity, making it appropriate for trend analyses involving NDVI data (de Beurs and Henebry, 2005). The significance of the amplitude and phase trends ($p \leq 0.05$) was identified using the non-parametric Mann-Kendall test. Since the primary growing season in the southern hemisphere occurs over two given calendar years, the trend analysis was repeated for the southern hemisphere by advancing the regression six months ahead each year. This resulted in one less year or a 29-year trend analysis for the southern hemisphere.

3.0 Results

3.1 First Experiment: performance of long-term records using Landsat F_{PAR} and *in situ* LAI

~~Of the original 1459 Landsat – LAI data pairs, only 242 were used for the final analysis. A small portion of the data loss was due to the fact that they were collected after the long-term records ended. Most of the data loss was due to considerable overlap of LAI data in space and time, because they were collected without remote sensing applications in mind: 1) LAI values that were captured by the same coarse resolution pixels were averaged along with Landsat NDVI/EVI2 and 2) due to the presence of missing values in the long-term records, LAI and Landsat NDVI/EVI2 were averaged on a 15-day basis. These reductions led to small sample sizes for each crop. The sample sizes for cotton and rice were so small that they were omitted to avoid over-fitting. In order to increase the sample size on a per-crop basis, two aggregations based on the presumed similarity of crop spectral/canopy characteristics were made: 1) barley and wheat (winter and spring varieties) were classified as wheat and 2) garlic, onion, potato, and sugar beet were classified as tuber.~~

The accuracy of each long-term record when compared to *in situ* LAI was mixed, but NDVI3g performed moderately better than NDVI3v and EVI3v. The scatterplots of predicted (downscaled) NDVI3g, NDVI3v, and EVI3v F_{PAR} versus Landsat F_{PAR} for wheat and pasture are shown in **Figure 3**, while the summary statistics of the linear models used to downscale the records for all crops with sufficient samples sizes and reasonable correlations are shown in **Table**

1. The models used to downscale NDVI3g yielded higher correlations and lower error than the models used to downscale NDVI3v for maize and wheat, while NDVI3v yielded higher correlations and lower error for soybean and pasture, and EVI3v was the most difficult to downscale of the three. Specifically, ΔR^2 for NDVI3g over NDVI3v was 0.04 for maize and 0.18 for wheat, while ΔR^2 for NDVI3v over NDVI3g was 0.06 and 0.04 for pasture and soybean. It is important to note however that the strength of the relationships were low across all records with the exception of pasture, which could be due to the homogeneity (consistent clumping) of pasture over large areas. The relationship for tuber was so poor that it was not included in the LAI evaluation. The relationship between the downscaled NDVI3g, NDVI3v, and EVI3v F_{PAR} and *in situ* LAI are shown for wheat and pasture is in **Figure 4**, while the model statistics and transformation for a linear comparison, are presented in **Table 2**. The NDVI3g-LAI models captured *in situ* variability better than NDVI3v and EVI3v for maize (ΔR^2 = 0.06), pasture (ΔR^2 = 0.11), and wheat (ΔR^2 = 0.10), with comparable results between NDVI3g and NDVI3v for soybean. EVI3v tended to perform better than NDVI3v for two of the crops: pasture (ΔR^2 = 0.05) and wheat (ΔR^2 = 0.04). As can be seen in **Figure 4**, however, the predictive power of EVI3v could be inflated by leveraging at high LAI, i.e. EVI3v tends to be more variable than NDVI3v at higher LAI.

3.2 Second experiment: similarity of NDVI3g and NDVI3v climatology

On a monthly basis, NDVI3g and NDVI3v showed a high level of consistency in terms of relative magnitude expressed as R^2 (**Figure 5**) and direction expressed as slope (**Figure 6**). Both metrics were computed with the slopes forced through the origin (0, 0). In the northern hemisphere, R^2 approached one after green-up (May) and progressively got stronger over the boreal summer months (June, July, and August). The poorest correlations ($R^2 < 0.7$) were seen primarily at the northern-most latitudes during the transition between boreal winter and spring. Correlations were more consistent in the Southern Hemisphere where snow and cloud cover was notably less than in the north. A glaring exception however was the Strut Stony Desert of South Central Australia, which showed poor correlations during the transition between Austral summer (December, January, and February) and fall. The tropics showed high and significant correlations throughout most of the year as well. The slopes followed a similar pattern as the correlations, with values approaching a one-to-one relationship (slope=1.0) after the transition from winter to spring in the northern hemisphere and consistently over much of the year in the tropics and southern hemisphere. The great deserts of the world and sparsely vegetated areas had slopes approaching zero throughout the year. Since the slopes were expressed with NDVI3v as the dependent variable and the slopes were always less than one, NDVI3g was always less than NDVI3v. The difference in NDVI3g and NDVI3v magnitudes is more clearly shown in **Figure 7**, which illustrates the monthly latitudinal mean and standard deviation for both. Mean NDVI3v was always higher and more variable than NDVI3g. In addition, large divergence in means between the two records occurred during senescence in the northern hemisphere. Other patterns were more consistent: NDVI3g and NDVI3v were high in the tropics throughout the year and peak and decline following the seasons in the northern and southern hemispheres; and the standard deviations for both were higher in the northern hemisphere than the southern hemisphere due to continentally.

3.~~2~~ 3 Third experiment: similarity of NDVI3g and NDVI3v trends in magnitude and phenology

The two NDVI records exhibited a high level of correspondence in maximum primary season NDVI (1st harmonic amplitude), both in direction and location (**Figure 8**). In terms of magnitude trends, however, NDVI3v was higher than NDVI3g. The figure was masked for pixels that had complete NDVI records to ~~guarantee accurate~~facilitate curve-fitting in a given year and then again for trends that were statistically significant over the 30-year period. This resulted in no trends over much of the northern latitudes. In addition, NDVI amplitudes ≥ 0.03 per year (or 1.0 over the 30-year period) and NDVI amplitudes ≤ -0.03 (or -1.0 over the 30-year period) were flagged as missing, since NDVI ranges from -1 to 1. In most cases, however, the increase in absolute amplitude per year was less than 0.01 or 0.3 over the 30-year period. Overall, the positive NDVI3g trends appeared to be more consistent spatially in several important cropping and grazing regions, including: the Great Plains of the United States; the Region del Norte Grande of Argentina; the Iberian Peninsula (particularly Portugal); Lesotho, South Africa (east), and Swaziland; Ganges (India) and Indus (Pakistan) Plains; the Sahel of West Africa; and Cape York Peninsula (Australia). Negative trends (also more consistent in NDVI3g) appeared to be primarily in the great deserts of the northern hemisphere. In the southern hemisphere, however, some negative trends were seen in the tropical forests of the Amazon and Congo River basins.

The two records in terms of primary season timing (1^{st} harmonic phase) showed a lower level of correspondence than for amplitude (**Figure 9**). As above, trends were not seen over much of the northern hemisphere. In addition, the NDVI phases ≥ 0.07 per year (or ~2 months over the 30-year period) and NDVI phases ≤ -0.07 (or ~-2 months over the 30-year period) were flagged as missing, because changes of more than two months were deemed aberrant. In most cases however, the absolute change in timing was less than two months. As with trends in amplitude, the trends in phase were more consistent spatially over both hemispheres from NDVI3g. Earlier green-up (negative trend) represented the majority of trends in the two datasets, though considerably less than the increase in amplitude shown in **Figure 8**. Negative trends were seen over many important cropping and grazing areas: California and the Southwestern United States; the Iberian Peninsula; the Sahel of sub-Saharan Africa; Iran (east); South Africa (west); Turkmenistan (north); and over much of the areas bordering the deserts of Australia. Later green-up (positive trend) was primarily concentrated in the great deserts (e.g. the Great Sandy and Gibson deserts of northwestern Australia).

4.0 Discussion

This study assessed the latest versions of two non-stationary and long-term vegetation index records used in global change studies. The assessment was performed with three experiments that addressed ~~the primary~~important global change applications, namely: ~~the estimation of~~ F_{PAR} and LAI; ~~estimation of SVAT~~ vegetation climatology; and trend analysis of vegetation ~~productivity~~ magnitude and ~~phenology~~timing. The results of the analysis highlight important similarities and differences between the two records that the global change community should be aware of before using them for these applications: 1) NDVI3v was consistently higher and more variable than NDVI3g, which in Tian et al. (2015) has been attributed to artificial jumps in the record between AVHRR and MODIS periods and may contribute to relatively lower correlations and higher errors with *in situ* LAI; 2) the performance of EVI3v with *in situ* LAI compared to NDVI3g was unexpectedly poor; 3) correlations between GIMMS and VIP were highest during the primary growing season, so trends in peak NDVI were fairly consistent between the two, both showing increases over much of the globe and decreases in tropical rainforests; and 4) correlations between GIMMS and VIP were lower during green-up and senescence, which were most pronounced at high latitudes where the NDVI3g product is expected to have much lower bias due SeaWifs inter-calibration. ~~so trends in NDVI timing were~~

~~less consistent between the two, however, both showed earlier green-up over much of the globe, particularly in the driest regions of the world.~~Overall, we recommend using NDVI3g over NDVI3v and EVI3v for vegetation climatology and trend analysis, because it is spatially and temporally more consistent. Unlike previous studies, however, the *in situ* LAI experiment revealed that NDVI3g is better suited for absolute measurements as well.

Formatted: Font: Italic

4.1 First Experiment: performance of long-term records using Landsat FPAR and in situ LAI

Formatted: Indent: First line: 0 cm

Unlike previous inter-comparison studies, a unique moderate resolution remote sensing and *in situ* LAI database for agro-ecosystems was used for accuracy assessment. ~~Although there was a spatial mismatch between *in situ* and AVHRR data, and the *in situ* data had a small sample size with a limited geographic extent,~~ In most cases, NDVI3g ~~appeared to be~~was more accurate than NDVI3v or EVI3v. EVI3v performed considerably worse than NDVI3g, which is surprising, because EVI tends to be better correlated than NDVI from other sensors with canopy structural properties (Huete et al., 2002). Earlier studies have suggested that the LTDR NDVI from which MODIS data is merged in the VIP product is more appropriate for modeling applications requiring absolute values (Beck et al., 2011), meaning NDVI3v should reproduce more accurate estimates of F_{PAR} and LAI than NDVI3g, but this was not the case in this study. Tian et al. (2015) assessed ~~the blended and smoothed LTDR and MODIS product~~NDVI3v. They attributed- jumps in the NDVI3v record ~~the relatively high and variable NDVI3v mainly~~ to poor orbital drift correction and the break in the LTDR and MODIS records in 2000. The reason for the poor performance of EVI2 is less clear, but clearly needs to be addressed in future work, given its potential importance to advancing global change research. ~~However, since the LTDR data appears to reproduce more accurate absolute values than GIMMS and a smoother was not used and there was a high level of correlation between NDVI3g and NDVI3v in this study,~~

~~orbital drift correction is likely not the culprit. Therefore, the blending of MODIS and LTDR is most likely the most important factor impacting the accuracy of biophysical estimates in NDVI3v and EVI3v and should be addressed in later product versions.~~

~~At the time of writing this manuscript, a VIP Version 4 is forthcoming. It will be interesting to see if this new version will produce more accurate results using the LAI-Landsat database. In the meantime, however,~~ I~~i~~f users require the higher spatial resolution offered by VIP and the added biophysical information afforded by EVI3v for application purposes, several options exist for improving their accuracy. Perhaps the most important would be to fill the remaining data gaps in the filtered VIP datasets generated here with ~~a~~ smoothed data~~r~~ (see Kandasamy et al. (2012) for examples), which will address some of the noise in the data observed in Tian et al. 2015 and this study. ~~NDVI3g has undergone extensive statistically smoothing~~. Another option widely used in the climate modeling community, ~~that could be combined with this option would be~~ is to generate an ensemble mean of NDVI3v and NDVI3g to account for some of the bias and uncertainties in each product. Finally, instead of using EVI3v, the red and NIR channels included in the VIP database could be used to calculate the Soil Adjusted Vegetation Index (SAVI) (Huete, 1988) instead. ~~The evaluation of EVI2 has so far been limited, whereas~~ Unlike EVI2, SAVI has undergone extensive evaluation.

~~The LAI-Landsat database should be combined with other databases in the future, such as the LAI for woody plant database (Iio et al., 2014), so that a large amount of data over multiple biomes are used to develop robust evaluations (Weiss et al., 2014). New databases should aim to extend the temporal ranges of biophysical data on a per-pixel basis, so that the ratio-based approach to downscaling as suggested in Hwang et al. (2011) can be performed, instead of the linear regression by crop type approach taken here. The downscaling procedure can also be~~

~~improved. In the Hwang et al. (2011) study, F_{PAR} was used to downscale MODIS data to Landsat resolution, representing a ratio of approximately 8 : 1 (250 m : 30 m), whereas in this study, Landsat F_{PAR} was used to downscale AVHRR data, representing a ratio of approximately 266 : 1 (8000 m : 30 m). The large discrepancy in resolution in this study could be resolved in the future by first downscaling AVHRR with MODIS F_{APAR} and then downscaling again using Landsat F_{APAR}.~~

4.2 Second experiment: similarity of NDVI3g and NDVI3v climatology

Formatted: Indent: First line: 0 cm

NDVI3g and NDVI3v showed a high level of agreement with one another at mid-latitudes during the primary growing season and in the densely vegetated tropics throughout most of the year, and a low level of agreement at high latitudes during winter months and in the sparsely vegetated sub-tropics throughout most of the year. The high level of agreement is expected, because data gaps, cloud contamination, and atmospheric water vapor, is less at mid-latitudes during summer months (Beck et al., 2011; Moulin et al., 1997). The high level of agreement in the tropics was more surprising, because data gaps and cloud contamination are persistent there throughout much of the year, typically leading to large discrepancies among records (Brown et al., 2006). However, as previously stated, ~~the standard smoothed VIP data was not used this study, so many of the potentially smoothed and~~ many contaminated pixels were omitted from the analysis. The large discrepancy at high latitudes could have been due to factors other than cloud contamination and other noise ~~data gaps~~, including the 1) presence of snow cover; 2) high frequency of off-nadir pixels, which would impact the results of the compositing algorithm (MVC versus CV-MVC); and perhaps most importantly, 3) use of SeaWiFS over SPOT for GIMMS inter-calibration (Hall et al., 2006). The large discrepancy in deserts and sparsely vegetated areas on the other hand was most likely due to the dominance of soil in the signal and sensitivity of NDVI to soil wetness (Jiang et al., 2006). With the high level of correlation during the primary growing season and higher and more variable NDVI3v, users should expect NDVI3v climatology during the primary growing season to be higher at mid-latitudes and in the tropics throughout most of the year, but consistent with changes in NDVI3g. During winter months, especially at high latitudes and in semi-arid to arid subtropical regions, where SeasWiFS inter-calibration is less biased, NDVI3v will be higher, more variable, and less ~~consistent with~~accurate than NDVI3g.

4.3 Third experiment: similarity of NDVI3g and NDVI3v trends in magnitude and timing

Formatted: Indent: First line: 0 cm

NDVI3g and NDVI3v both showed greening (positive NDVI amplitude) globally, with localized browning (negative NDVI amplitude) over a 30+ year time frame, but the magnitude of the trends in the latter was higher. Therefore, trend analyses of peak NDVI or annual means will be higher in NDVI3v than NDVI3g, but the direction will be the same. The direction of change in general corroborated previous global studies. The gain or loss of plant productivity is generally attributed to biophysical drivers (temperature and precipitation), human-related change, and discontinuities in the long-term record (de Jong et al., 2012). At mid-latitudes, warming (cooling) at the beginning of the growing season can lead to greening (browning) in areas where water supplies are ample. In North America east of the Great Plains, for example, greening was observed in NDVI3g and NDVI3v, which has been attributed to temperature-driven increases in plant productivity in previous studies (Wang et al., 2011). Increased rainfall (droughts) proceeding or during the growing season can lead to greening (browning) particularly in water-limited regions such as the Sahel. As shown here, the Sahel has experienced greening over the past 30+ years. This greening, typically referred to as the “re-greening of the Sahel” is defined in other studies as the increase in woody biomass (Brandt et al., 2015) that followed the recovery of rains in the 1990's after two decades of severe droughts driven by below normal sea surface temperatures in the subtropical North Atlantic (Giannini et al., 2013). Deforestation is perhaps the most~~ recognized~~ appreciated human driver of plant productivity. Browning in the Amazon and Congo River basins, as was shown in this study, has been attributed to widespread deforestation in previous studies (Hansen et al., 2010; Mayaux et al., 2013), though other drivers, such as shift in Walker circulation potentially contribute to the loss as well (Zhou et al., 2014). Greening was observed in the areas tropical rainforests as well, but this has been attributed in previous studies to rapid regrowth after deforestation, the way VIs are composited, and the methods by which trends are detected (Beck et al., 2011). Some of the trends disagree with previous research and should be addressed in future studies. Most prominent were that no trend was detected at extreme northern latitudes, though previous studies have shown summer drought-driven declines in boreal forest productivity (Goetz et al., 2005), and positive trends were detected for the Region del Norte Grande of Argentina, though previous studies have shown negative trends attributed to the rapid encroachment of agriculture into subtropical forests of the region (Paruelo et al., 2004).

NDVI3g and NDVIv both showed earlier green-up (negative NDVI phase) more than later green-up (positive NDVI phase), but they were less consistent with one another compared to trends in peak NDVI. NDVI3g and NDVI3v showed low correlations during green-up and diverging climatology during senescence, which could lead to discrepancies in the timing of start of season (SOS) and end of season (EOS). Global studies seldom analyze trends in vegetation timing. On a regional basis, however, t~~T~~he findings appear to be less consistent with previous studies~~ with the timing trends in other studies~~. Over the majority of northern regions, for example, ~~the start of season (~~SOS~~)~~ has been retreating as shown, however unlike this study, previous studies have shown that ~~the end of the season (~~EOS~~)~~ has been advancing. The combination of the two processes has led to a longer growing season attributed primarily to asymmetric and rising global temperatures. One of the limitations of the harmonic approach taken in this study is that it is rigid, i.e. it assumes that the time series oscillates at a regular interval over each year. In the future, a harmonic or other phenological model that accounts for SOS and EOS asymmetry may be more appropriate for accurate trend analysis.

5.0 Conclusion

This paper revealed important similarities and differences of two new long-term vegetation databases: Global Inventory Modeling and Mapping Studies Normalized Difference Vegetation Index Version 3 (NDVI3g) and 2) Vegetation Index & Phenology Lab Version 3 NDVI (NDVI3v) and Enhanced Vegetation Index 2 (EVI3v). Overall, NDVI3g performed better and more consistently than NDVI3v and EVI3v in three experiments designed to evaluate the two products in absolute terms and changes in magnitude and timing. ~~when downscaled with Landsat 30 m resolution fraction of photosynthetically active radiation intercepted by the canopy and compared to *in situ* Leaf Area Index (LAI). VIP processing and the approach taken to synthesize data streams contributed to higher and more variable values that adversely affected the predictive ability of the database.~~ VIP tended to be higher in magnitude, more variable, and less consistent in terms of trends, due primarily to the blending of two sensors with different attributes (AVHRR with MODIS). GIMMS, on the other hand only uses AVHRR. ~~However, the~~The two databases showed a high level of consistency during the primary growing season, which contributed to similar changes in the relative magnitude and direction of plant productivity climatology and dynamics, which are critical to global change research. The two products were less consistent in timing, especially at the start and end of the primary growing seasons at high latitudes. It is suspected that these poor correlations are attributed to the higher resolution sensors each product uses for intercalibration. ~~due in part to their poorer correlation at the start and end of growing season. New opportunities exist for improving the two products that can account for the discrepancies highlighted here. In the meantime~~In conclusion, it is suggested users requiring a long-term product to measure biophysical parameters, vegetation climatology, and trends in plant productivity magnitude and timing to use NDVI3g and to avoid using EVI3v.

Acknowledgements

This work was supported primarily through donor contributions to the Consortium of International Agricultural Research (CGIAR) Centers Research Program (CRP) on Policies, Institutions and Markets for the project, titled “Enhancing IMPACT Simulations for Climate Change Adaptation and Mitigation with innovations in remote sensing and biophysical modelling. Additional funds came from the CGIAR CRP Forest, Trees, and Agroforestry to facilitate the collection and processing of the global datasets. The LAI-Landsat database was developed with support from the National Aeronautics and Space Administration Earth and Space Science Fellowship, National Science Foundation Water Sustainability & Climate Program (DEB-1038759), North Temperate Lakes Long-Term Ecological Research Program (DEB-0822700), and University of Wisconsin-Madison Anna Grant Birge Award. Special thanks to Dr. Molly E. Brown and Dr. Kamel Didan for providing the GIMMS and VIP datasets, respectively.

List of References

Asrar, G., Myneni, R. B. and Choudhury, B. J.: Spatial heterogeneity in vegetation canopies and remote sensing of absorbed photosynthetically active radiation: A modeling study, Remote Sens. Environ., 41(2–3), 85–103, 1992.

Baret, F., Weiss, M., Allard, D., Garrigues, S., Leroy, M., Jeanjean, H., Myneni, R., Privette, J., Morisette, J., Bohbot, H., Bosseno, R., Dedieu, G., Di Bella, C., Duchemin, B., Espana, M., Gond, V., Gu, X. F., Guyon, D., Lelong, C., Maisongrande, P., Mougin, E., Nilson, T., Veroustraete, F. and Vintilla, R.: VALERI: a network of sites and a methodology for the validation of medium spatial resolution land satellite products, Institut national de la recherche agronomique, Avignon, France., 2014.

Barichivich, J., Briffa, K. R., Myneni, R. B., Osborn, T. J., Melvin, T. M., Ciais, P., Piao, S. and Tucker, C.: Large-scale variations in the vegetation growing season and annual cycle of atmospheric CO2 at high northern latitudes from 1950 to 2011, Glob. Change Biol., 19(10), 3167–3183, 2013.

Beck, H. E., McVicar, T. R., van Dijk, A. I. J. M., Schellekens, J., de Jeu, R. A. M. and Bruijnzeel, L. A.: Global evaluation of four AVHRR–NDVI data sets: Intercomparison and assessment against Landsat imagery, Remote Sens. Environ., 115(10), 2547–2563, 2011.

Beer, A.: Bestimmung der Absorption des rothen Lichts in farbigen Flüssigkeiten, Ann Phys U Chem, 1852.

De Beurs, K. M. and Henebry, G. M.: A statistical framework for the analysis of long image time series, Int. J. Remote Sens., 26(8), 1551–1573, 2005.

Brandt, M., Mbow, C., Diouf, A. A., Verger, A., Samimi, C. and Fensholt, R.: Ground- and satellite-based evidence of the biophysical mechanisms behind the greening Sahel, Press, 2015.

Brown, M. E., Pinzon, J. E., Didan, K., Morisette, J. T. and Tucker, C. J.: Evaluation of the consistency of long-term NDVI time series derived from AVHRR,SPOT-vegetation, SeaWiFS, MODIS, and Landsat ETM+ sensors, IEEE Trans. Geosci. Remote Sens., 44(7), 1787–1793, 2006.

Brown, M. E., de Beurs, K. M. and Marshall, M.: Global phenological response to climate change in crop areas using satellite remote sensing of vegetation, humidity and temperature over 26 years, Remote Sens. Environ., 126, 174–183, 2012.

Cihlar, J., Ly, H., Li, Z., Chen, J., Pokrant, H. and Huang, F.: Multitemporal, multichannel AVHRR data sets for land biosphere studies—Artifacts and corrections, Remote Sens. Environ., 60(1), 35–57, 1997.

Didan, K.: Multi-Satellite Earth Science Data Record for Studying Global Vegetation Trends and changes, The University of Arizona, Tucson, AZ., 2014.

Eastman, R., Sangermano, F., Ghimire, B., Zhu, H., Chen, H., Neeti, N., Cai, Y., Machado, E. A. and Crema, S. C.: Seasonal trend analysis of image time series, Int. J. Remote Sens., 30(10), 2721–2726, 2009.

El Saleous, N. Z., Vermote, E. F., Justice, C. O., Townshend, J. R. G., Tucker, C. J. and Goward, S. N.: Improvements in the global biospheric record from the Advanced Very High Resolution Radiometer (AVHRR), Int. J. Remote Sens., 21(6-7), 1251–1277, 2000.

Fensholt, R., Sandholt, I. and Stisen, S.: Evaluating MODIS, MERIS, and vegetation indices using in situ measurements in a semiarid environment, IEEE Trans. Geosci. Remote Sens., 44(7), 1774–1786, 2006.

Fisher, J. B., Tu, K. P. and Baldocchi, D. D.: Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites, Remote Sens. Environ., 112(3), 901–919, 2008.

Friedl, M. A., Davis, F. W., Michaelsen, J. and Moritz, M. A.: Scaling and uncertainty in the relationship between the NDVI and land surface biophysical variables: An analysis using a scene simulation model and data from FIFE, Remote Sens. Environ., 54(3), 233–246, 1995.

Funk, C. C. and Brown, M. E.: Intra-seasonal NDVI change projections in semi-arid Africa, Remote Sens. Environ., 101(2), 249–256, 2006.

Gao, X., Huete, A. R., Ni, W. and Miura, T.: Optical–Biophysical Relationships of Vegetation Spectra without Background Contamination, Remote Sens. Environ., 74(3), 609–620, 2000.

Giannini, A., Salack, S., Lodoun, T., Ali, A., Gaye, A. T. and Ndiaye, O.: A unifying view of climate change in the Sahel linking intra-seasonal, interannual and longer time scales, Environ. Res. Lett., 8(2), 024010, 2013.

Gilbert, R. O.: Statistical Methods for Environmental Pollution Monitoring, John Wiley & Sons, New York, NY., 1987.

Glenn, E. P., Huete, A. R., Nagler, P. L. and Nelson, S. G.: Relationship Between Remotely-sensed Vegetation Indices, Canopy Attributes and Plant Physiological Processes: What Vegetation Indices Can and Cannot Tell Us About the Landscape, Sensors, 8(4), 2136–2160, 2008.

Goetz, S. J., Bunn, A. G., Fiske, G. J. and Houghton, R. A.: Satellite-observed photosynthetic trends across boreal North America associated with climate and fire disturbance, Proc. Natl. Acad. Sci. U. S. A., 102(38), 13521–13525, 2005.

Guay, K. C., Beck, P. S. A., Berner, L. T., Goetz, S. J., Baccini, A. and Buermann, W.: Vegetation productivity patterns at high northern latitudes: a multi-sensor satellite data assessment, Glob. Change Biol., 20(10), 3147–3158, 2014.

Gutman, G. and Ignatov, A.: The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models, Int. J. Remote Sens., 19(8), 1533–1543, 1998.

Hall, F., Masek, J. G. and Collatz, G. J.: Evaluation of ISLSCP Initiative II FASIR and GIMMS NDVI products and implications for carbon cycle science, J. Geophys. Res. Atmospheres, 111(D22), D22S08, 2006.

Hall, F. G., Huemmrich, K. F., Goetz, S. J., Sellers, P. J. and Nickeson, J. E.: Satellite remote sensing of surface energy balance: Success, failures, and unresolved issues in FIFE, J. Geophys. Res. Atmospheres, 97(D17), 19061–19089, 1992.

Hansen, M. C., Stehman, S. V. and Potapov, P. V.: Quantification of global gross forest cover loss, Proc. Natl. Acad. Sci., 107(19), 8650–8655, 2010.

Holben, B. N.: Characteristics of maximum-value composite images from temporal AVHRR data, Int. J. Remote Sens., 7(11), 1417–1434, 1986.

Huete, A. .: A soil-adjusted vegetation index (SAVI), Remote Sens. Environ., 25(3), 295–309, 1988.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X. and Ferreira, L. G.: Overview of the radiometric and biophysical performance of the MODIS vegetation indices, Remote Sens. Environ., 83(1-2), 195–213, 2002.

Huete, A., Kim, H.-J. and Miura, T.: Scaling dependencies and uncertainties in vegetation index - biophysical retrievals in heterogeneous environments, in Geoscience and Remote Sensing Symposium, 2005. IGARSS ’05. Proceedings. 2005 IEEE International, vol. 7, pp. 5029–5032., 2005.

Huete, A., Didan, K., Leeuwen, W. van, Miura, T. and Glenn, E.: MODIS Vegetation Indices, in Land Remote Sensing and Global Environmental Change, edited by B. Ramachandran, C. O. Justice, and M. J. Abrams, pp. 579–602, Springer New York, 2010.

Hwang, T., Song, C., Bolstad, P. V. and Band, L. E.: Downscaling real-time vegetation dynamics by fusing multi-temporal MODIS and Landsat NDVI in topographically complex terrain, Remote Sens. Environ., 115(10), 2011.

Iio, A., Hikosaka, K., Anten, N. P. R., Nakagawa, Y. and Ito, A.: Global dependence of field-observed leaf area index in woody species on climate: a systematic review, Glob. Ecol. Biogeogr., 23(3), 274–285, 2014.

Jakubauskas, M. E., Legates, D. R. and Kastens, J. H.: Harmonic analysis of time-series AVHRR NDVI data, Photogramm. Eng. Remote Sens., 67(4), 461–470, 2001.

Jeong, S.-J., Ho, C.-H., Gim, H.-J. and Brown, M. E.: Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008, Glob. Change Biol., 17(7), 2385–2399, 2011.

Jiang, L., Kogan, F. N., Guo, W., Tarpley, J. D., Mitchell, K. E., Ek, M. B., Tian, Y., Zheng, W., Zou, C.-Z. and Ramsay, B. H.: Real-time weekly global green vegetation fraction derived from advanced very high resolution radiometer-based NOAA operational global vegetation index (GVI) system, J. Geophys. Res. Atmospheres, 115(D11), D11114, 2010.

Jiang, Z., Huete, A. R., Chen, J., Chen, Y., Li, J., Yan, G. and Zhang, X.: Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction, Remote Sens. Environ., 101(3), 366–378, 2006.

Jiang, Z., Huete, A. R., Didan, K. and Miura, T.: Development of a two-band enhanced vegetation index without a blue band, Remote Sens. Environ., 112(10), 3833–3845, 2008.

De Jong, R., de Bruin, S., de Wit, A., Schaepman, M. E. and Dent, D. L.: Analysis of monotonic greening and browning trends from global NDVI time-series, Remote Sens. Environ., 115(2), 692–702, 2011.

De Jong, R., Verbesselt, J., Schaepman, M. E. and de Bruin, S.: Trend changes in global greening and browning: contribution of short-term trends to longer-term change, Glob. Change Biol., 18(2), 642–655, 2012.

Kandasamy, S., Baret, F., Verger, A., Neveux, P. and Weiss, M.: A comparison of methods for smoothing and gap filling time series of remote sensing observations: application to MODIS LAI products, Biogeosciences Discuss, 9(12), 17053–17097, 2012.

Kang, Y., Ozdogan, M., Zipper, S. C., Roman, M. O., Walker, J., Youn Hong, S., Marshall, M., Magliulo, V., Moreno, J., Alonso, L., Miyata, A., Kimbal, B. and Loheide, S. P.: How Universal is the Relationship between Remotely Sensed Vegetation Indices and Crop Leaf Area Index? A Global Assessment, Remote Sens. Environ. Press, 2015.

Karnieli, A., Bayasgalan, M., Bayarjargal, Y., Agam, N., Khudulmur, S. and Tucker, C. J.: Comments on the use of the Vegetation Health Index over Mongolia, Int. J. Remote Sens., 27(10), 2017–2024, 2006.

Van Leeuwen, W. J. D., Orr, B. J., Marsh, S. E. and Herrmann, S. M.: Multi-sensor NDVI data continuity: Uncertainties and implications for vegetation monitoring applications, Remote Sens. Environ., 100(1), 67–81, 2006.

Marshall, M. and Thenkabail, P.: Developing in situ Non-Destructive Estimates of Crop Biomass to Address Issues of Scale in, Remote Sens., 7(1), 2015.

Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., Gao, F., Kutler, J. and Lim, T.-K.: A Landsat surface reflectance dataset for North America, 1990-2000, IEEE Geosci. Remote Sens. Lett., 3(1), 68–72, 2006.

Mayaux, P., Pekel, J.-F., Desclée, B., Donnay, F., Lupi, A., Achard, F., Clerici, M., Bodart, C., Brink, A., Nasi, R. and Belward, A.: State and evolution of the African rainforests between 1990 and 2010, Philos. Trans. R. Soc. B Biol. Sci., 368(1625), 20120300, 2013.

Monsi, M. and Saeki, T.: Über den Lichtfaktor in den Pflanzengesellschaften und seine Bedeutung für die Stoffproduktion, Jpn. J. Bot., 14, 22–52, 1953.

Moulin, S., Kergoat, L., Viovy, N. and Dedieu, G.: Global-Scale Assessment of Vegetation Phenology Using NOAA/AVHRR Satellite Measurements, J. Clim., 10(6), 1154–1170, 1997.

Mu, Q., Heinsch, F. A., Zhao, M. and Running, S. W.: Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, Remote Sens. Environ., 111(4), 519–536, 2007.

Myneni, R. ., Hoffman, S., Knyazikhin, Y., Privette, J. ., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G. ., Lotsch, A., Friedl, M., Morisette, J. ., Votava, P., Nemani, R. . and Running, S. .: Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, Remote Sens. Environ., 83(1–2), 214–231, 2002.

Nagol, J. R., Vermote, E. F. and Prince, S. D.: Effects of atmospheric variation on AVHRR NDVI data, Remote Sens. Environ., 113(2), 392–397, 2009.

Nemani, R. R., Keeling, C. D., Hashimoto, H., Jolly, W. M., Piper, S. C., Tucker, C. J., Myneni, R. B. and Running, S. W.: Climate-Driven Increases in Global Terrestrial Net Primary Production from 1982 to 1999, Science, 300(5625), 1560–1563, 2003.

Norman, J. M., Kustas, W. P. and Humes, K. S.: Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature, Agric. For. Meteorol., 77(3–4), 263–293, 1995.

O'ishi, R. and Abe-Ouchi, A.: Influence of dynamic vegetation on climate change arising from increasing CO2, Clim. Dyn., 33(5), 645–663, 2009.

Paruelo, J. M., Garbulsky, M. F., Guerschman, J. P. and Jobbágy, E. G.: Two decades of Normalized Difference Vegetation Index changes in South America: identifying the imprint of global change, Int. J. Remote Sens., 25(14), 2793–2806, 2004.

Pedelty, J., Devadiga, S., Masuoka, E., Brown, M., Pinzon, J., Tucker, C., Roy, D., Ju, J., Vermote, E., Prince, S., Nagol, J., Justice, C., Schaaf, C., Liu, J., Privette, J. and Pinheiro, A.: Generating a long-term land data record from the AVHRR and MODIS Instruments, in Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International, pp. 1021–1025., 2007.

Peischl, S., Walker, J. P., Rüdiger, C., Ye, N., Kerr, Y. H., Kim, E., Bandara, R. and Allahmoradi, M.: The AACES field experiments: SMOS calibration and validation across the Murrumbidgee River catchment, Hydrol Earth Syst Sci, 16(6), 1697–1708, 2012.

Pinzon, J. E. and Tucker, C. J.: A Non-Stationary 1981–2012 AVHRR NDVI3g Time Series, Remote Sens., 6(8), 6929–6960, 2014.

Quillet, A., Peng, C. and Garneau, M.: Toward dynamic global vegetation models for simulating vegetation–climate interactions and feedbacks: recent developments, limitations, and future challenges, Environ. Rev., 18(NA), 333–353, 2010.

Ramankutty, N., Evan, A. T., Monfreda, C. and Foley, J. A.: Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000, Glob. Biogeochem. Cycles, 22(1), GB1003, 2008.

Rao, C. R. N. and Chen, J.: Inter-satellite calibration linkages for the visible and near-infared channels of the Advanced Very High Resolution Radiometer on the NOAA-7, -9, and -11 spacecraft, Int. J. Remote Sens., 16(11), 1931–1942, 1995.

Rao, C. R. N. and Chen, J.: Post-launch calibration of the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on the NOAA-14 spacecraft, Int. J. Remote Sens., 17(14), 2743–2747, 1996.

Rocha, A. V. and Shaver, G. R.: Advantages of a two band EVI calculated from solar and photosynthetically active radiation fluxes, Agric. For. Meteorol., 149(9), 1560–1563, 2009.

Rouse, J. W.: Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. [online] Available from: http://ntrs.nasa.gov/search.jsp?R=19740022555 (Accessed 6 April 2015), 1974.

Scheftic, W., Zeng, X., Broxton, P. and Brunke, M.: Intercomparison of Seven NDVI Products over the United States and Mexico, Remote Sens., 6(2), 1057–1084, 2014.

Scheiter, S., Langan, L. and Higgins, S. I.: Next-generation dynamic global vegetation models: learning from community ecology, New Phytol., 198(3), 957–969, 2013.

Sellers, P. J.: Canopy reflectance, photosynthesis and transpiration, Int. J. Remote Sens., 6(8), 1335–1372, 1985.

Tian, F., Fensholt, R., Verbesselt, J., Grogan, K., Horion, S. and Wang, Y.: Evaluating temporal consistency of long-term global NDVI datasets for trend analysis, Remote Sens. Environ., XXX, XXX–XXX, 2015.

Tucker, C. J.: Red and photographic infrared linear combinations for monitoring vegetation, Remote Sens. Environ., 8(2), 127–150, 1979.

Tucker, C. J., Newcomb, W. W. and Dregne, H. E.: AVHRR data sets for determination of desert spatial extent, Int. J. Remote Sens., 15(17), 3547–3565, 1994.

Tucker, C. J., Pinzon, J. E., Brown, M. E., Slayback, D. A., Pak, E. W., Mahoney, R., Vermote, E. F. and El Saleous, N.: An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data, Int. J. Remote Sens., 26(20), 4485–4498, 2005.

Vermote, E., Saleous, N. E., Kaufman, Y. J. and Dutton, E.: Data pre-processing: Stratospheric aerosol perturbing effect on the remote sensing of vegetation: Correction method for the composite NDVI after the Pinatubo eruption, Remote Sens. Rev., 15(1-4), 7–21, 1997.

Wang, X., Piao, S., Ciais, P., Li, J., Friedlingstein, P., Koven, C. and Chen, A.: Spring temperature change and its implication in the change of vegetation growth in North America from 1982 to 2006, Proc. Natl. Acad. Sci., 108(4), 1240–1245, 2011.

Weiss, M., Baret, F., Block, T., Koetz, B., Burini, A., Scholze, B., Lecharpentier, P., Brockmann, C., Fernandes, R., Plummer, S., Myneni, R., Gobron, N., Nightingale, J., Schaepman-Strub, G., Camacho, F. and Sanchez-Azofeifa, A.: On Line Validation Exercise (OLIVE): A Web Based Service for the Validation of Medium Resolution Land Products. Application to FAPAR Products, Remote Sens., 6(5), 2014.

Xiao, X., Braswell, B., Zhang, Q., Boles, S., Frolking, S. and Moore III, B.: Sensitivity of vegetation indices to atmospheric aerosols: continental-scale observations in Northern Asia, Remote Sens. Environ., 84(3), 385–392, 2003.

Xin, Q., Gong, P., Yu, C., Yu, L., Broich, M., Suyker, A. E. and Myneni, R. B.: A Production Efficiency Model-Based Method for Satellite Estimates of Corn and Soybean Yields in the Midwestern US, Remote Sens., 5(11), 5926–5943, 2013.

Zeng, X., Dickinson, R. E., Walker, A., Shaikh, M., DeFries, R. S. and Qi, J.: Derivation and Evaluation of Global 1-km Fractional Vegetation Cover Data for Land Modeling, J. Appl. Meteorol., 39(6), 826–839, 2000.

Zhou, L., Tian, Y., Myneni, R. B., Ciais, P., Saatchi, S., Liu, Y. Y., Piao, S., Chen, H., Vermote, E. F., Song, C. and Hwang, T.: Widespread decline of Congo rainforest greenness in the past decade, Nature, 509(7498), 86–90, 2014.

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R. R. and Myneni, R. B.: 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 Sens., 5(2), 927–948, 2013.

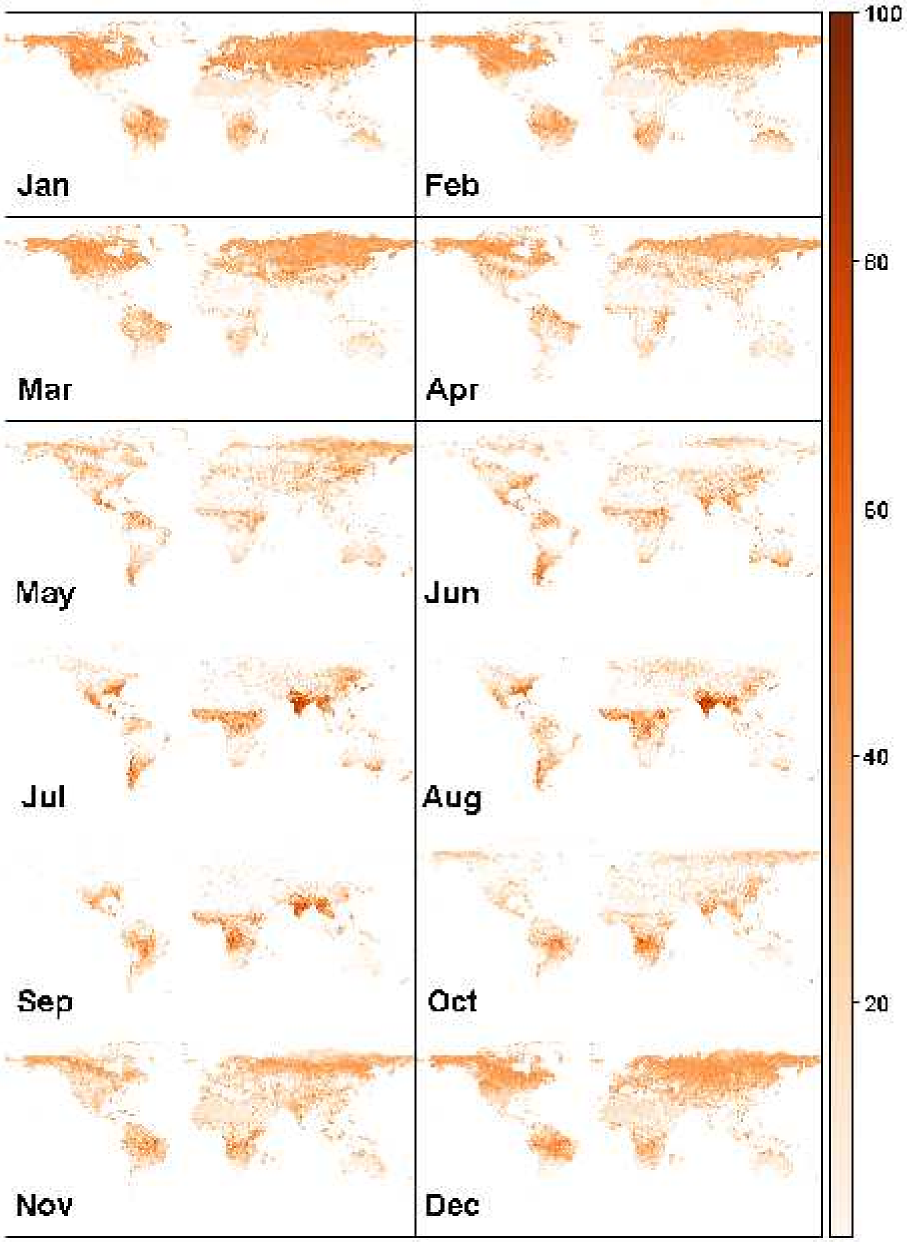

Figure 1. Percentage increase in pixels added (i.e. gaps filled) after applying the temporal filter to Vegetation Index & Phenology Lab Version 3 records.

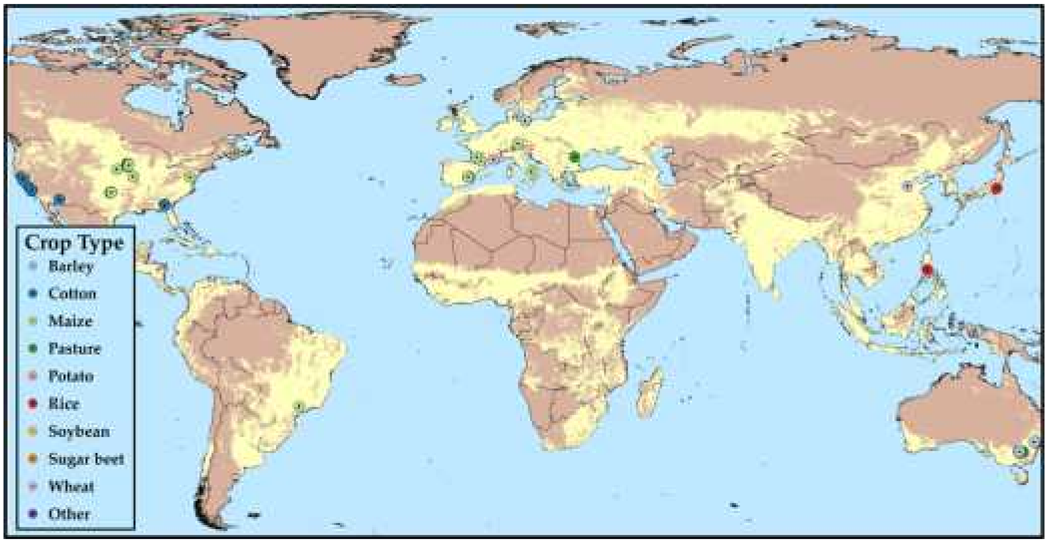

Figure 2. Sites where *in situ* (destructive or optical) measurements and Landsat Thematic Mapper/The Enhanced Thematic Mapper Plus ground reflectance data were compiled, resulting in more than 1,400 data pairs. The sites are overlaid with 1 km grid cells that contain 5% or more crop area (Ramankutty et al., 2008).

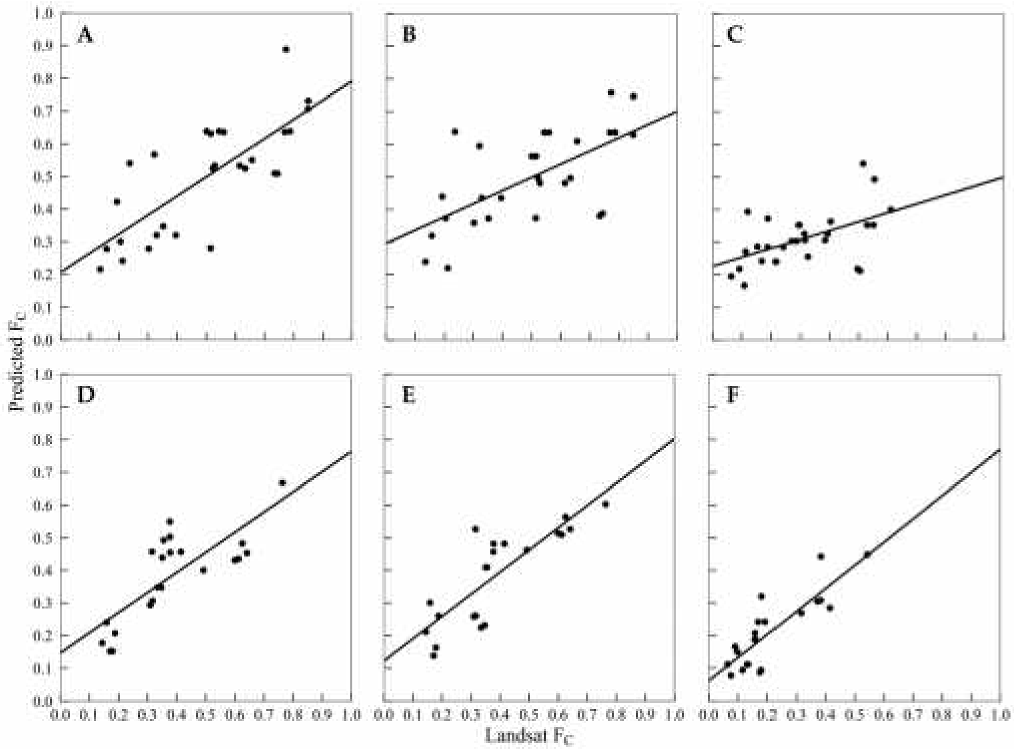

Figure 3. Scatterplots of the Fraction Absorbed of Photosynthetically Active Radiation (F_{APAR}) Landsat versus F_{APAR} for wheat **(a-c)** and pasture **(d-f)** estimated by the Global Inventory Modeling and Mapping Studies Normalized Difference Vegetation Index Version 3; Vegetation Index & Phenology Lab Version 3 Normalized Difference Vegetation Index; and Vegetation Index & Phenology Lab Version 3 Enhanced Vegetation Index 2, respectively. The solid lines represent the linear model used to downscale the vegetation record for evaluation with *in situ* leaf area index.

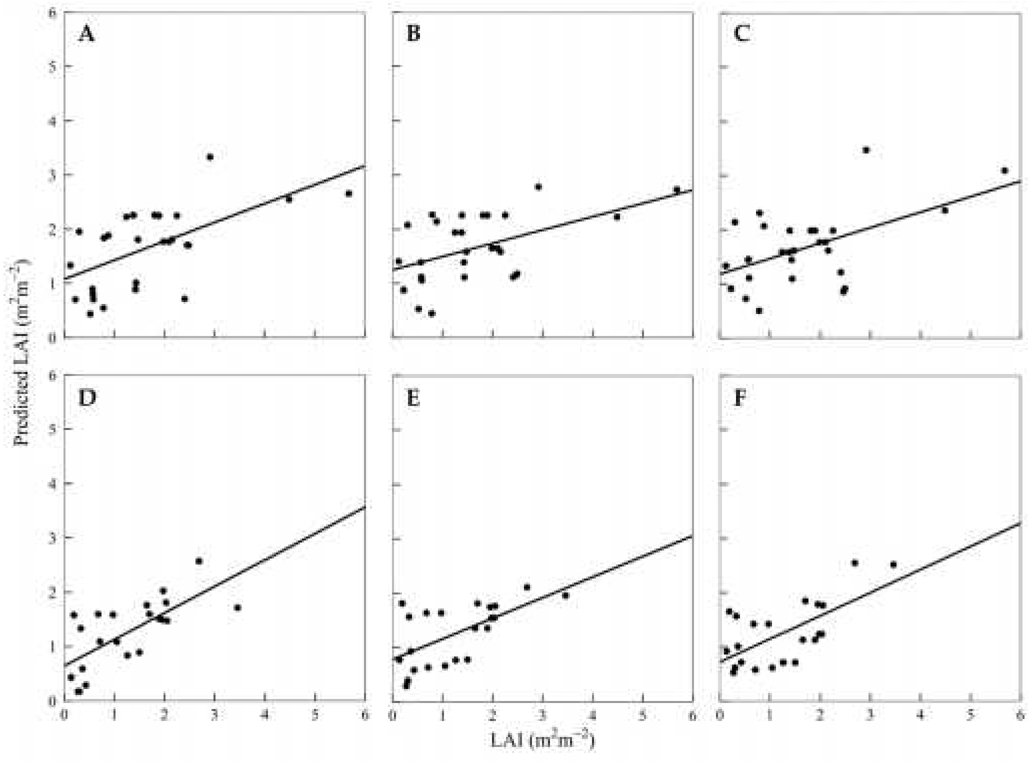

Figure 4. Scatterplots of *in situ* leaf area index for wheat **(a-c)** and pasture **(d-f)** versus corresponding Landsat resolution pixels downscaled from the Global Inventory Modeling and Mapping Studies Normalized Difference Vegetation Index Version 3; Vegetation Index & Phenology Lab Version 3 Normalized Difference Vegetation Index; and Vegetation Index & Phenology Lab Version 3 Enhanced Vegetation Index 2 datasets, respectively. The solid lines represent the best model fit.

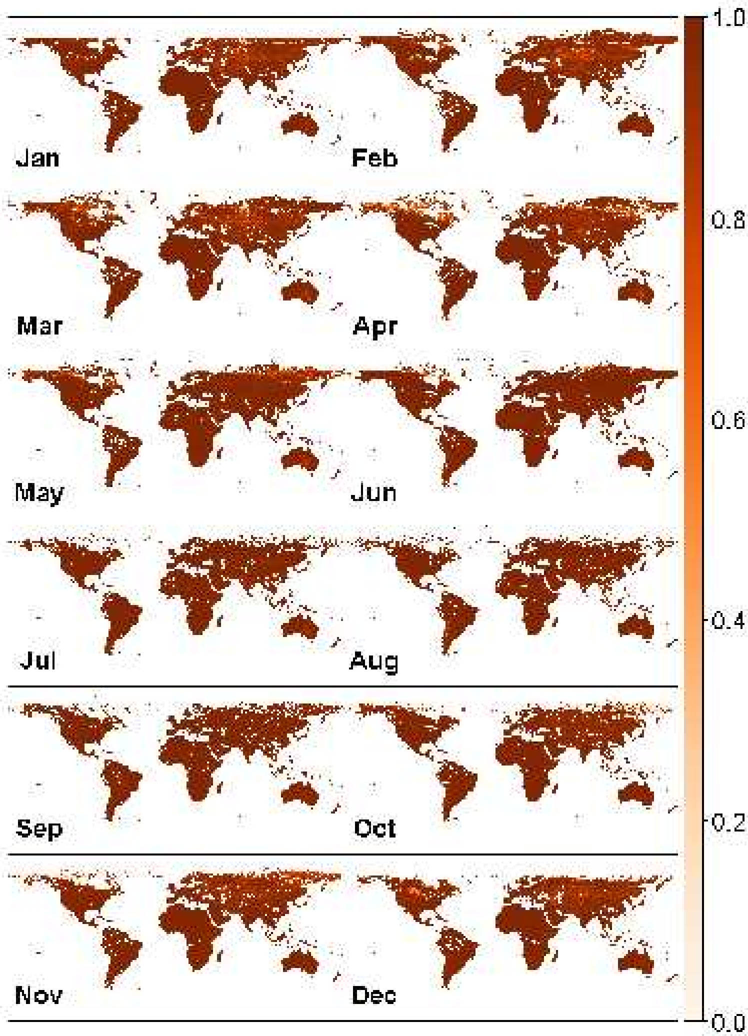

Figure 5. The coefficient of determination (R^2) on a per-pixel basis for the Vegetation Index & Phenology Lab Version 3 Normalized Difference Vegetation Index versus the Global Inventory Modeling and Mapping Studies Normalized Difference Vegetation Index Version 3. R^2 was determined using a 30-year time series of 15-day composites for each month. The images have been masked for significance ≤ 0.05 and latitudes ranging from 60°N - 60°S.

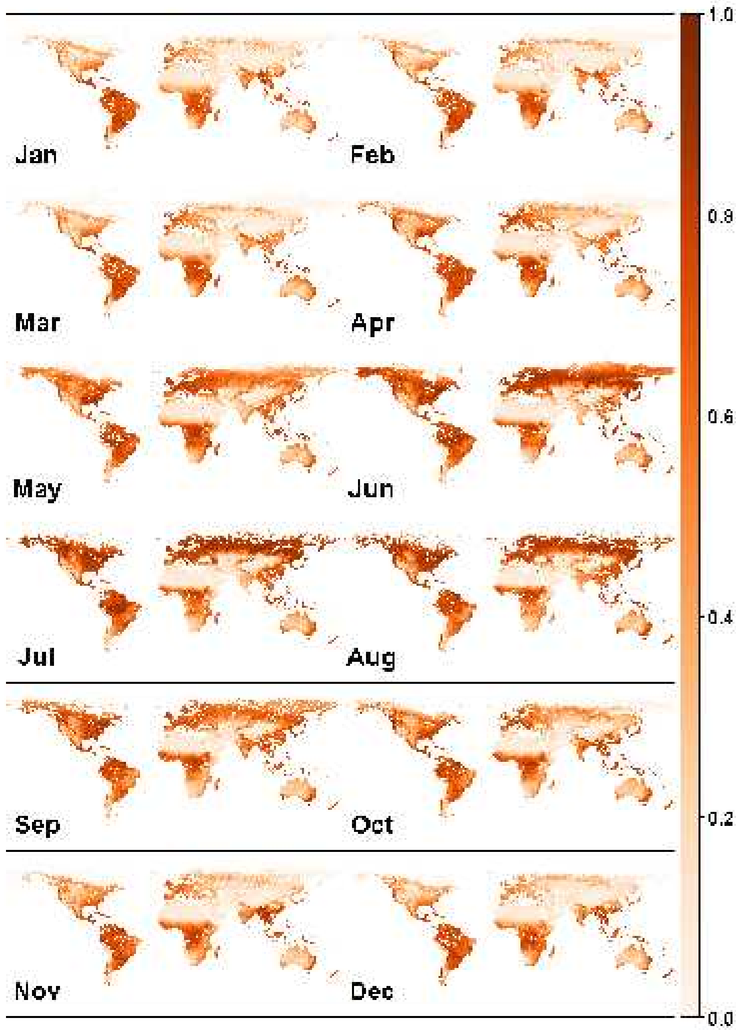

Figure 6. The slope (intercept = 0) determined from linear regresion on a per-pixel basis for the Vegetation Index & Phenology Lab Version 3 Normalized Difference Vegetation Index versus the Global Inventory Modeling and Mapping Studies Normalized Difference Vegetation Index Version 3. Slope was determined using a 30-year time series of 15-day composites for each month. The images have been masked for significance ≤ 0.05 and latitudes ranging from 60°N - 60°S.

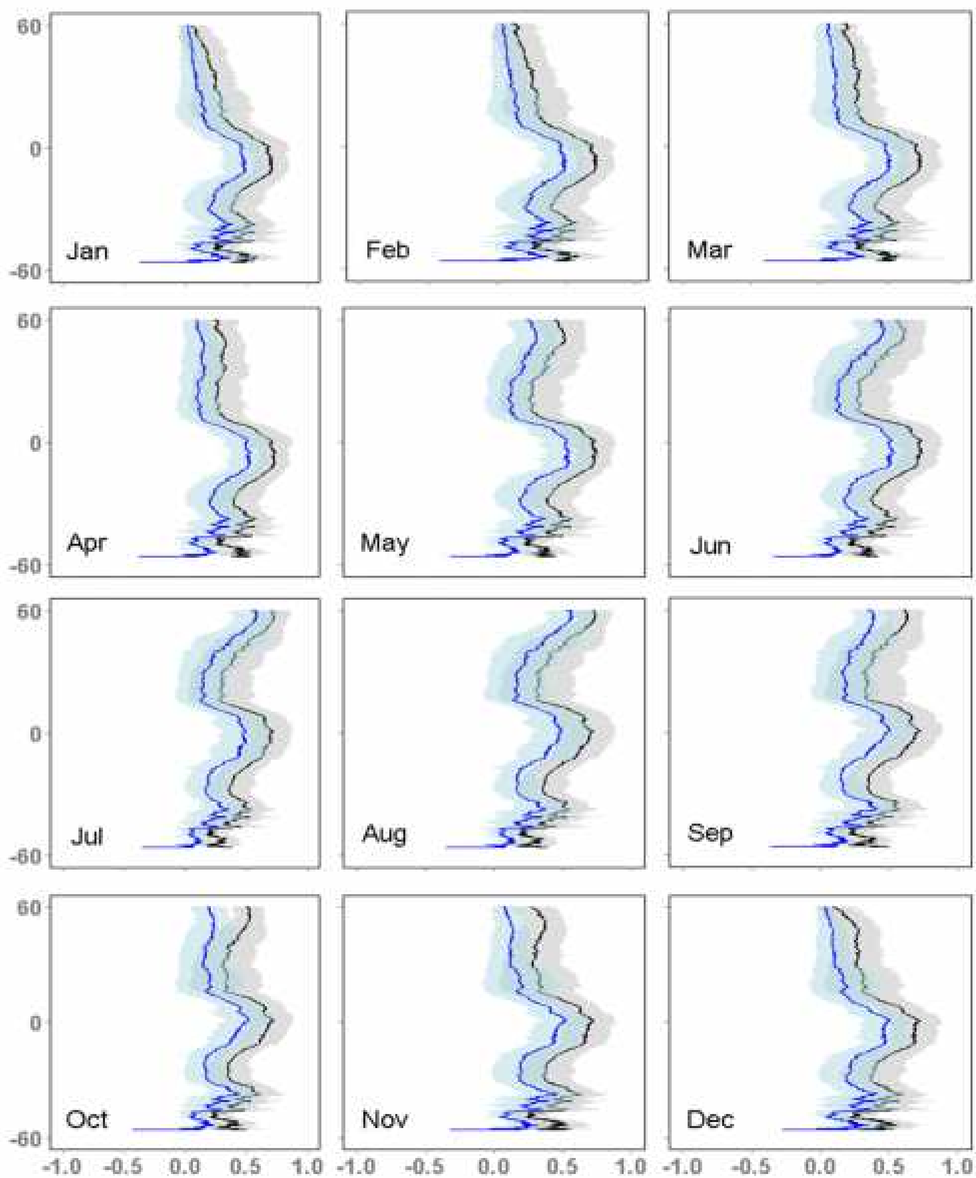

Figure 7. The latitudinal mean (solid line) and standard deviation (ribbon) of the Global Inventory Modeling and Mapping Studies Normalized Difference Vegetation Index Version 3 (blue) and Vegetation Index & Phenology Lab Version 3 Normalized Difference Vegetation Index (black) over 30 years. Values are shown from 60°N - 60°S.

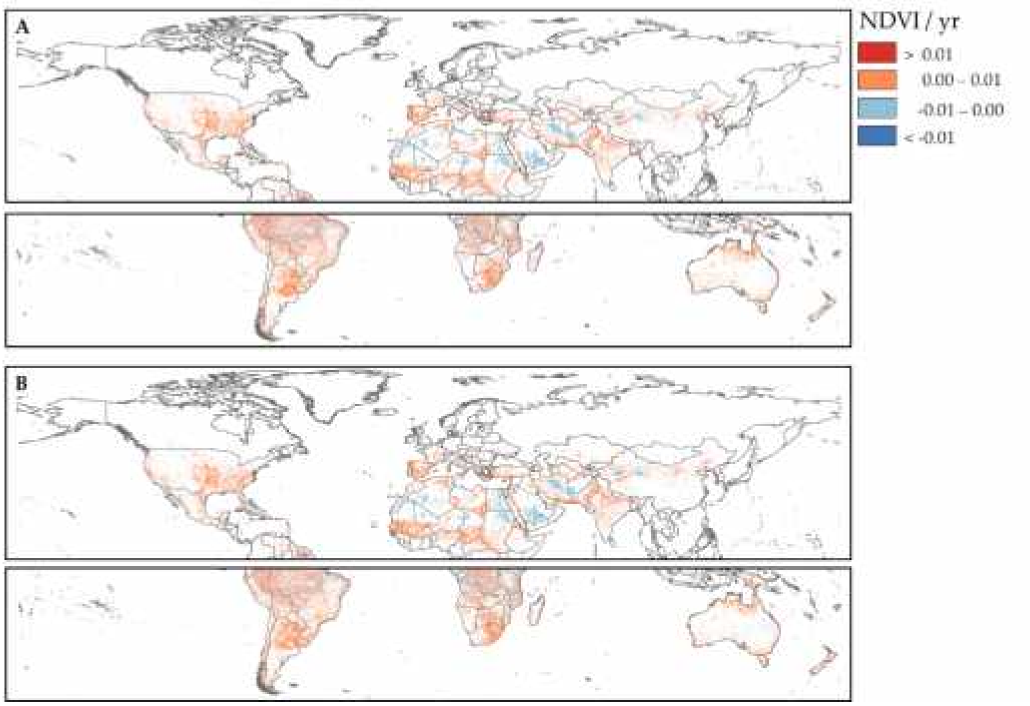

Figure 8. The change in maximum Normalized Difference Vegetation Index (NDVI) per year (yr) from the a) Global Inventory Modeling and Mapping Studies (GIMMS) and b) Vegetation Index & Phenology Lab (VIP) records. The upper panals represent the northern hemisphere (30 year change) and the lower panels represent the southern hemisphere (29 year change). The trends have been masked for significance ≤ 0.05.

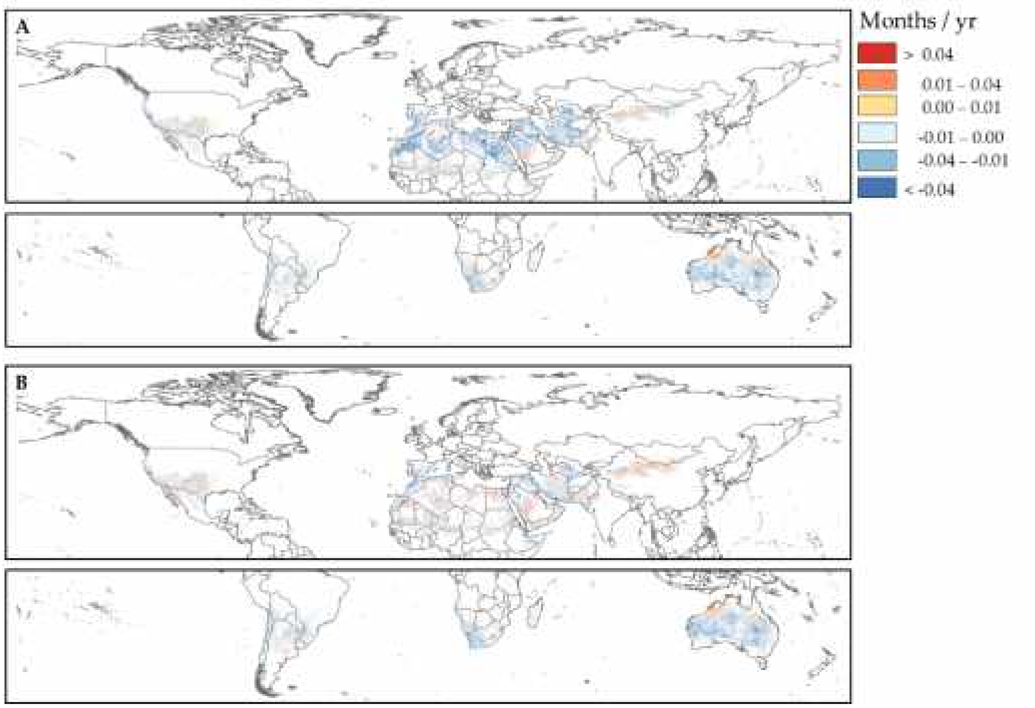

Figure 9. The change in timing of the Normalized Difference Vegetation Index (NDVI) per year (yr) from the a) Global Inventory Modeling and Mapping Studies (GIMMS) and b) Vegetation Index & Phenology Lab (VIP) records. The upper panals represent the northern hemisphere (30 year change) and the lower panels represent the southern hemisphere (29 year change). Negative values indicate earlier green-up/scenence, while positive values indicate later green-up/scenence. The trends have been masked for significance ≤ 0.05.

Table 1. Summary statistics (**R^2** = coefficient of determination, **m** = slope, **b** = intercept, **p** = significance, and **RMSE** = root-mean-square error) of the linear relationships between the Fraction of Photosynthetically Active Radiation intercepted by the canopy (F_{PAR}) estimated by Landsat Thematic Mapper or Enhanced Thematic Mapper Plus and F_{PAR} estimated by the long-term vegetation records (NDVI3g = Global Inventory Modeling and Mapping Studies Normalized Difference Vegetation Index Version 3, NDVI3v = Vegetation Index & Phenology Lab Version 3 Normalized Difference Vegetation Index, and EVI3v = Vegetation Index & Phenology Lab Enhanced Vegetation Index 2).

Crop	Product	R^2	m	b	p	RMSE
Maize	NDVI3g	0.33	0.61	0.416	<0.001	0.178
N = 98	NDVI3v	0.29	0.73	0.201	<0.001	0.183
	EVI3v	0.26	0.65	0.178	<0.001	0.163
Pasture	NDVI3g	0.62	0.72	0.106	<0.001	0.110
N = 22	NDVI3v	0.68	0.85	-0.100	<0.001	0.101
	EVI3v	0.71	0.81	-0.038	<0.001	0.071
Soybean	NDVI3g	0.40	0.82	0.146	<0.001	0.168
N = 39	NDVI3v	0.47	1.09	-0.212	<0.001	0.158
	EVI3v	0.40	0.86	0.086	<0.001	0.125
Wheat	NDVI3g	0.59	0.86	0.222	<0.001	0.148
N = 28	NDVI3v	0.40	0.84	0.058	<0.001	0.177
	EVI3v	0.27	0.74	0.096	0.004	0.140

Table 2. Summary statistics (**R^2** = coefficient of determination, **m** = slope, **b** = intercept, **p** = significance, and **RMSE** = root-mean-square error) of the relationships between *in situ* Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation intercepted by the canopy (F_{PAR}) estimated by the downscaled long-term vegetation records (NDVI3g = Global Inventory Modeling and Mapping Studies Normalized Difference Vegetation Index Version 3, NDVI3v = Vegetation Index & Phenology Lab Version 3 Normalized Difference Vegetation Index, and EVI3v = Vegetation Index & Phenology Lab Enhanced Vegetation Index 2). A logarithmic transformation was performed for soybean to meet the assumptions of normality, while the *in situ* LAI from the other crops were not transformed.

Crop	Product	R^2	m	b	p	RMSE	Transformation
Maize	NDVI3g	0.28	7.02	-1.942	<0.001	1.405	Linear
N = 98	NDVI3v	0.22	6.67	-1.695	<0.001	1.461	Linear
	EVI3v	0.21	7.87	-0.739	<0.001	1.474	Linear
Pasture	NDVI3g	0.49	4.65	-0.532	<0.001	0.665	Linear
N = 22	NDVI3v	0.38	3.90	-0.244	0.002	0.733	Linear
	EVI3v	0.43	5.46	0.097	<0.001	0.704	Linear
Soybean	NDVI3g	0.50	5.56	-3.264	<0.001	0.756	Logarithmic
N = 39	NDVI3v	0.51	5.12	-2.991	<0.001	0.753	Logarithmic
	EVI3v	0.39	6.89	-2.713	<0.001	0.838	Logarithmic
Wheat	NDVI3g	0.35	4.29	-0.482	<0.001	1.029	Linear
N = 28	NDVI3v	0.25	4.34	-0.504	0.007	1.107	Linear
	EVI3v	0.29	7.92	-0.806	0.003	1.077	Linear