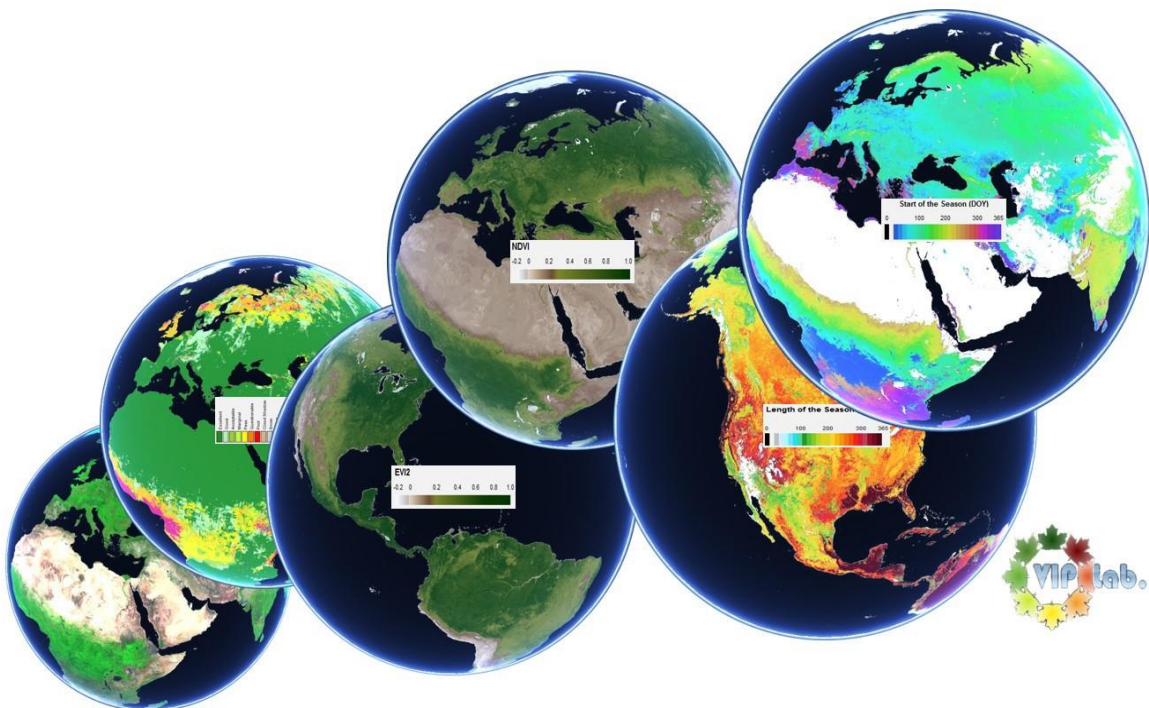# Multi-Sensor Vegetation Index and Phenology Earth Science Data Records

Algorithm Theoretical Basis Document and User Guide

Version 4.0

Kamel Didan, Armando Barreto Munoz, Tomoaki Miura, Javzandulm Tsend-Ayush, Xiaoyang Zhang, Mark Friedl, Josh Gray, Willem Van Leeuwen, Jeffrey Czapla-Myers, Calli Jenkerson, Tom Maiersperger, David Meyer

Vegetation Index and Phenology Lab.
vip.arizona.edu
The University of Arizona
-Dec. 2015-

Science Team

The Vegetation Index and Phenology (**VIP**) Earth Science Data Records Project is a multi-institution effort lead by the University of Arizona and funded by NASA **MEaSUREs** (Making Earth System Data Records for Use in Research Environments) Cooperative Agreement # NNX08AT05A.

The University of Arizona

- Kamel Didan (PI) [didan@email.arizona.edu]
- Armando Barreto Munoz
- Willem Van Leeuwen
- Jeffrey Czapla-Myers

Boston University

- Mark Friedl
- Josh Gray

South Dakota State University

- Xiaoyang Zhang

University of Hawaii-Manoa

- Tomoaki Miura
- Javzandulm Tsend-Ayush

USGS-EROS and LP-DAAC

- Calli Jenkerson
- Tom Maiersperger
- David Meyer

Note: This is a live document that may and will change to update the status of the product suite and any necessary corrections. All updates will be posted on the website [vip.arizona.edu].

Table of Contents

List of Tables

List of Figures

Foreword

Several recent events have highlighted the need for long-term satellite observations of the Earth. Climate change is expected to significantly impact the functioning of terrestrial ecosystems and thereby alter fluxes of energy, mass and momentum between the land surface and the atmosphere. There has been partial success in closing this feedback loop of climate–vegetation–interactions, however, the accurate characterization of the land surface vegetation and its seasonal timing and annual sequence of events, is crucial to this effort, and to link land surface-atmosphere interactions in models. Vegetation phenology is a characteristic property of ecosystem functioning and predictor of ecosystem processes. Numerous studies have demonstrated that climate processes operating at seasonal and interannual time scales (e.g. ENSO) are identifiable in the phenology of vegetation.

To accurately understand current trends and anomalies we must have a better characterization of long term normal. Land surface vegetation and all its derivative measurements are a central component of Earth observing and an integrator of climate and anthropogenic drivers. While land surface vegetation is measured by various direct and indirect observational metrics, the vegetation index time series product is by far the most successful data record from the various Earth Observing Systems. The international Committee on Earth Observing Satellites (CEOS) convened a Vegetation Index/Phenology workshop, in summer 2006, to bring together producers and users of global VI time-series data and discuss the current state of global VI records, their accuracy, and methods used to quantify their uncertainties in phenology and long-term land surface process studies. The community recognizes the value of the long term AVHRR-NDVI data record and the importance of backward compatibility so that scientific analyses can utilize the AVHRR record. Any reprocessing of the AVHRR record should consider steps to integrate important and significant improvements made with new sensors and algorithms to allow forward compatibility with newer sensors and products. Reprocessing of AVHRR should consider other VIs-such as the Enhanced Vegetation Index -wherever possible and should quantitatively address the uncertainty in these records.

In addition, an important new development in the field of phenology in the United States is the creation of the National Phenology Network (NPN). The US-NPN began as a grass roots, interdisciplinary effort involving botanical gardens, academia, and government agencies with the goal of systematically collecting and analyzing phenological data and making use of satellite observations. Providing satellite based observations of phenology is a key component of these efforts and will further enhance our observational capabilities and data holdings.

The science data records discussed in this document and the methods proposed for their consistent generation and characterization will contribute greatly toward these objectives.

1.   Introduction

One of the primary interests of observing the Earth surface with global imagers is to characterize and measure the role of vegetation in large-scale global processes with the key goal of understanding how the Earth functions as a system. This requires an understanding of the global distribution of vegetation types as well as their biophysical, functional, structural properties, and spatial/temporal variations. While many direct spectral images interpretation methods exist the simpler method of spectral bands ratioing, or Vegetation Indices (VI), remains one of the most robust empirical methods for characterizing land surface vegetation health and activity (Huete et al. 2002, Tucker et al. 2005). Vegetation indices are designed to enhance the vegetation reflected signal from measured spectral responses by making use of the distinctive soil-vegetation characteristic in the red-edge area of the spectrum. Vegetation indices combine two (or more) spectral bands in the **red (0.6 - 0.7 μm)** and **NIR wavelengths (0.7-1.1 μm)** regions (Tucker 1979). Vegetation indices time series inform us about the status of vegetation health during the growing season and as it changes in response to environmental, climate, and anthropogenic drivers. Time series measures of vegetation index have been shown highly correlated with flux tower photosynthesis measurement and integrate the response of vegetation to change in environmental factors providing valuable information to global change research.

A recent development of the study of land surface vegetation with remote sensing time series data is the characterization of vegetation growing season or phenology. While phenology is the study of change of all living things over time, in this context phenology is the study of vegetation change over time using remote sensing data and tools (Beaubien, et al., 2003). Because vegetation phenology affects terrestrial carbon cycle across a wide range of ecosystem and climate regimes (Baldocchi et al., 2001; Churkina et al., 2005; Richarson et al., 2009), accurate information related to phenology is important to studies of regional-to-global carbon budgets. The presence of leaves also influences land surface albedo (Moore et al., 1996; Ollinger et al., 2008) and exerts strong control on surface radiation budgets and the partitioning of net radiation between latent and sensible heat fluxes (Chen and Dudhia, 2001; Yang et al., 2001). Thus, the phenological dynamics of vegetated ecosystems influence a host of eco-physiological processes that affect hydrologic processes (Hogg et al., 2000), nutrient-cycling, (Cooke and Weih, 2005), and land-atmosphere interactions (Heimann et al., 1998).

Many data sets related to plant growing season have been collected at specific sites or in networks focused on individual plants or plant species, still remote sensing provides the only way to observe and monitor phenology at global scale and at consistent and regular intervals. Satellite phenology encompasses the analysis of the timing and rates of vegetation growth, senescence, and dormancy at seasonal and interannual time scales. To that end vegetation indices, which capture the aggregate functioning of a canopy (Asrar et al., 1984), are the most robust and are widely used proxies for extracting phenology information.

1.1.   Motivation and Background: From Science to Earth Science Data Records

Climate change is expected to significantly impact the functioning of terrestrial ecosystems and thereby alter fluxes of energy, mass and momentum between the land surface and the atmosphere (Melillo *et al.*, 1996; Watson *et al.*, 1996; Mintz, 1984; Dickinson & Henderson-Sellers, 1988; Rowntree, 1988; Bonan et al., 1992). There has been limited success in closing this feedback loop of climate–vegetation– interactions, however, the accurate characterization of land surface phenology, i.e., the seasonal timing and annual sequence of events in plant life (Fig. 1), is crucial to

this effort, and to link land surface – atmosphere interactions in models (Claussen, 1994). Vegetation phenology is a characteristic property of ecosystem functioning and predictor of ecosystem processes. Numerous studies have demonstrated that climate processes operating at seasonal and interannual time scales (e.g., ENSO) are identifiable in the phenology of vegetation (Braswell et al., 1996; Asner and Braswell, 2000, Myneni et al. 1997).

Recent findings indicate that the effects of climate change are manifested in landscape phenology (Randerson et al., 1999), hence this has emerged as a key area of research in biosphere-atmosphere interactions, climate change, and global change biology. Shifts in phenology depict an integrated vegetation response to environmental change and influence local biogeochemical processes, including nutrient dynamics, photosynthesis, water cycling, soil moisture depletion, transpiration, and canopy physiology (Reich & Borchert 1988; Herwitz 1985). Knowledge of phenologic variability and the environmental conditions controlling their activity are further prerequisite to inter-annual studies and predictive modeling of land surface responses to climate change (Myneni et al., 1997; Shabanov et al., 2002; White et al., 2002, Huete et al., 2006, Saleska et al. 2007, Huete et al, 2008, Keeling 1996a, 1996b). With major shifts in global temperature and precipitation patterns anticipated (ICCP, 2006), there is increased concern on how land surface phenology will change in response to global warming, land cover change, and shifts in land use activities (Schwartz & Reed, 1999; deBeurs and Henebry, 2005; Cochrane et al., 1999; Gedney & Valdes, 2000; Houghton et al., 2000; Lambin et al., 2003).
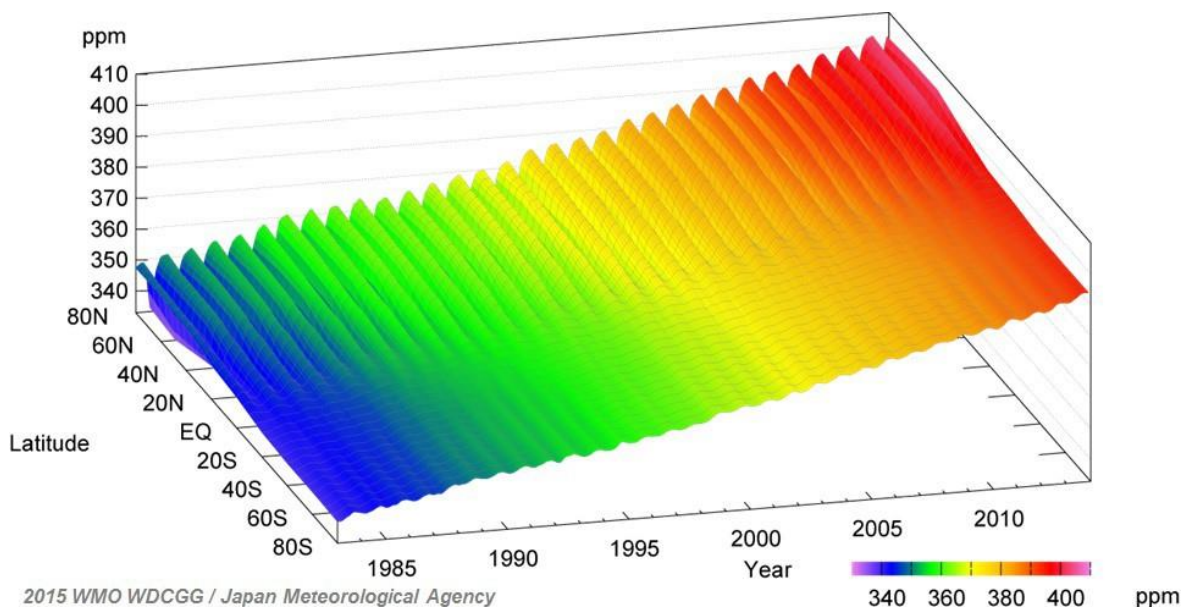
Figure 1. Mean distribution of atmospheric CO2 by time and latitude (2015 WMO WDCGG/Japan Meteorological Agency http://ds.data.jma.go.jp/ghg/kanshi/ghgp/co2_e.html). The noticeable difference between the northern and southern hemisphere is driven by ve driven by vegetation phenology, volume and timing.

Satellite vegetation indices (VI's) have played a major role in monitoring seasonal vegetation dynamics (Henderson-Sellers, 1993 and 1995) and interannual comparisons of vegetation activity. Satellite studies using vegetation index time series seasonal profiles have shown how broad-scale changes in land use and land cover change affect land surface phenology (White et al., 2002, 2009). The temporal profile of the normalized difference vegetation index (NDVI) has been shown to depict phenologic events such as, length of the growing season, peak greenness, onset of greenness, and

leaf turnover or 'dry-down' period and the time integral of the VI over the growing season has been correlated with NPP/GPP (Running and Nemani, 1988; Prince, 1991; Justice et al., 2000; Goward et al., 1991; Tucker and Sellers, 1986; Huete et al., 2008). There is evidence from satellite data that the phenology of key biomes is changing in response to shifts in climate (Myneni et al., 1997; Keeling et al., 1996; White et al., 2002; Bogaert et al., 2002; Jia et al., 2003; Huete et. al., 2006 & 2008, Saleska et al., 2007), e.g., Myneni et al. (1997) used a 10 year AVHRR-NDVI data record of northern Boreal forests to show a warming trend, whereby the length of the growing season had increased by nearly 2 weeks. Whether these trends will persist, change direction, or disappear altogether requires accurate observation and the compilation of long term data records. The efforts described by this document will detail how a long term multi-sensor data record about vegetation and phenology will be assembled, generated, and characterized.

1.2. Need for Long Term Data Records

Whereas single mission or sensor specific measurements of vegetation index and phenology exist, the length of these records is usually limited due to the mission life expectancy usually being few years, engineering and technological changes which necessitates new designs and improvements, and changes in data processing methods and approaches which render the data undesirable. In practice these limitations impose a restriction on the data usefulness in particular when addressing long term phenomenon and trends because they lack representation, or in statistical context they cannot support the generation of an accurate and representative long term normal. Extending these records beyond the short life time framework of the sensor has been both a goal and a challenge.

The data records discussed in this user guide were proposed within the framework of NASA's Making Earth System data records for Use in Research Environments project (MEaSUREs). In this project we developed two global data records about vegetation index and phenology. These records were generated from multiple sensors spanning the AVHRR and MODIS eras. They provide the longest and most consistent satellite based measurement of land surface vegetation. These records were developed to meet scientific community needs for consistent, global, and multi-decadal satellite-derived data of land surface vegetation health and dynamic. The products are based on standard science algorithms for vegetation index and land surface phenology. In this document we provide an overview of the science supporting these product records, followed by a description of the product algorithms and records specifications.

2. Science Background

Two key Earth Science Data Records (ESDR) identified in a NASA white paper on the Vegetation Index (Huete et al, 2006) and Phenology (Friedl et al., 2006) are the goal of this effort. Both of these data records are now standard products generated from MODIS since 2000. The Normalized Difference Vegetation Index (NDVI) and a backward compatible version of the Enhanced Vegetation Index (EVI), called EVI-2, and the Vegetation Phenology product were generated through backward extension to the AVHRR data record and forward compatibility with the Visible Infrared Imager/Radiometer Suite (VIIRS) sensor, that is part of the suite of sensors in the Joint Polar Satellite System (JPSS) mission (Welsch et al., 2001).

The theoretical basis for empirical-based vegetation indices is derived from examination of typical spectral reflectance signatures of leaves. The reflected energy in the visible is very low as a result of high absorption by photosynthetically active pigments, with maximum absorption values in the blue (470 nm) and red (670 nm) wavelengths. Nearly all of the near-infrared radiation (NIR) is scattered (reflected and transmitted) with very little absorption, in a manner dependent upon

the structural properties of a canopy (LAI, leaf angle distribution, leaf morphology). As a result, the contrast between red and near-infrared responses is a sensitive measure of vegetation amount, with maximum red−NIR differences occurring over a full canopy and minimal contrast over targets with little or no vegetation (Fig. 2). For low and medium amounts of vegetation, the contrast is a result of both red and NIR changes, while at higher amounts of vegetation, only the NIR contributes to increasing contrasts as the red band becomes saturated due to chlorophyll absorption.
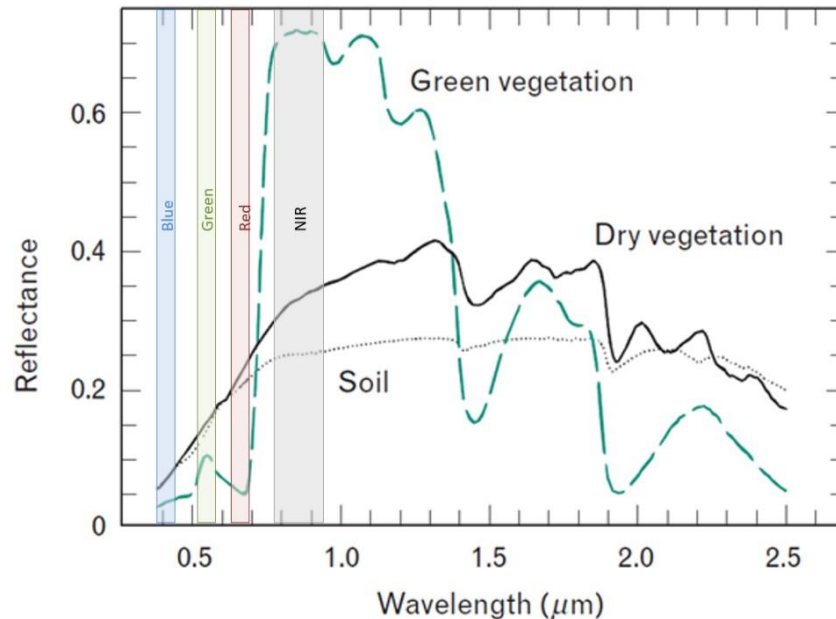
Figure 2. Plants absorb and reflect light differently depending on the wavelength and plant health status. The photosynthetic process absorbs most of the visible light (blue-red region) and vegetation reflect much of the near-infrared (NIR). These differences permit the separation of healthy from stressed plants and/or other objects.

The red-NIR contrast can be quantified through the use of ratios (NIR/red), differences (NIR−red), weighted differences (NIR−k•red), linear band combinations (x_1•red+x_2•NIR), or a hybrid combination. Vegetation indices are measures of this contrast and thus are integrative functions of canopy structural (%cover, LAI, LAD) and physiological (pigments, photosynthesis) parameters.

2.1. Vegetation Index ESDR Algorithms

2.1.1. Normalized Difference and Enhanced Vegetation Indices

Spectral vegetation indices are among the most widely used satellite data products providing key measurements for climate, phenology, hydrologic, and biogeochemical studies, and land cover/land cover change detection. There is currently a consistent NDVI record extending for more than 3 decades from the NOAA AVHRR series (Gutman et al., 1995), which have contributed significantly to the advancement of Earth System Science, in particular to global biome, agricultural primary production; interannual fluctuations and impacts of ENSO and other climatic disturbances, especially droughts, on primary production; phenology; and climate change and variability. Compared with other land products, and due to their simplicity, VI's are more readily fused across sensor systems facilitating an underlying need to ensure continuity of critical data sets to study climate-related processes. Recent cross-sensor analyses and studies have shown the potential to empirically fuse medium and coarse resolution NDVI measurements from new and advanced sensor

systems (MODIS, SPOT-VEGETATION (VGT), SeaWiFS, etc...) in order to extend the existing long-term NDVI data record (Tucker et al., 2005, Brown et al., 2006). This has been accomplished by converting one sensor-specific time series into another, either by degrading newer data to the AVHRR data record or through processing improvements of the older data, e.g., implementation of better atmosphere corrections to AVHRR (DeFelice, et al., 2003; Vermote et al., 1995, 2006), or by simple statistical and correlative analyses.

The NDVI is a normalized transform of the NIR to red reflectance ratio, ρ_{NIR}/ρ_{red}, designed to standardize VI values to between [−1 and +1], it is expressed as:

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \quad (1)$$

As a ratio, the NDVI has the advantage of minimizing certain types of band-correlated noise (positively-correlated) and influences attributed to variations in direct/diffuse irradiance, clouds and cloud shadows, sun and view angles, topography, and atmospheric attenuation. Ratioing can also reduce, to a lesser degree, calibration (Rao et al., 1994; Vermote et al., 1994) and instrument-related errors. The extent to which ratioing can reduce noise is dependent upon the correlation of noise between red and NIR responses and the degree to which the surface exhibits Lambertian behavior.

The main disadvantage of ratio-based indices tends to be their non-linearities exhibiting asymptotic behaviors, which leads to insensitivities to vegetation variations over certain land cover conditions. Ratios also fail to account for the spectral dependencies of additive atmospheric (path radiance) effects, canopy-background interactions, and canopy bidirectional reflectance anisotropies, particularly those associated with canopy shadowing.

The biophysical performance of satellite VI measures of greenness has been consistently tested and proved useful and well correlated with continuous flux tower measurements of photosynthesis (Huete et al., 2006; Xiao et al., 2005, Rahman et al., 2005), which provide valuable information about the carbon cycle, phenology, and the seasonal and inter-annual changes in ecosystems. An accurate depiction of seasonal vegetation dynamics is a desired prerequisite for accurate ecosystem modelling, and improves confidence in ESDR/CDR products and model capabilities to predict longer term, inter-annual vegetation responses to climate variability. Comparisons of temporally aggregated flux tower measures of photosynthesis with satellite VI measures of greenness have shown a strong seasonal correspondence with the Enhanced Vegetation Index (EVI) from MODIS and SPOT-VGT sensors (Xiao et al., 2004, 2005; Rahman et al., 2005; Sims et. al., 2006, Huete et al. 2008). An example of this tight coupling at the Harvard Forest site is shown in Fig. 3. In the case of NDVI, there is some saturation and an overestimation of GPP. MODIS and SPOT-VGT EVI were also shown to depict phenology cycles in dense Amazon rainforests for the first time, confirmed by a strong linear and consistent relationship between seasonal EVI and tower-calibrated GPP measurements of carbon fluxes in both intact rainforest and forest conversion to pasture/agriculture sites in the Amazon (Huete et al., 2006; Xiao et al., 2005).
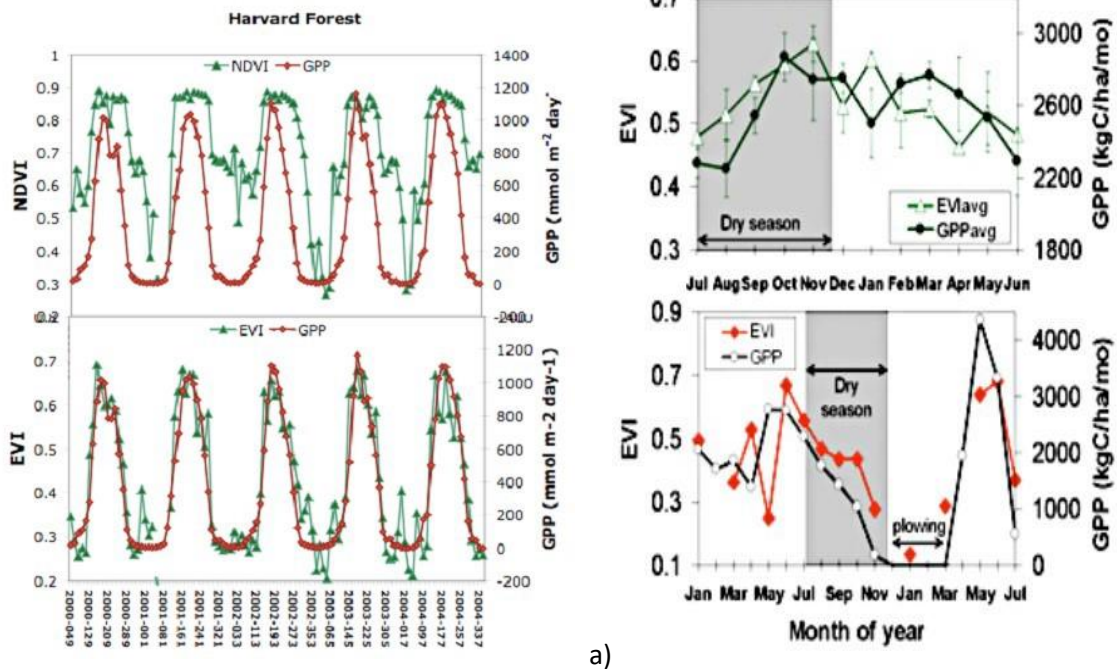
Figure 3. MODIS and SPOT VGT EVI are consistent in their phenological depiction of temperate and tropical ecosystems, providing in-situ based methods for assessment of VI performance and capabilities. a) 16-day MODIS VI's plotted with in-situ 16-day GPP flux measures at Harvard forests. b) Seasonal correspondence of MODIS EVI with tower flux measures of GPP in both intact rainforest (top) and forest conversion to pasture/agriculture (bottom). Huete et al., 2006

2.1.2.   Development of 2-band EVI (EVI2)

While the utility of the NDVI has been well established in climate science, one major weakness is its nonlinear behavior and saturation in high biomass vegetated areas (Huete et al., 2002; Ünsalan & Boyer, 2004; Gitelson, 2004; Vaiopoulos et al., 2004). Reduction of saturation effects and improved linearity adds to the observed accuracy in estimating biophysical parameters from the VI values and provides a mechanism for multi-sensor (resolution) scaling of VI values. The enhanced vegetation index (EVI) was developed to optimize the vegetation signal with improved sensitivity in high biomass regions and improved vegetation monitoring through a de-coupling of the canopy background signal and a reduction in atmosphere influences (Running et al., 1994, Huete et al. 2002).

To minimize the impact of turbid atmosphere on the VI one can use the difference in blue and red reflectances as an estimator of the atmosphere influence level. This concept is based on the wavelength dependency of aerosol scattering cross atmosphere sections. In general, the scattering cross section in the blue band is larger than that in the red band. When the aerosol concentration is higher, the difference in the two bands becomes larger. This information is used to stabilize the index value against variations in aerosol concentration levels. EVI incorporates this atmospheric resistance concept as in the Atmospheric Resistant Index (ARVI, Kaufman et al., 1995), along with the removal of soil-brightness induced variations in VI as in the Soil Adjusted Vegetation Index (SAVI, Huete, 1988). The EVI additionally decouples the soil and atmospheric influences from the vegetation signal by including a feedback term for simultaneous correction (Huete et al., 1994). The 3-band EVI is expressed as:

7

$$EVI = G \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + C1 * \rho_{red} - C2 * \rho_{Blue} + L} \quad (2)$$

Where ρ_x are the full or partially atmospherically corrected (for Rayleigh scattering and ozone absorption) surface reflectances; *L* is the canopy background adjustment that addresses nonlinear, differential NIR and red radiant transfer through a canopy (based on Beer's law), and C_1, C_2 are the coefficients of the aerosol resistance term, which uses the blue band to correct for aerosol influences in the red band (Huete et al., 2002). The coefficients adopted for MODIS EVI are L=1, C_1=6, C_2=7.5, and G (gain factor) =2.5.

EVI has been used recently in a wide variety of studies, including those on land cover/land cover change (Wardlow et al., 2007), estimation of vegetation biophysical parameters (Chen et al., 2004; Houborg et al., 2007), phenology (Zhang et al., 2003, 2006; Xiao et al., 2006; Ahl et al., 2006, Huete et al., 2008), Evapotranspiration (Nagler et al., 2005), biodiversity (Waring et al., 2006), and the estimation of gross primary production (GPP) (Rahman, et al., 2005; Sims et al., 2006).

Recent cross-sensor studies have shown the feasibility of NDVI and EVI translation across several sensors systems (Gallo et al., 2005; Miura et al., 2006; Brown et al., 2006, Huete et al., 2006). EVI extension, however, is limited to only sensors that carry a blue channel, which includes SPOT-VGT, SeaWiFS, and other instruments. In contrast to the red and NIR bands, sensor-dependent blue channels are generally not as compatible and often do not overlap, e.g., the MODIS (459-479 nm), MERIS-blue (440-450 nm), and VIIRS-blue (478-498 nm) channels do not overlap, a spectral issue that restricts the compatibility of cross-sensor EVI values. Thus, it is recommended that cross-sensor algorithms should be based on VIs without a blue band (Fensholt et al., 2006, Jiang et al., 2008).

Since the blue band in the EVI does not provide additional biophysical information about vegetation properties, rather is aimed at reducing noise and uncertainties associated with highly variable atmospheric aerosols, a 2-band adaptation of EVI was developed to be compatible with EVI (Huete et al., 2006, Jiang et al., 2008). An earlier version of the 2-band EVI (EVI2) was used as the "backup algorithm" for MODIS EVI product for cases when the blue band yields problematic VI values, mainly over dense snow, or pixel with extensive subpixel clouds. The EVI2 remains functionally equivalent to the EVI, although slightly more prone to aerosol noise, which is becoming less significant with continuing advancements in atmosphere correction (Vermote et al., 2002, Lyapustin et al., 2012).

The EVI2 is based on a linearization method and by geometrical analysis of spectral angles in the red-near infrared reflectance space (Fig. 4).
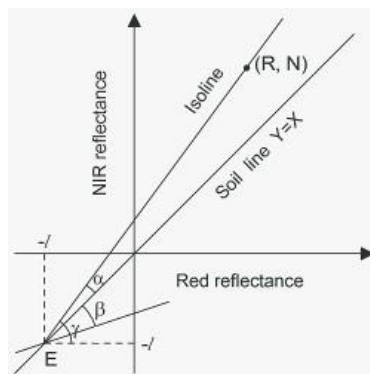
Figure 4. The isolines of the EVI/SAVI and their angles in red-NIR reflectance space

A linearized vegetation index (LVI) comparable to the EVI is obtained by adjusting the constant angle $\pi/4$ to a variable angle β, or soil background adjustment factor,

$$LVI(\beta) = \tan\left[arctan\left(\frac{SAVI}{1+L}\right) + \beta\right] \quad (3)$$

Where: β describes a line across E deviating from the soil line in clockwise direction in Fig. 4. The LVI value of the soil line, Y=X, (LVI_0) is, $LVI_0 = \tan(\beta)$, which is described as,

$$\begin{aligned} LVI &= G'\left[\tan(\alpha + \beta) - \tan\beta\right] \\ &= G\frac{(N-R)}{N+R\tan(\pi/4 + \beta) + L(1 - \tan\beta)} \end{aligned} \quad (4)$$

Where a gain factor, G', is multiplied in order to maintain the amplitude of the LVI as that of the EVI,

$$G = \frac{G'\sec^2\beta}{(1 - \tan\beta)} \quad (5)$$

With optimal β and G, the differences between the LVI values and the EVI values would be very small when atmospheric effects are insignificant and this optimal LVI is used as the 2-band EVI, i.e. EVI2 (Jiang et al., 2008). For a given combination of L and β, there is a single, optimal G that minimizes mean absolute difference (MAD) between EVI and EVI2, which results in G = 2.5 similar to the standard 3-band EVI, and the optimal parameter values for the EVI2 equation are,

$$EVI2 = 2.5\frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + 2.4*\rho_{red} + 1} \quad (6)$$

The resulting relationship between EVI and EVI2 show their very close correspondence for the entire range of values (Fig. 5). The coefficient of determination between EVI and EVI2 is high (R^2=0.9986) with the Mean Absolute Difference (MAD) of 0.00346 reflectance units. It is important to note that because the 2-band EVI lacks the blue band it becomes prone to atmosphere contamination, although with modern atmosphere correction this issue is minimal, while maintaining the other advantages of EVI, being the minimization of background variation and the additional canopy sensitivity.
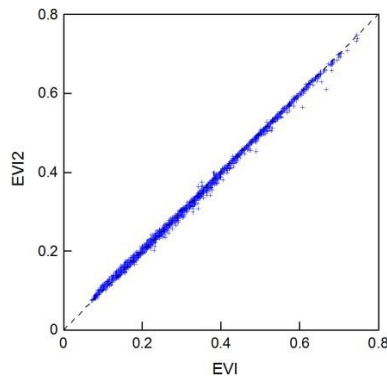
Figure 5. Relationship between EVI and EVI2 with MODIS data from 40 test validation sites (Jiang et al., 2008).

2.1.3. Continuity Algorithm

In order to generate a sensor independent value for any measurement one of three things must exist: 1) Same sensor specifications (e.g. MODIS on Terra or Aqua), 2) A sensor independent formulation (easier for a physical entity), or 3) a transfer function capable of translating across two or more sensors (Tsend-Ayush, et al, 2010a, 2010b). Since the data for this work derive from sensors with very different band passes (Fig. 6) we are left with option 3 and a continuity algorithm is required to translate the data.
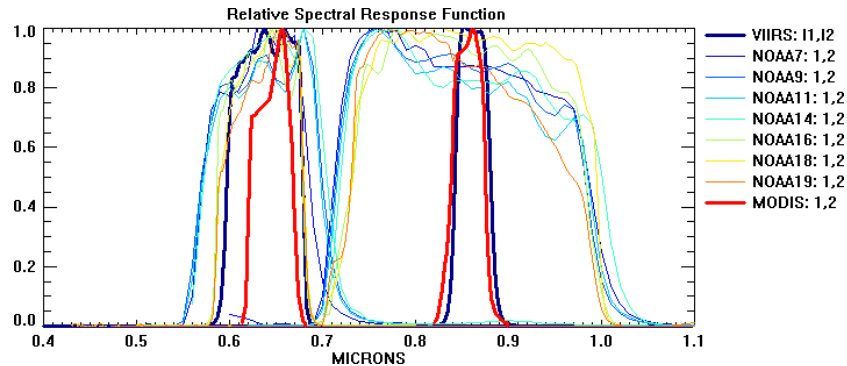
Figure 6. VIS and NIR bands spectral response functions for AVHRR, MODIS and VIIRS

Whereas many of the continuity and translation algorithm efforts to date attempt to use a quasi-physically based approach that try to address the continuity at the band pass difference level (Yoshioka et al. 2005, Steven et al 2003; Tsend-Ayush, et al., 2010a, 2010b, Trishchenko et al, 2002) here we proposed and have developed a hybrid per-pixel simpler method based on the statistical correlation analysis across the two sensors (AVHRR and MODIS). In theory this method requires that only data from simultaneous observation be used so they can be compared, however in practice the approach is developed using data from a short period of transition from AVHRR to MODIS. For this we limited our data to the periods 1995-1999 and 2000-2004, which provides long enough record while limiting the impact of change. Many of the proposed methods require a bridge data record and a two-step translation process using SPOT-VGT. The first step is to translate between AVHRR and SPOT-VGT then from SPOT-VGT to MODIS, with SPOT-VGT providing observations to close the gap between the other two sensors. These methods are cumbersome and prone to ancillary issues, especially the lack of an effective cloud mask for SPOT-VGT and due to differences in data processing approaches, which creates a host of complex challenges.

For the current methodology (V4) we limited the AVHRR and MODIS data sets to the 5 years before and after transition (1995-2004) from each sensor assuming minimal disturbances. The resulting time series is then filtered to retain only cloud free and high quality data as defined by the per pixel Quality Assurance information available in each data record. The remaining data is then aggregated into a one-year cycle for each sensor and a regression analysis performed to extract three forms of the "Slope and intercept" equation for each pixel (Fig. 7):

- A general Slope and Intercept (α,β),
- A Slope only with Intercept set to ZERO ($\alpha,\beta=0$)
- A Translation only and Slope is set to 1 ($\alpha=1,\beta$)

Although, the intercept is expected to be zero when the input is equal to zero, a pre-analysis of the AVHRR vs. MODIS data indicates that an intercept was required for certain snow covered areas

(high latitude mostly) due to the erratic AVHRR data behavior over snow. On the other hand, tropical forests time series data and due to the excessive presence of residual clouds (noise) (Donahue, et al., 2005) and lack of a consistent growing season (usually a flat profile) required special handling where a slope only maybe required to effectively translate between AVHRR and MODIS. Other areas require the application of a slope and intercept. To further consider the impact of seasonality on the relationship between MODIS and AVHRR we modified this algorithm to work on a per-pixel monthly adjusted step. The resulting continuity algorithm follows this general form:

$$MODIS\ (AVHRR_{VI}) = \ \alpha * AVHRR_{VI} + \ \beta\ (7)$$

Where:

The left side represents the desired continuity (new) VI value for the AVHRR sensor. $AVHRR_{VI}$ is the VI value measured by AVHRR. α is the slope of the correlation, and β is an intercept (dynamically set for each pixels). To account for seasonal changes, the method was refined by statistically analyzing the data at a monthly interval. The resulting monthly translation functions are then applied to the corresponding AVHRR data. To minimize deviation and preserve the trends in the resulting data the algorithm is applied dynamically to each pixel following these conditions and logic:

If Greenland then
  Use Slope with Intercept algorithm
Else
  If Slope-Z > 0.3 then
    Use Slope Only Algorithm
  Else
    If (Slope-Z<-0.9)
      Use translation Algorithm
    Else
      Use Slope with Intercept Algorithm
Where Slope-Z is the slope resulting from setting the intercept to ZERO.
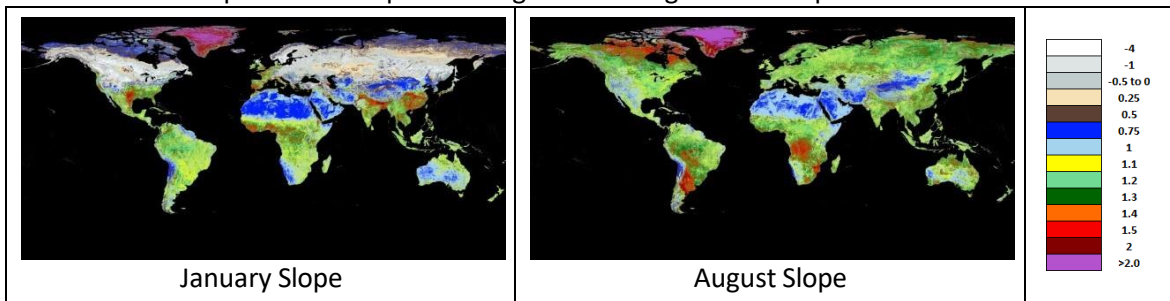
A: Monthly continuity algorithm based on a Slope and setting Intercept=ZERO
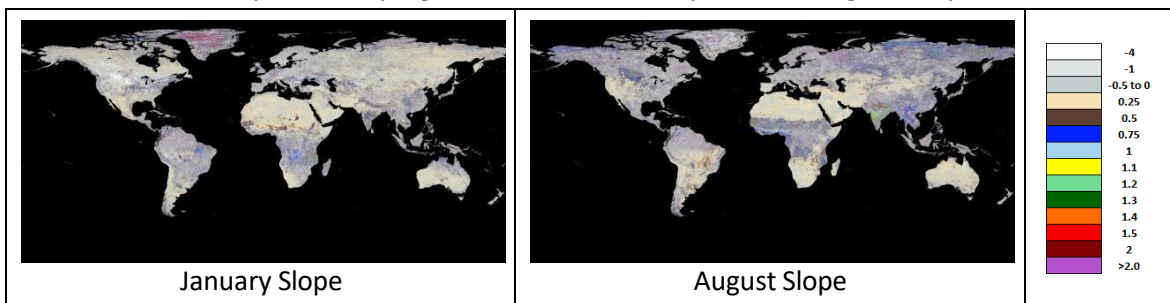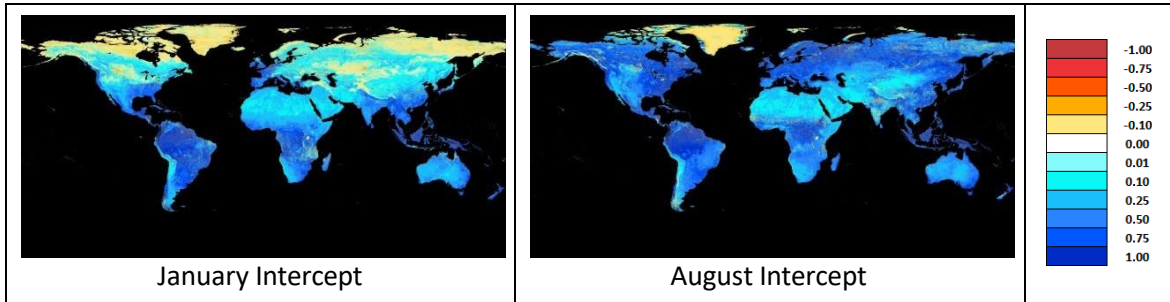
B: Monthly continuity algorithm based on a Slope and Intercept
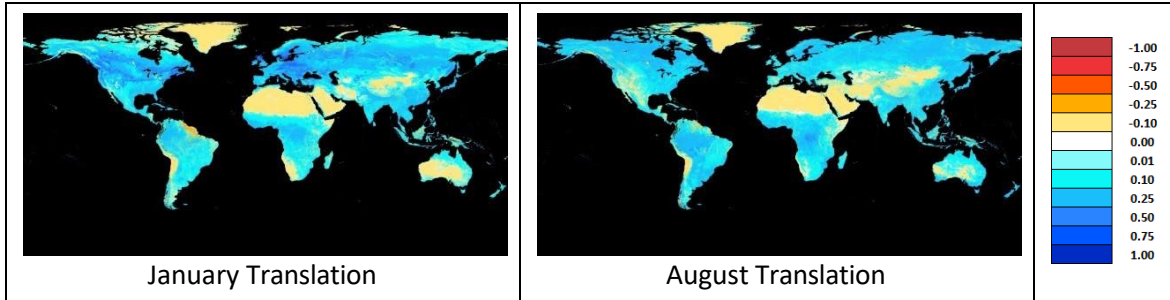
C: Monthly continuity algorithm based on a Simple Translation

Figure 7. Global continuity algorithm following the proposed methods (full regression, regression with slope only, and translation only). The algorithm dynamically decides what form of the continuity to apply based on the analysis of the location of the pixel and its VI profile
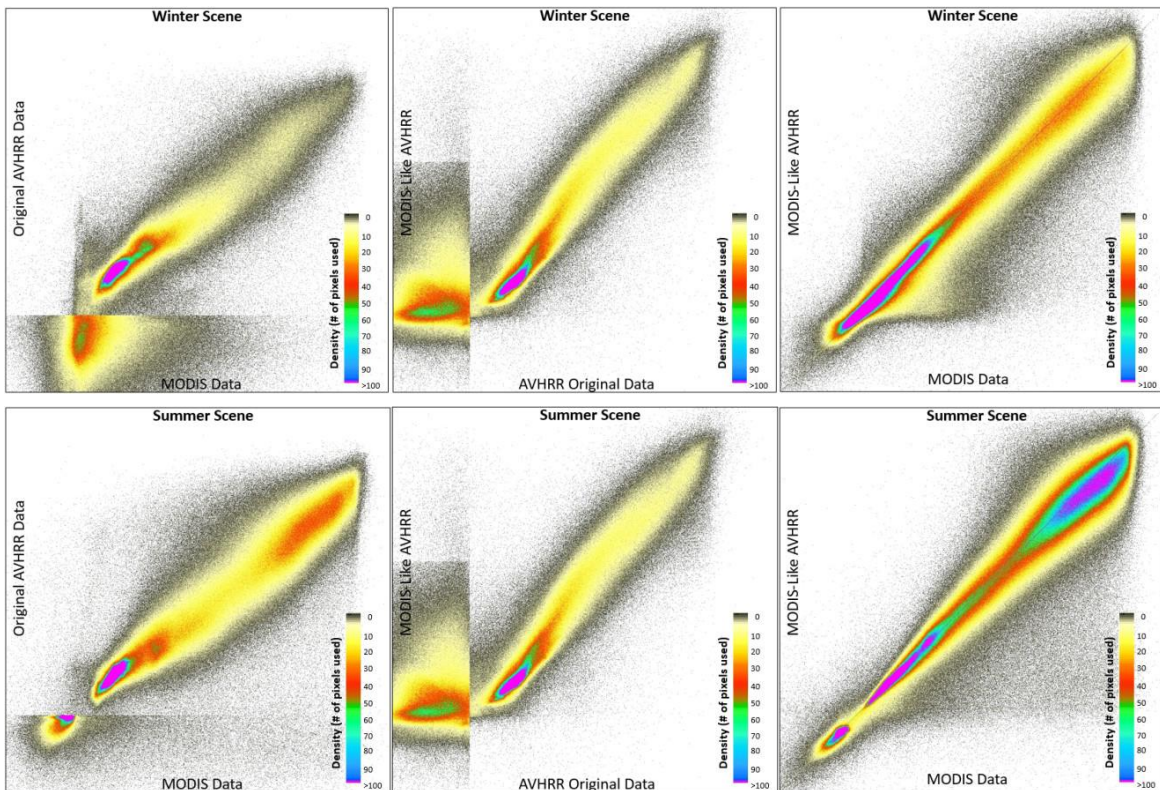
Figure 8. AVHRR-MODIS Correlations (winter and summer data) demonstrating the process of going from AVHRR to MODIS like data

Only cloud free (Stowe et al., 1999) and atmosphere free input AVHRR and MODIS data are used with this algorithm. The remaining gaps (the strict filter leaves large areas without data) are then filled using a simple temporally based gap filling algorithm that gives more weight to neighboring pixels. The algorithm is constrained to using no more than a month search window to gap fill the missing data, otherwise the long term average is used instead to fill the gap (more on this in later sections). Figure 8 shows the various correlations steps and confirms the algorithm performance using winter and summer data.

2.2.   Satellite Phenology ESDR Algorithm

Studying phenology with remote sensing involves the estimation of the vegetation dynamic, start of growth, and senescence/dormancy at seasonal and interannual time scales. The consistency of the resulting phenology product is dependent on the accuracy of the vegetation index input and the phenology extraction methodology. A land surface phenology Earth Science Data Record (ESDR) should accurately characterize the land surface temporal behavior through time in a manner consistent with what is observed on the ground and specie level.

In general, vegetation has a strong and well defined growing season. The annual vegetation cycle has a bell shape (Fig. 9) that is defined by the following set of parameters:

- Start of season (a), End of the season (b), Length of the season (d),
- Day of peak (c, time),
- Magnitude of peak (e, VI),
- Rate of greening (ΔVI/Δt between (a) and (c)),
- Rate of senescence (ΔVI/Δt, between (c) and (b)),
- Cumulative green (area under the curve)

A phenology algorithm is designed to estimate these parameters by analyzing the time series of an annual VI profile in a consistent fashion.
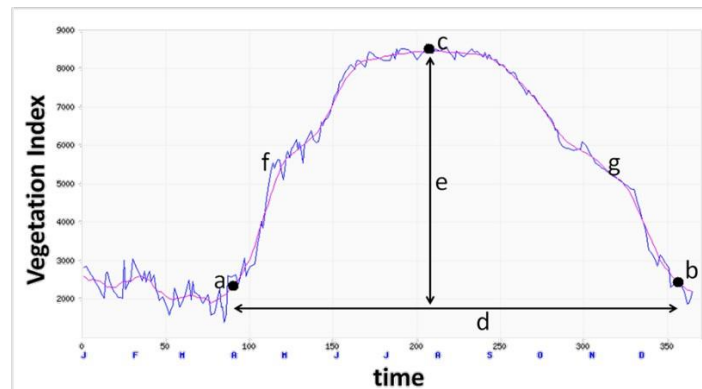
Figure 9. VI time series profile and different phenology parameters.

A variety of methods have been proposed and used to identify different phenology parameters (dates of onset and offset of growing season) using satellite data. Table 1, summarizes some of the most widely proposed and used algorithms. For the proposed phenology ESDR we opted for the use of an algorithm based on a variation of the Half-Maximum VI approach proposed by White (1997). The Half-Max algorithm provides a consistent and simple methodology (White et al. 1997, 2009) for identifying transition dates that define the key phenological phases of vegetation at annual time scales (Zhang et al., 2003). This algorithm is driven by either NDVI or EVI/EVI2 and requires VI specific parameterization. The bulk of the algorithm itself has been successfully

validated and used to support a number of studies related to climate-vegetation relationships (White et al. 2009, White and Nemani, 2006, Huete et al., 2008).

Table 1. Remote sensing VI based phenology parameters extraction methods.

Approach	Reference	Remote Sensing Data Input
Inflection points on fitted, bell-shaped curve	Badhwar (1984)	NDVI driven
Threshold techniques	Fischer (1994)	NDVI driven
Divergence of smoothed curve from autoregressive moving average	Reed et al. (1994)	NDVI driven
Largest NDVI increase after air temperature exceeds 5°C	Kaduk and Heimann (1996)	NDVI driven
VI Half-Max approach	White et al (1997)	NDVI driven
TIMESAT Curve fitting approach	Jönsson &Eklundh (2002, 2004)	VI driven
Piecewise logistic functions	Zhang et al. (2003)	EVI driven
Pheno-cluster approach	White et al. (2005, 2009) Didan & Huete (2004, 2005) Didan et al. (2010)	NDVI/EVI driven
Modified Half-Max algorithm	White et al. (2009) Barreto et al. (2015) Didan et al. (2010)	NDVI or EVI

To estimate the phenology transition dates, the daily NDVI or EVI2 continuity time series is used. Figure 33, illustrates the parameterization of phenology algorithm and growing season cycle. The model uses a set of parameters to guide the time series profile analysis for the pixel (Fig. 9). To account for the relatively noisy quality of the data, in particular form the AVHRR era, we used a data smoothing technique based on a moving window approach (defined during the model parameterization, Barreto, et al. 2015). This algorithm is then applied to both the NDVI and EVI2 time series separately leading to two different phenology records.

While, phenology is generally thought of as a single growing season cycle, some natural and managed lands (irrigated agricultural lands in particular) are characterized by more than one growing cycle. To account for this the algorithm is adjusted to consider up to three separate growing seasons (Fig. 10)

Single Season with one unique well defined profiles

Dual seasons with two well defined and separate profiles

Triple seasons, with up to three unique and separate profiles
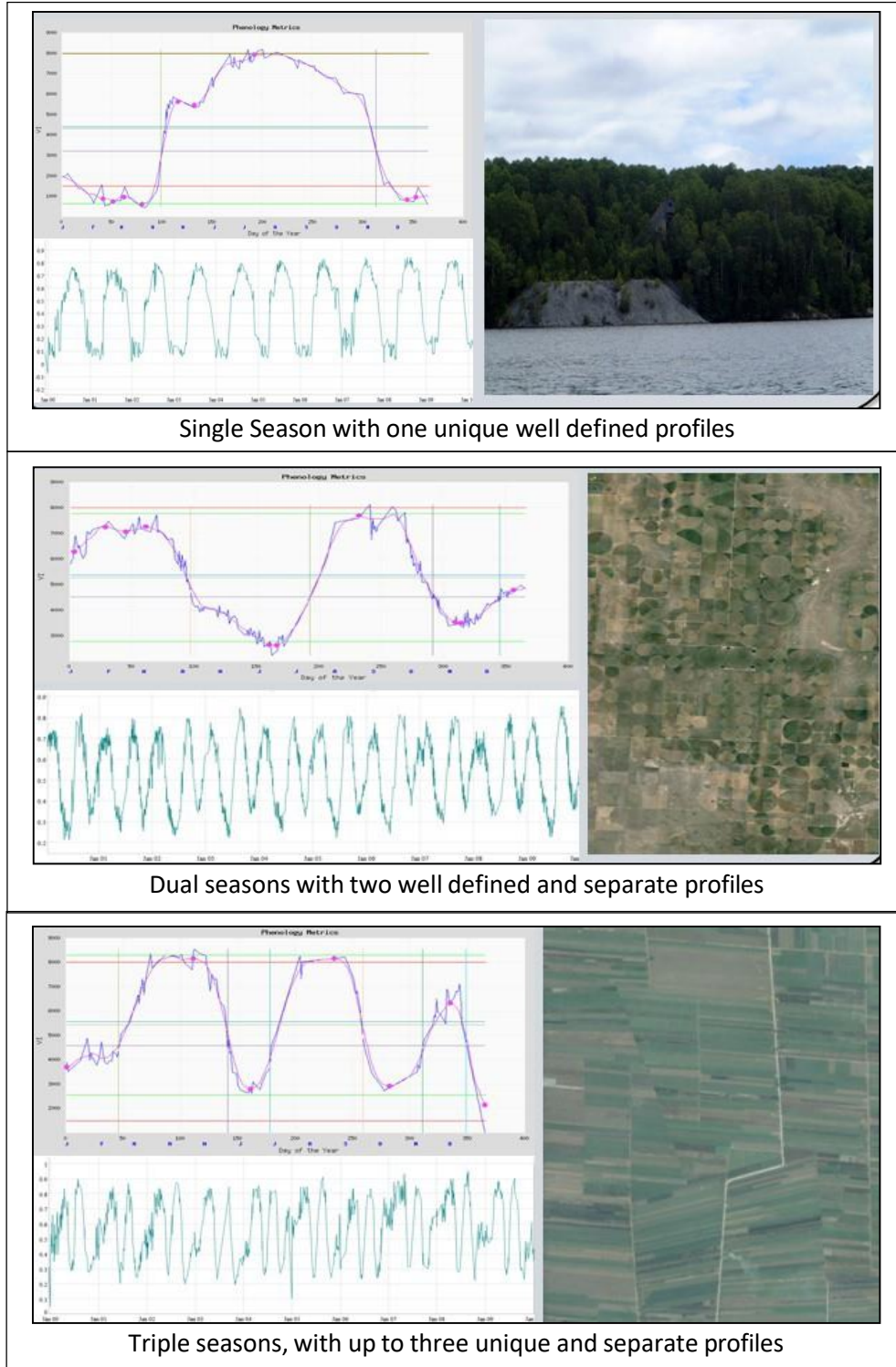
Figure 10. Typical growing season profiles with single, double, or triple cycles

2.3. What is New in Release 4

Since Version 1.0 (deprecated) the data record has gone through three additional reprocessing exercises. This modular and incremental reprocessing approach aims at addressing issues and

implementing improvements. Since release 3.0 we've made changes and/or additions to the data and methods along these major categories:

- This V4 release is based on input daily data from AVHRR LTDR v4 (1981-1999) and MODIS C5 (2000-2014) data records
- We started using the new MODIS based MOD44W Land water mask product, with many improvements in the wetlands, small lakes, rivers, and riparian classes (Fig. 11).
- Science changes
- Product files structural changes
- Improved Processing rules
- Additional VI byproducts
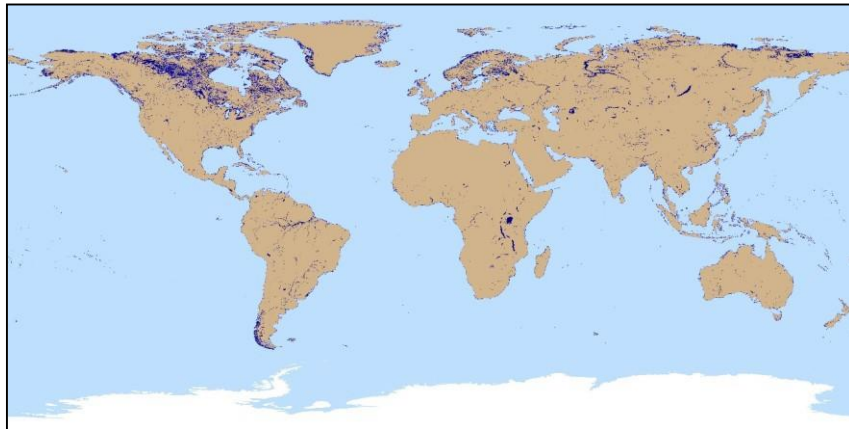
Figure 11. MOD44W based new global Land Water mask

A key improvement to the data record is the adoption of a two-step filtering approach of the input data and the use of a new per pixel continuity algorithm at a monthly step. These changes have positively impacted the VI time series and improved its quality through a better and more consistent continuity between AVHRR and MODIS, and improved spatial consistency over traditionally problem areas (high latitude boreal and tropical forests).

The phenology ESDR is now based on a 3-year moving window average which enables the removal of persistent and residual noise in the time series leading to an improved phenology parameters extraction. Starting V4 we have also included a special "Vegetation background signal" which defines the average minimum background landscape VI value above which vegetation shows seasonal dynamic. In addition to these standard daily VI and phenology ESDR products we have generated a suite of multi-year long term average records. All ESDRs now include the simple per-pixel reliability index that summarizes and ranks the data quality in an ordinal fashion useful for automated data post processing.

3.   VIP ESDRs Description

3.1.   File Format

The sensor independent Vegetation Index and Phenology ESDRs are stored in the Hierarchical Data Format-Earth Observing System (HDF-EOS) structure, which is the standard archive format for all NASA EOS Data Information System (EOSDIS) products. Each VIP file contains two separate structures:

- Scientific data sets (SDS) which are the actual data stored in a 2-D array

- Three sets of metadata
  - Structural metadata that describes the scientific content of the file,
  - Core metadata that describes the projection and grid name,
  - Archive metadata that describes miscellaneous aspects of the data and file such as date, time, and statistics about quality. These are useful for archiving and searching the product.

All VIP products are in a geographic grid structure, defined as Latitudinal-Longitudinal projected, fixed-area size. The use of metadata enhances the self-describing nature of HDF files and is useful to the end user, facilitating the archiving and searching of files.

Metadata provides the users with general information about the file contents, its characteristics and general quality, which aids in deciding if a particular day/file is useful. There are two types of metadata attributes:
- Global attributes common to all VIP products, and
- Product specific attributes

3.2. File Naming Convention

There are two types of VIP ESDR products, Vegetation indices and Phenology. The file naming convention for the VIP products suite is similar to the file naming used for MODIS and other EOS products. The file names are structured left to right following this nomenclature:

3.2.1. Vegetation Indices

Example File Name: **VIP01.A2008227.004.2015195220005.hdf**

In addition, the VIP Lab distributes other Vegetation Index products with varying degrees of processing. These VIP distributed files follow this naming style:

Example File Name: **VIP01P1.A2008227.004.2015195220005.hdf**

Table 2. Vegetation indices file naming convention

Sequence	Meaning
VIP	Identifies the product as a VIP product
01	Indicates the compositing interval (2 digits)
	01= daily product
	07 = 7 days
	15 = 15days
	30 = monthly product
P1*	Identifies the product type
	No P? : Continuity gap filled data (available from VIP and LP-DAAC)
	P1 = Preprocessed input data (only from the VIP website)
	P2 = Preprocessed filtered data (only from the VIP website)
	P3 = Continuity data (only from the VIP website)
A2008227	Is the year (4-digits) of the observation followed by the day of year (3-digits)
004	Identifies the data product version
2015195	Is the year the file was processed followed by the day of year
220005	Is the hour, minute and second the file was processed
.hdf	Indicates that output file is in HDF format

3.2.2. Phenology

Example File Name:

VIPPHEN_NDVI.A2008.004.2015190220005.hdf

VIPPHEN_EVI2.A2008.004.2015190220005.hdf

In addition, the VIP Lab distributes other Phenology products based on multiday composited Vegetation Index. These VIP distributed files follow this naming style:

Example File Name:

VIPPHEN15_NDVI.A2008.004.2015190220005.hdf

VIPPHEN30_EVI2.A2008.004.2015190220005.hdf

Table 3. Phenology file naming convention

Sequence	Meaning
VIP	Identifies the product as a VIP product
PHEN	Identifies the dataset as Phenology
XX	Indicates product is based on composited data with interval (2 digits)
	No XX: Daily (available from VIP and LP-DAAC)
	07 = 7 days (only from the VIP website)
	15 = 15days (only from the VIP website)
	30 = monthly product (only from the VIP website)
_NDVI _EVI2	Indicates the vegetation index used to produce the phenology metrics
A2008	Is the year (4-digits) of the observations
004	Identifies the data product version
2015195	Is the year the file was processed followed by the day of year
220005	Is the hour, minute and second the file was processed
.hdf	Indicates that output file is in HDF format

3.3. VIP Product Sequence

There are 5 products in the VIP series that are organized by temporal frequency as follows:
- VI Daily products
  - VIP01: daily 0.05-deg VI
- VI Composited products
  - VIP07: 7-day 0.05-deg VI
  - VIP15: 15-day 0.05-deg VI
  - VIP30: monthly 0.05-deg VI
- Phenology products
  - VIPPHEN: daily/yearly 0.05-deg Phenology

All VIP products are based on surface reflectance data from AVHRR (N07, N09, N11 and N14), and MODIS MOD09, which are daily level 3G (L3G – G stands for gridded and projected in this case) products. The VI algorithms ingest the level 3G surface reflectance and generates the VI products. To keep products consistent, all VIP composited datasets contains the same number of SDSs and metadata structures. They follow the Daily product suite with the exception of the addition of the Day of the Year layer that stores information about the retained day of composite. The Phenology

product is reported yearly and uses the VIP VI Daily data as input. Figure 12 shows the data production flow diagram and the resulting product records.
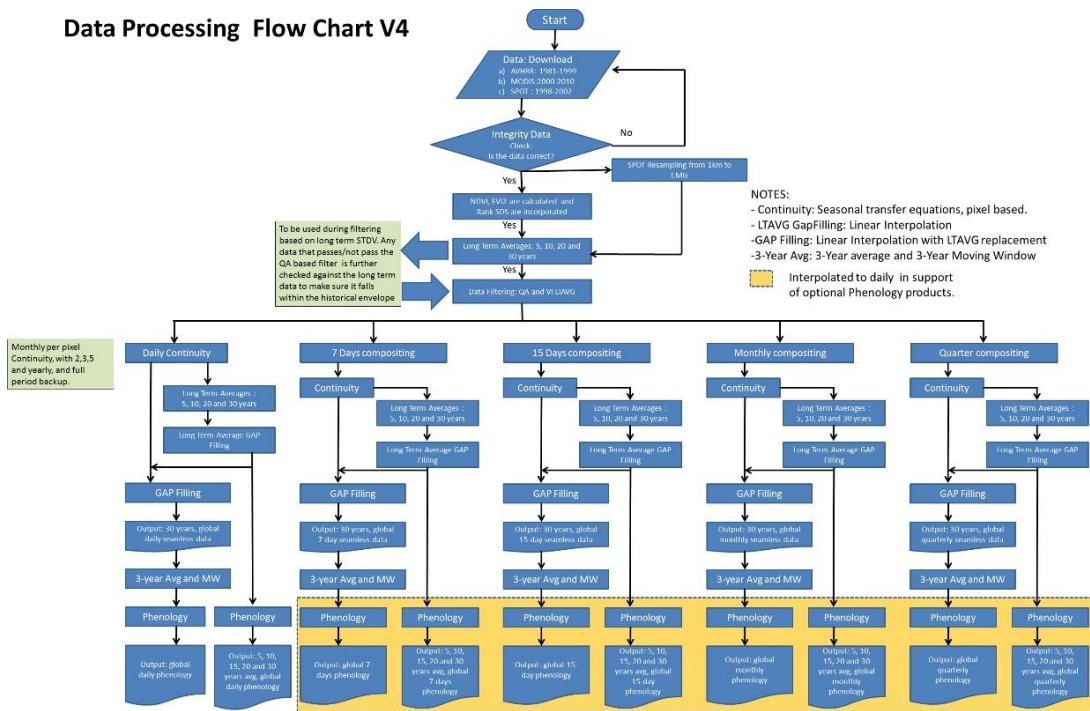
Figure 12. VIP products processing flow diagram

4.   VIP01 (daily 0.05-deg) Vegetation Index ESDR

This product is generated using the daily AVHRR and MODIS Level-3G (L3G) surface reflectance data. Through a series of processing algorithms (Fig 12, Fig. 13) a sensor independent gap filled dataset is obtained. This process starts by filtering the input data to retain high quality cloud free data, which results in large spatial gaps. These gaps are addressed last in this processing chain to avoid biasing the final product.

| 1. Input surface reflectance, VIs calculation | 2. Ranking according to QA internal flags |

| 3. 2nd step ranking and filtering | 4. Filtered VI |
| 5. Continuity & Gap filling | 6. Final rank status |
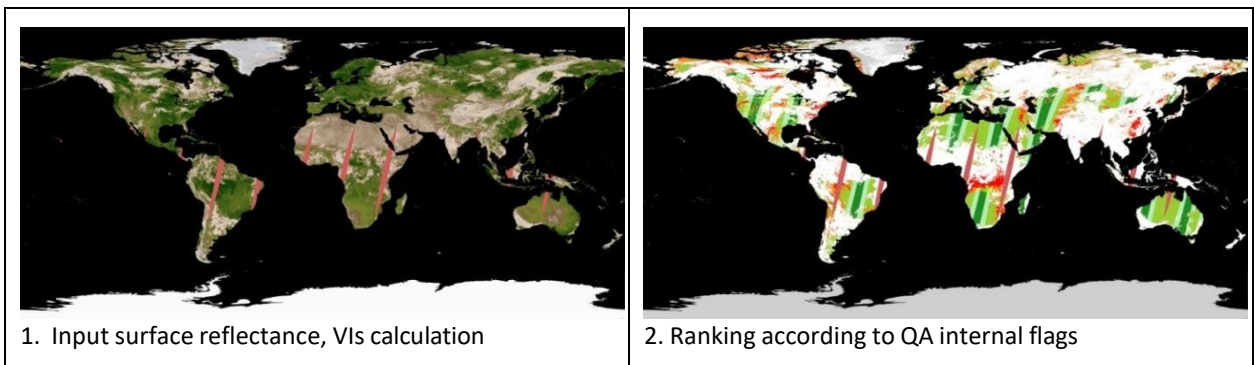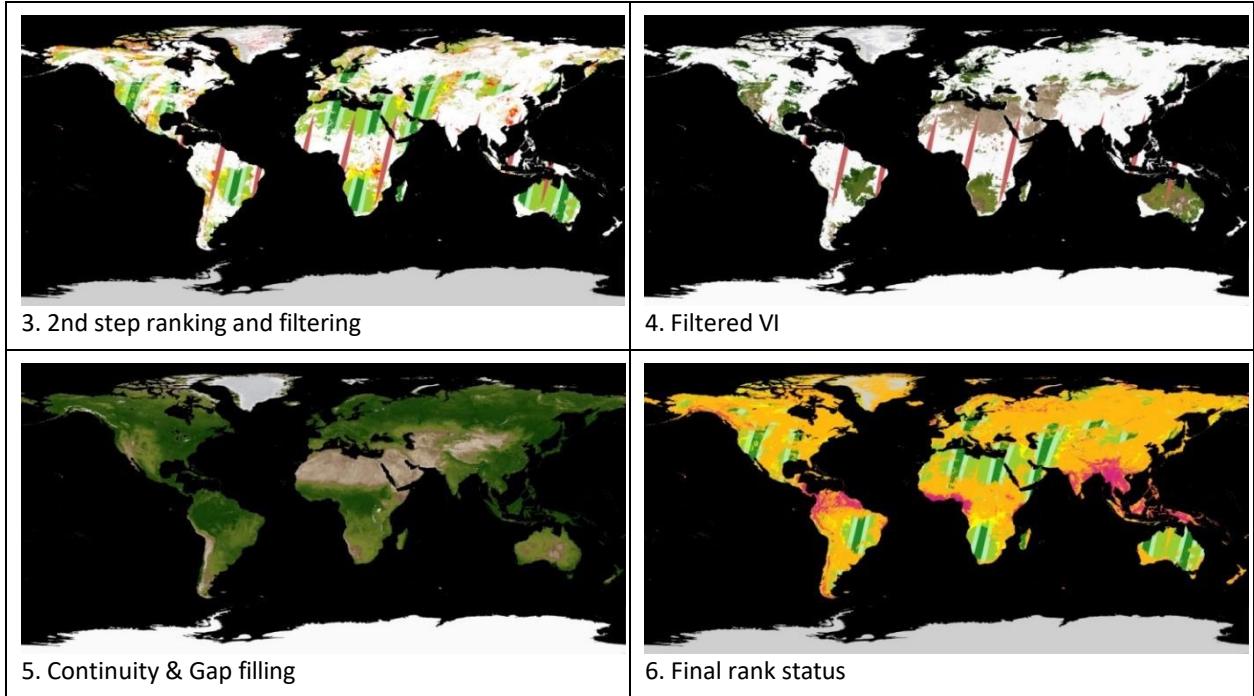
Figure 13. Processing steps used by the VIP processing chain

4.1. Algorithm Description

The VIP processing algorithms chain (Fig. 14) operates on a per-pixel basis and consists of several stages of preprocessing and data quality evaluation. It starts by the ingestion of the AVHRR and MODIS daily data, the creation of temporary pre-continuity vegetation index, then data quality analysis and ranking, data filtering, continuity and translation and finally spatial gap filling.
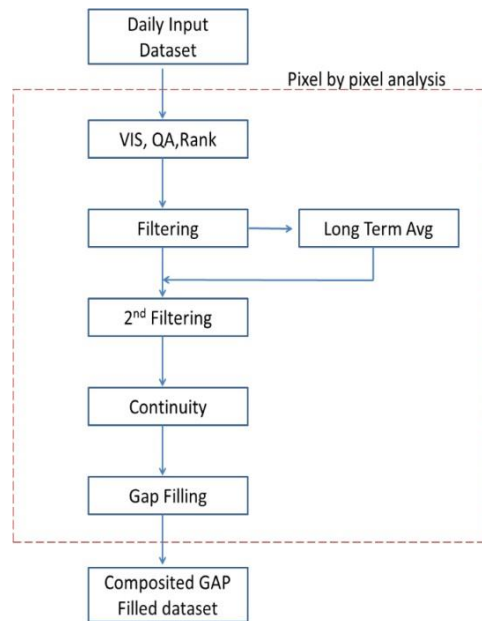
Figure 14. VIP01 data processing flow diagram

4.1.1. Input Data

Daily surface reflectances, viewing geometry, and quality assurance data from AVHRR LTDR-V4 (1981-1999) and MODIS C5 (2000-2014) are the input to this algorithm chain. Three pre-processing algorithms are applied to each land pixel.

- Vegetation indices calculation
- Quality assurance standardization, and
- Data Quality Ranking.

NDVI and EVI2 vegetation indices are calculated using the Red and NIR surface reflectances following the science algorithms described earlier. Only reflectance values between 0-1.0 (scaled by 10,000) are retained. Resulting VI values are also scaled by 10,000 and range from -10,000 to 10,000 (this is a slight deviation from MODIS VI Data range being -2000 to 10000). Pixels for which no data could be obtained (i.e. red or NIR outside the normal range, gaps between orbits, bad VI values, etc...) are assigned a fixed fill value,

FILL_VALUE= -12,000 over land pixels at high latitude during the dark winter months,
FILL_VALUE= -13,000 over land to help identify missing data,
FILL_VALUE= -14,000 over Antarctica, and
FILL_VALUE= -15,000 over water/oceans

The adoption of different fill values helps separate between the reasons a pixel value may be missing.
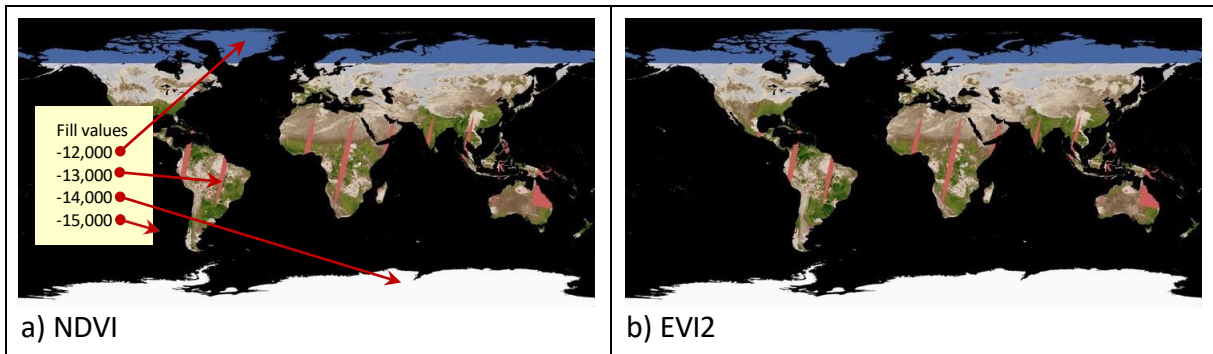
Figure 15. Typical NDVI and EVI2 preprocessed daily input image (DOY 001, 2008). The arrows indicate the different "Fill Values", their values, and their potential locations.

Quality flags from AVHRR and MODIS vary greatly, even within the same MODIS sensor era earlier quality flags from 2000-2002 are different from the period 2003 onward, therefore we synchronized the QA bit flags into a single standard structure. The VIP products adopted a QA structure similar to MODIS since it is the newer and reference sensor for all continuity work. AVHRR data does not inform about Snow/Ice or Aerosol loads, so in order to compensate for this missing information, a statistically based Snow/Ice and Aerosol quality data was created from MODIS long term QA data and then calibrated by AVHRR NDVI values and time series. This approach uses a probabilistic model (frequency analysis of MODIS QA) to infer the likelihood of cloud, snow, ice and aerosol load presence. The probability is further constrained by the NDVI value, since typically NDVI value would be low when there are snow/ice, heavy aerosols, or clouds. When NDVI is close to the long term average NDVI value the snow/ice or aerosol value is set to clear. While this method is purely statistical and does not account for the actual overpass conditions for the day it does help in identifying poor quality data that are due to snow/ice and heavy aerosols (Fig. 16, Barreto et al.

2010).
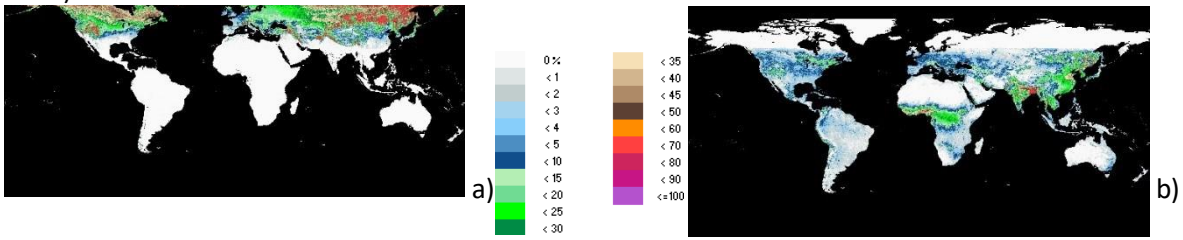
Figure 16. Probability analysis of snow/ice and High Aerosol occurrence based on MODIS QA data for the month of January. (a) Snow probability during the month of January, (b) High aerosol load during the month of January. This data is used to guide and restrain the AVHRR data QA.

Finally, pixels are ranked (Fig. 17) based on data usefulness, the VIP QA layer, and the viewing geometry. This pixel reliability or rank summarizes the pixel quality based on its QA and how likely it will serve the end user. The categories of this ranking scheme are in table 4.

Table 4. VIP rank classes

Rank	Label	Rank	Label
0	Excellent	8	Snow/Ice
1	Good	9	Cloud
2	Acceptable	10	Estimated
3	Marginal	11	LTAVG
4	Pass	-1	NO_DATA
5	Questionable	-2	NO_DATA High latitude
6	Poor	-3	Antarctica
7	Cloud Shadow	-4	Water/Ocean
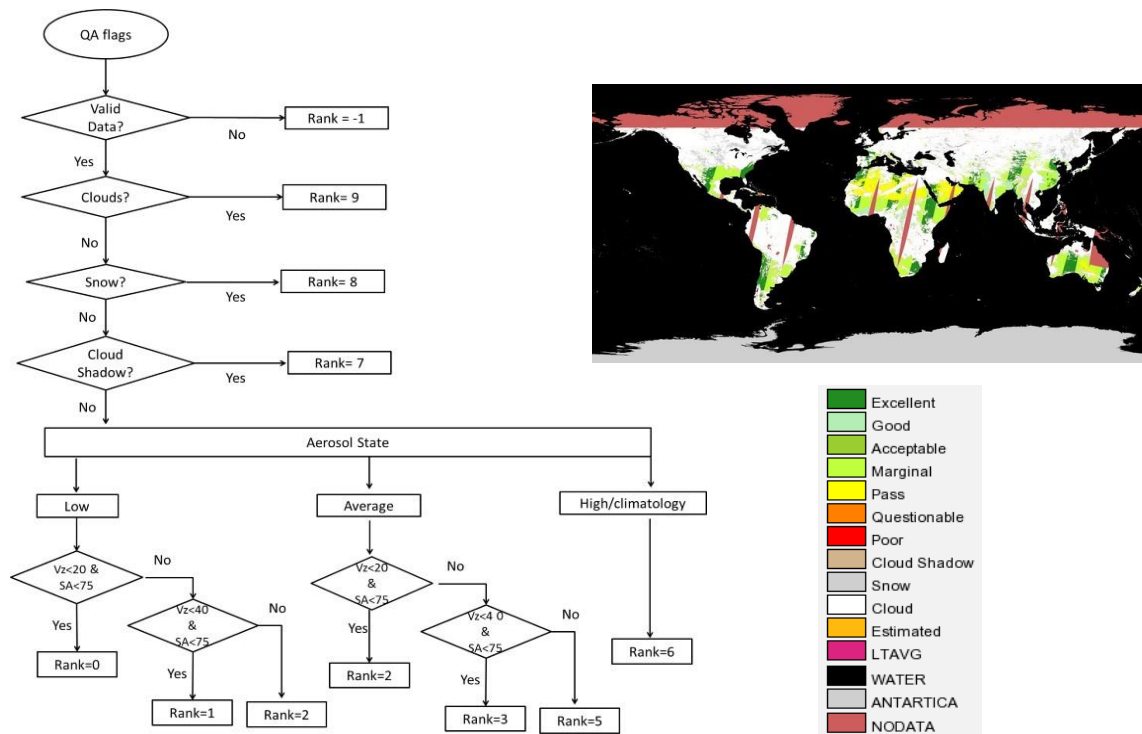
Figure 17. Data raking algorithm flow diagram and example Rank image (DOY 001, 2008, inset)

23

4.1.2.   Filtering

In order to minimize the impact of poor quality input data on the continuity algorithm, data must be filtered and only cloud free and high quality data retained. This insures the data used for the derivation of the continuity equations are free from bias and noise. However, time series analysis of filtered data from previous VIP data product versions indicated that the QA information for filtering MODIS, and especially AVHRR, is quite inconsistent leaving large amounts of poor quality data. Many pixels are either miss classified clear when they are cloudy and vice versa. Similarly, pixels with high aerosol loads were classified as having low or average aerosols. To minimize the impact of these QA omissions and commissions a second filtering model was implemented and a re-ranking of the data is performed. The second filtering is based on the analysis of the retained high quality data that resulted from the first step QA filtering (using the VIP Rank SDS). From this filtered data we generate long term statistics for NDVI and EVI2. Confidence intervals are then created from this long term average NDVI and EVI2 time series and each VI pixel value is then compared to this envelope. An ad-hoc 1 or 2 (σ) standard deviation is used to construct these confidence intervals. A good quality pixel needs to reside in the **±2σ** envelop otherwise it will be rejected, and any average to questionable quality pixel will need to reside within the tighter interval of **±1σ** envelope otherwise it will be rejected. The logic behind this is that a good quality pixel is permitted to deviate more since that could result from natural change or disturbance, as opposed to a low quality pixel where deviation could be due to noise and hence limited to only **1σ**. All VI values are re-evaluated after this second filter algorithm and assigned a new rank to reflect the new quality assessment. Special considerations were given to trends in order to safeguard legitimate changes and disturbances by looking at the following observations during this analysis for consistency in the trend, if data drops below the long term average consistently it is likely due to a disturbance rather than noise and the pixel is retained.
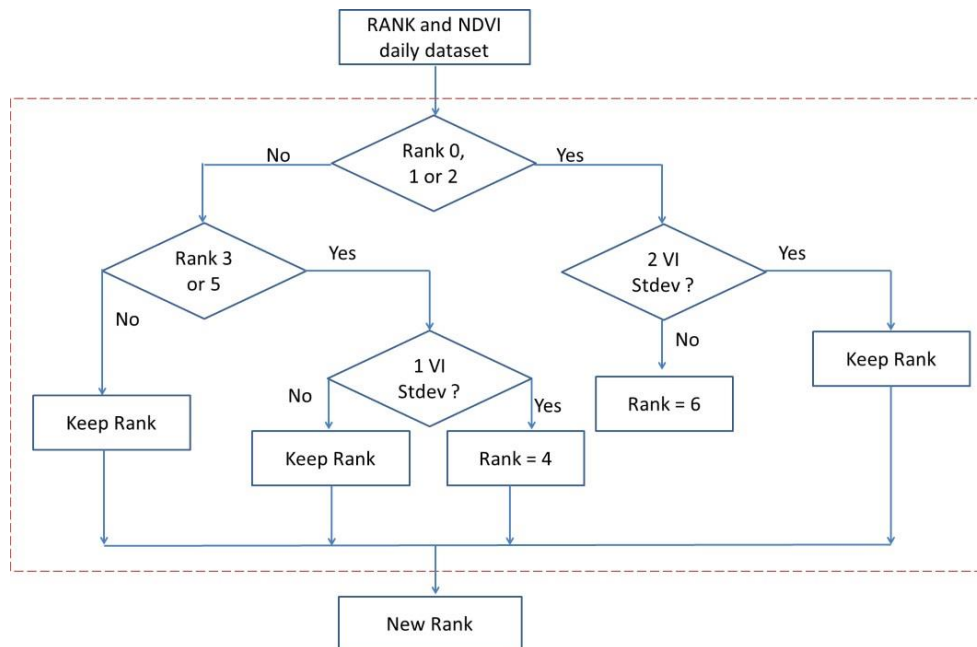
Figure 18. VIP Re-ranking flow chart

4.1.3.   Continuity

The operational continuity algorithm is based on a set of transfer equations that translate

AVHRR observations to MODIS like values. The continuity equations are based on the MODIS time series and a linear regression that correlates AVHRRR to MODIS. While most of the literature calls for a single global equation or land cover based translation equations our initial analysis (VIP V1 and V2 based on those methods) shows that these methods are prone to bias, over-translation, do not capture the regional conditions which leads to unusable value for many areas. When we tested the SPOT-VGT (1998-2002) bridge-based method the results were still quite poor due to the lack of a consistent SPOT-VGT cloud and the inadequate QA information. A continuity method based on a direct AVHRR MODIS comparison during the 5-year transition period (1995-2004) with special attention to QA and land cover disturbances was tested and adopted. The final operational algorithm was based on a statistical analysis of each pixel across the two sensors and further refined using a monthly step. The method was applied separately for NDVI and EVI2. And while this is not fully accurate, we assumed that the bulk of the Earth's vegetation land cover observed by AVHRR 1995-1999 did not change considerably when observed by MODIS during 2000-2004 provided the impact of disturbances is minimized. Only high quality data was considered and for cases where no valid observations were available the periods were extended to AVHRR 1990-1999 and MODIS 2000-2010 (only few pixels required this extension).

Monthly data sequences were dynamically expanded (before and after the month) by 15, 30 and 45 days to ensure an adequate number of good quality observations were available for each month. Expanding the month also helps in smoothing the transition between adjacent months. The resulting continuity/translation equations are based on the minimum period length required for adequate data availability. That is, in regions where good quality observations are obtained consistently the period length equals the given month (ex: US Southwest) over areas with persistent clouds and other issues resulting in large temporal gaps (ex: Amazon Basin) the data period could be up to 3 months or more. Following this approach, the continuity algorithm for each pixel takes one of three forms:

- Zero Intercept: MODIS($AVHRR_{VI}$) = Slope * $AVHRR_{VI}$
  The majority of pixels follows this equation, with a slope value of around 1.2 for vegetated areas and 0.8 for semiarid areas.
- Slope and Intercept: MODIS($AVHRR_{VI}$) = Slope * $AVHRR_{VI}$ + Intercept
  The equation takes on this full form when the statistical analysis results in a Slope > 2 or < 0, indicating the presence of snow/ice which renders the native observation problematic making it impossible to match the time series between AVHRR and MODIS with a smooth regression without a correction by an intercept.
- Translation: MODIS($AVHRR_{VI}$) = $AVHRR_{VI}$ + Correction Translation Coefficient
  Additional challenges remained when pairing AVHHR to MODIS over ice/snow covered regions. With a positive VI value for Snow/Ice in the MODIS record and negative in the AVHRR record. The statistical analysis over the input data reports a slope values from the 'Slope and Intercept' method close to zero and large intercept values. When applying the full Slope/Intercept to this data (slope being close to Zero and Intercept high) it did bring AVHRR values close to MODIS however the trend and variance of the data was removed due to the close to ZERO slope. To try and preserve the time series trend only a translation model was proposed, where the data from AVHRR was simply bumped up to MODIS like dynamic range using the translation coefficient, resulting in a better data behavior while preserving the data trends.

| a) Slope of the Slope/Intercept method | b) Intercept of Slope/Intercept method |
| c) Intercept and Slope= Zero | d) Translation |
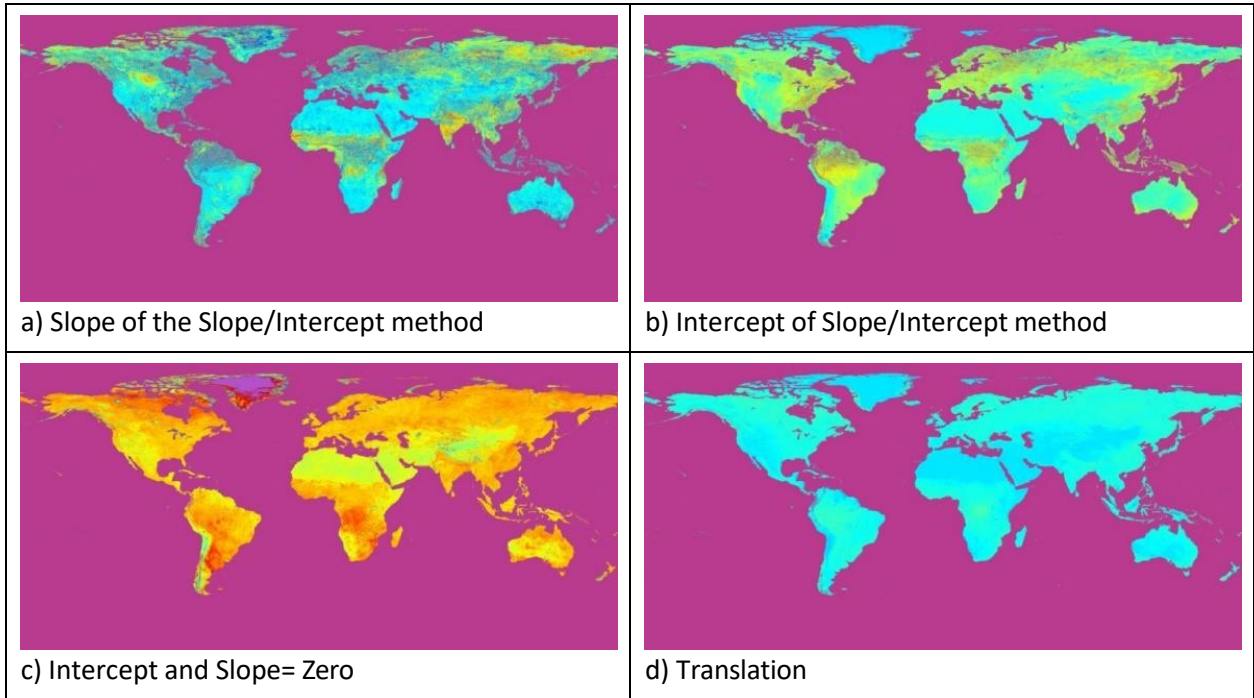
Figure 19. Global distribution of continuity equation parameters for the month of August.

4.1.4.　Continuity Algorithm Performance

To capture the efficiency of this continuity approach we show some results illustrating the performance of the different continuity algorithms for NDVI and for key land cover types. These time series are based on a single CMG pixel. The first 20 years (1981-1999) are from AVHRR the remaining years are from MODIS. The graphs show the single sensor input data (AVHRR, MODIS), the filtered input, and the translated and gap filled result.

Deserts and Arid areas NDVI time series (Sahara Desert)

Semi arid NDVI time series (New Mexico)
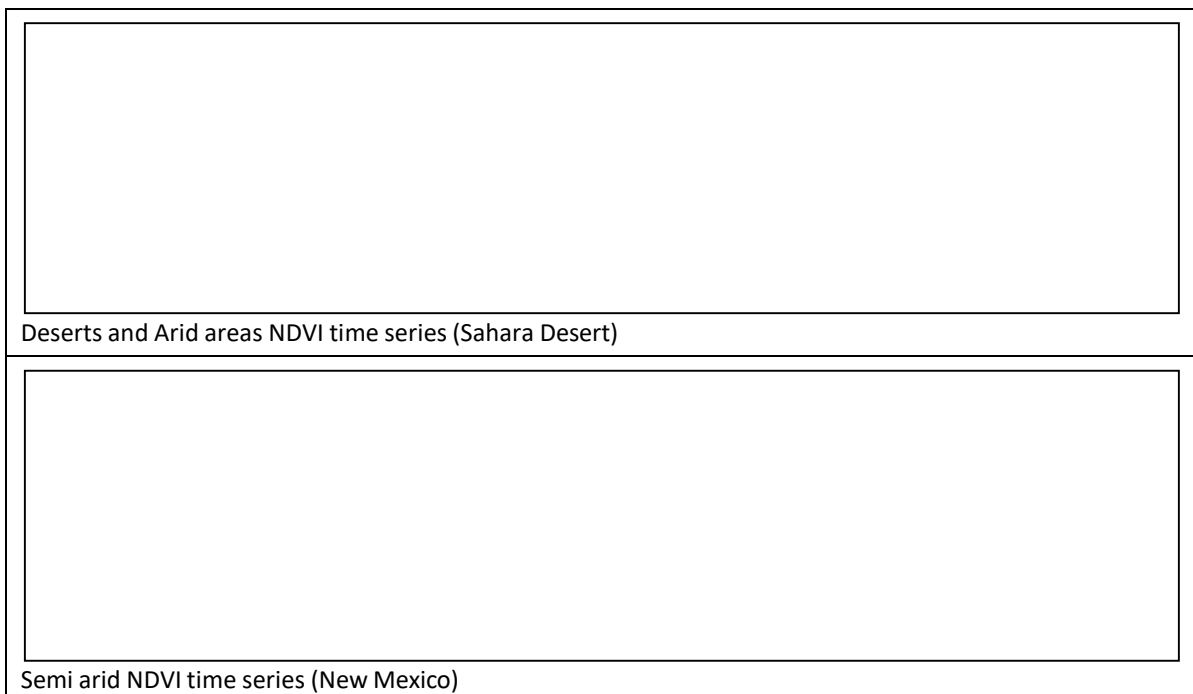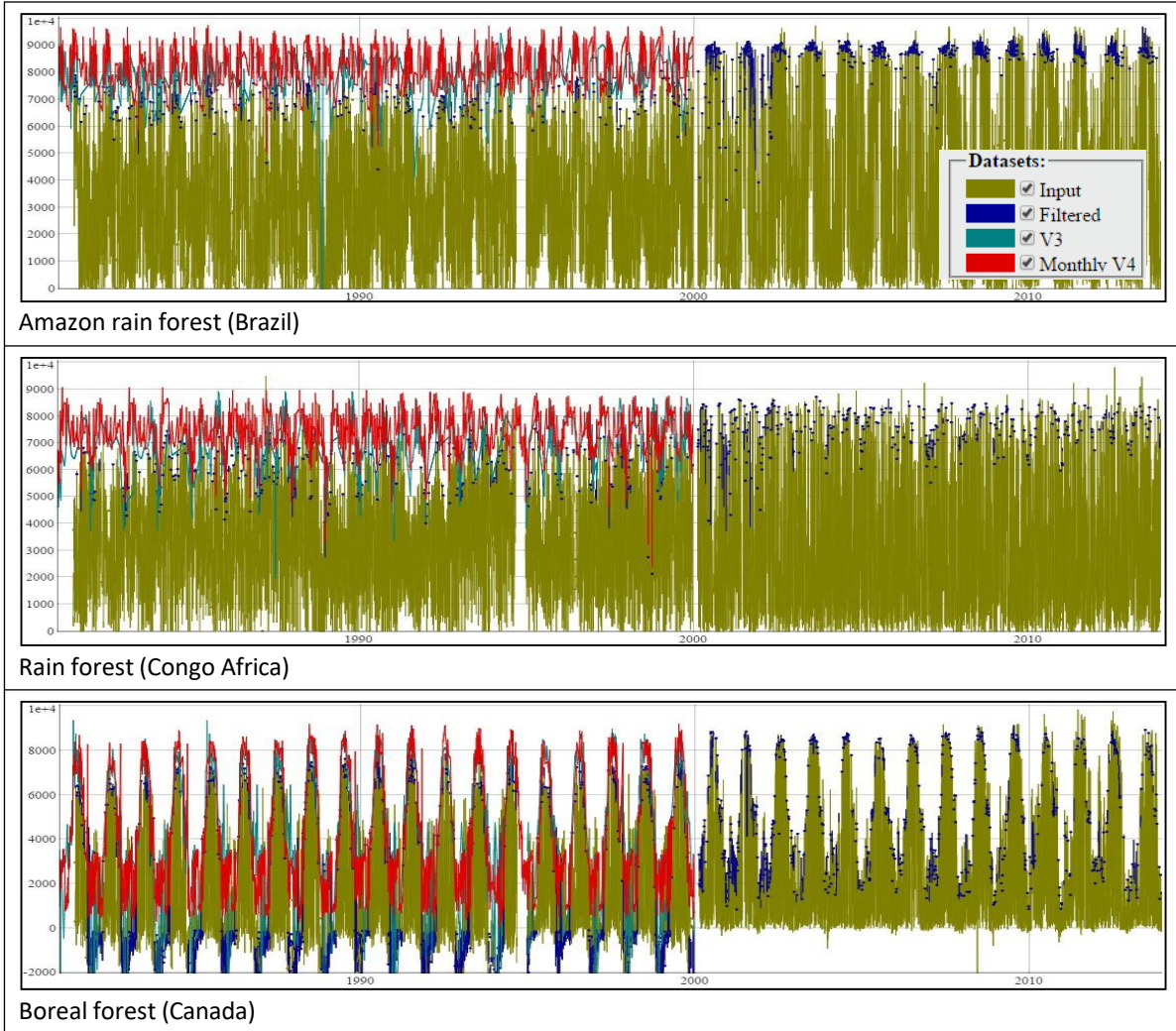
Figure 20. Continuity algorithm performance over some land cover types. We also show V3 and V4 continuity algorithms performance differences.

4.1.5.   Gap Filling

Once the data is translated into MODIS like values the long term sensor independent ESDR VI record is obtained. However, due to the rigorous filtering methods used the resulting data records will contain large spatial gaps. The final ESDR need to be spatially complete and an approach to address these gaps is required. A hybrid gap filling (Jacobson et al., 2004) approach was designed using temporal interpolation between the missing value and the closest available observations or from the long term average record when the gaps are persistent for periods longer than 3 months.

Using the continuity dataset various long term average VI records were calculated at a 5, 10, 20 and 30 year intervals. Each long term average is then gapfilled with a simple linear interpolation. Once a missing observation is identified we search for the two closest observations within a 4-month interval in the 30-year average, and 3 months in the other periods. While 3 and 4 months seems like a long period, gaps are usually filled with interpolated data from the most immediate observations, the need for3-4 months is only limited to areas with very persistent clouds. Getting good quality observations in sequence is almost impossible in these area unless the period is expanded. At the end of this process the 30-year dataset is gap filled. Once the 30-year long term

average is gap filled, we gap fill the remaining 20, 10 and 5-year long term average records. First the 20year is gapfilled from the 30-yr, the 10-yrs is gapfilled from the 20-yr, and finally the 5-yr is gapfilled from the 10-yr. This ensures the use of data from the shortest possible long term period (we note that when the 20-year long term has gaps the 10 and 5 year will also have gaps).

A similar process is applied to the daily ESDR VI time series. First attempt to gap fill is based on a simple linear interpolation with data from the current year and most immediate available observations. If no data exist, the pixel is replaced with a value from the 5yr long term average that overlaps the year in question. In case the user wishes not to use gap filled data the interpolated pixels in the record are identified with their VIP Rank layer which informs the users if the data was gap filled using a linear interpolation or from long term average.
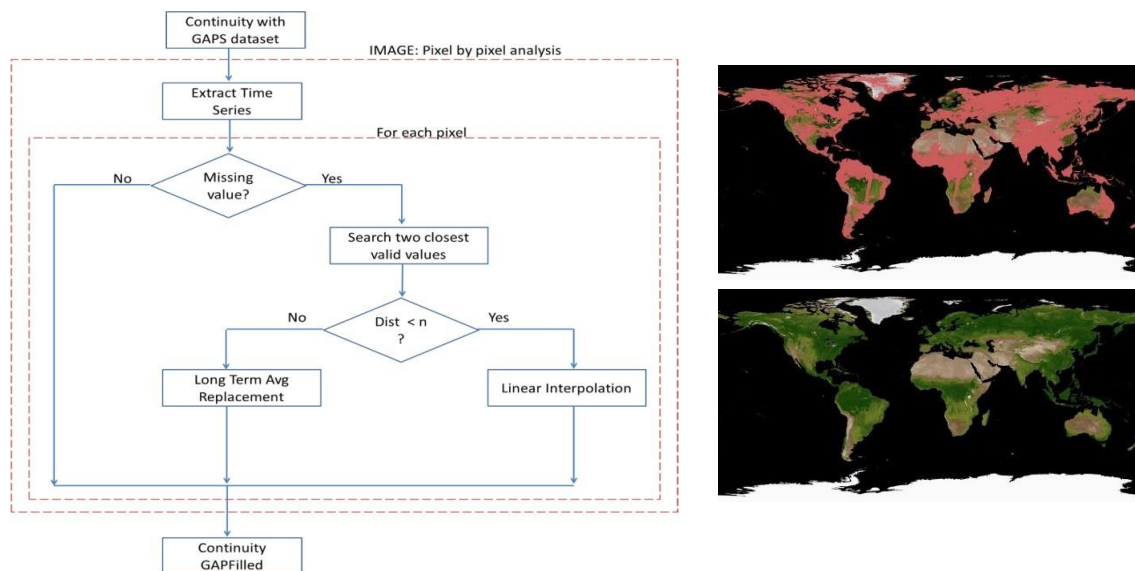
Figure 21. Hybrid gap filling algorithm flow diagram and example images (DOY 209, 2008)

4.2.   Scientific Data Sets

4.2.1.   VIP01 SDS Structure

The CMG 0.05 Deg VIP01 VI product contains the following scientific data

Table 5. VIP01 File SDS structure

Science Data set	Units	DataType	Valid Range	Fill[a]	Scale
CMG 0.05 Deg Daily NDVI	NDVI	INT16	-10,000 10,000	-15,000	10,000
CMG 0.05 Deg Daily EVI2	EVI2	INT16	-10,000 10,000	-15,000	10,000
CMG 0.05 Deg Daily VI Quality	bits	UINT16	0 -65535	65535	N/A
CMG 0.05 Deg Daily Pixel Reliability	Rank	INT8	-4 - 11	-4	1
CMG 0.05 Deg Daily RED reflectance	Reflectance	INT16	-100 16,000	-28,672	10,000
CMG 0.05 Deg Daily NIR reflectance	Reflectance	INT16	-100 16,000	-28,672	10,000
CMG 0.05 Deg Daily BLUE reflectance[b]	Reflectance	INT16	-100 16,000	-28,672	10,000
CMG 0.05 Deg Daily MIR reflectance[b]	Reflectance	INT16	-100 16,000	-28,672	10,000
CMG 0.05 Deg Daily Solar Zenith Angle	Degree	INT16	0 18,000	0	100
CMG 0.05 Deg Daily View Zenith Angle	Degree	INT16	0 18,000	0	100
CMG 0.05 Deg Daily Relative Azimuth Angle[c]	Degree	INT16	-18,000 18,000	0	100

a While we are listing only a single Fill_Value in practice there are multiple Fill Values to separate between the different reasons of missing data. For example, the Pixel Reliability has 4 different value for Fill depending of where the pixel is located (-1 over land, -2 over high latitude, -3 over Antarctica, and -4 over Water/Ocean).

b *BLUE and MIR SDSs are only available for MODIS (2000-2014). On the VIP data distribution server (vip.arizona.edu) AVHRR shows False Color composite images that are based on MNR. The Middle Infrared reflectance data used for AVHRR was generated using a MODIS based model that correlates the Red and NIR with the MIR band. The model is then applied to AVHRR Red/NIR to generate a false MIR reflectance data useful for image generation.*

*c The Relative Azimuth Angle (RAA) from the MOIS era, which is simply passed along from the MODIS data, was computed based on an absolute value of all the finer resolution pixels, resulting in only positive values and minor usefulness. For AVHRR data relative Azimuth Angle, which is also passed along in our processing from Input provided by the LTDR group, has an issue with the valid range. The actual range is [-360° to 360°] and it should have been [-180° to 180°]. To correct the range users can convert the value using the following simple routine (*https://ltdr.nascom.nasa.gov/ltdr/docs/AVHRR_LTDR_V4_Document.pdf*):*

SinRelativeAz=sin(RAA)

CosRelativeAz=cos(RAA)

Correct-RAA = atan2(SinRelativeAz,CosRelativeAz)

4.2.2. Pixel Reliability SDS Description

While a comprehensive VI Quality SDS is provided with each product, the complexity of the bit layout (inherited from MODIS) could be difficult or inaccessible to average users looking for a quick method to evaluate the data quality and decide what to reject for their specific applications. To address this complexity, the MODIS VI product, and starting C5, proposed the implementation of a pixel reliability metric that provides an easy approach to assess the quality of the pixel and decide on its usefulness (Didan and Huete, 2006). This metric provides a simple and direct mask for filtering the datasets (Table 4).

4.3. Product Specific Metadata

The metadata fields used in the VIP01 product suite are listed below. They provide a summary of the file quality based on the frequency of the reliability rank analysis.

- QAPERCENTEXCELLENT
- QAPERCENTGOOD
- QAPERCENTACCEPTABLE
- QAPERCENTMARGINAL
- QAPERCENTPASS
- QAPERCENTQUESTIONABLE
- QAPERCENTPOOR
- QAPERCENTCLOUDSHADOW
- QAPERCENTSNOW
- QAPERCENTCLOUD
- QAPERCENTESTIMATED
  QAPERCENTNODATA

4.4. Global and Local Metadata Attributes

All VIP products contain global metadata that is written during the product generation. This global metadata is useful for archiving, searching, and ordering the product and provides other key attributes about the file.

4.4.1. Global Metadata Attributes

Each HDF file contains the following attribute object structure (slightly abridged).

```
Global attributes: 4
HDFEOSVersion: HDFEOS_V2.19GROUP=SwathStructure
END_GROUP=SwathStructure
GROUP=GridStructure
    GROUP=GRID_1
        GridName="VIP_CMG_GRID"
        XDim=7200
        YDim=3600
        UpperLeftPointMtrs=(-180000000.000000,90000000.000000)
        LowerRightMtrs=(180000000.000000,-90000000.000000)
        Projection=GCTP_GEO
        GROUP=Dimension
        END_GROUP=Dimension
        GROUP=DataField
            OBJECT=DataField_1
                DataFieldName="CMG 0.05 Deg Daily NDVI"
                DataType=DFNT_INT16
                DimList=("YDim","XDim")
            END_OBJECT=DataField_1
            OBJECT=DataField_2
                DataFieldName="CMG 0.05 Deg Daily EVI2"
                DataType=DFNT_INT16
                DimList=("YDim","XDim")
            END_OBJECT=DataField_2
            OBJECT=DataField_3
                DataFieldName="CMG 0.05 Deg Daily VI Quality"
                DataType=DFNT_UINT16
                DimList=("YDim","XDim")
            END_OBJECT=DataField_3
            OBJECT=DataField_4
                DataFieldName="CMG 0.05 Deg Daily Pixel Reliability"
                DataType=DFNT_INT8
                DimList=("YDim","XDim")
            END_OBJECT=DataField_4
            OBJECT=DataField_5
                DataFieldName="CMG 0.05 Deg Daily RED reflectance"
                DataType=DFNT_INT16
                DimList=("YDim","XDim")
            END_OBJECT=DataField_5
            OBJECT=DataField_6
                DataFieldName="CMG 0.05 Deg Daily NIR reflectance"
                DataType=DFNT_INT16
                DimList=("YDim","XDim")
```

```
                    END_OBJECT=DataField_6
                    OBJECT=DataField_7
                        DataFieldName="CMG 0.05 Deg Daily BLUE reflectance"
                        DataType=DFNT_INT16
                        DimList=("YDim","XDim")
                    END_OBJECT=DataField_7
                    OBJECT=DataField_8
                        DataFieldName="CMG 0.05 Deg Daily MIR reflectance"
                        DataType=DFNT_INT16
                        DimList=("YDim","XDim")
                    END_OBJECT=DataField_8
                    OBJECT=DataField_9
                        DataFieldName="CMG 0.05 Deg Daily Solar Zenith Angle"
                        DataType=DFNT_INT16
                        DimList=("YDim","XDim")
                    END_OBJECT=DataField_9
                    OBJECT=DataField_10
                        DataFieldName="CMG 0.05 Deg Daily View Zenith Angle"
                        DataType=DFNT_INT16
                        DimList=("YDim","XDim")
                    END_OBJECT=DataField_10
                    OBJECT=DataField_11
                        DataFieldName="CMG 0.05 Deg Daily Relative Azimuth Angle"
                        DataType=DFNT_INT16
                        DimList=("YDim","XDim")
                    END_OBJECT=DataField_11
                END_GROUP=DataField
                GROUP=MergedFields
                END_GROUP=MergedFields
            END_GROUP=GRID_1
        END_GROUP=GridStructure
        GROUP=PointStructure
        END_GROUP=PointStructure
        END
        GROUP = INVENTORYMETADATA
         GROUPTYPE = MASTERGROUP
         GROUP  = ECSDATAGRANULE
         OBJECT = LOCALGRANULEID
           NUM_VAL = 1
           VALUE = "VIP01.A2013256.004.2016074222238.hdf"
          END_OBJECT = LOCALGRANULEID
          OBJECT = PRODUCTIONDATETIME
           NUM_VAL = 1
           VALUE = "2016-03-14T22:22:38.000Z"
          END_OBJECT = PRODUCTIONDATETIME
          OBJECT = REPROCESSINGACTUAL
          NUM_VAL = 1
           VALUE = "reprocessed"
          END_OBJECT = REPROCESSINGACTUAL
          OBJECT = LOCALVERSIONID
           NUM_VAL = 1
           VALUE = "4.1.0"
          END_OBJECT = LOCALVERSIONID
```

```
END_GROUP = ECSDATAGRANULE
GROUP = COLLECTIONDESCRIPTIONCLASS
OBJECT = VERSIONID
  NUM_VAL = 1
  VALUE = 4
 END_OBJECT = VERSIONID
 OBJECT = SHORTNAME
 NUM_VAL = 1
  VALUE = "VIP01"
 END_OBJECT = SHORTNAME
END_GROUP = COLLECTIONDESCRIPTIONCLASS
GROUP = INPUTGRANULE
 OBJECT = INPUTPOINTER
  NUM_VAL = "1"
  VALUE = "MOD09CMG.A2013256.005.2013261001331.hdf"
 END_OBJECT = INPUTPOINTER
END_GROUP = INPUTGRANULE
GROUP = SPATIALDOMAINCONTAINER
 GROUP = HORIZONTALSPATIALDOMAINCONTAINER
  GROUP = BOUNDINGRECTANGLE
   OBJECT = EASTBOUNDINGCOORDINATE
    NUM_VAL = 1
    VALUE = 180.0
   END_OBJECT = EASTBOUNDINGCOORDINATE
   OBJECT = WESTBOUNDINGCOORDINATE
   NUM_VAL = 1
    VALUE = -180.0
   END_OBJECT = WESTBOUNDINGCOORDINATE
   OBJECT = SOUTHBOUNDINGCOORDINATE
   NUM_VAL = 1
    VALUE = -90.0
   END_OBJECT = SOUTHBOUNDINGCOORDINATE
   OBJECT = NORTHBOUNDINGCOORDINATE
   NUM_VAL = 1
    VALUE = 90.0
   END_OBJECT = NORTHBOUNDINGCOORDINATE
  END_GROUP = BOUNDINGRECTANGLE
 END_GROUP = HORIZONTALSPATIALDOMAINCONTAINER
 GROUP = GRANULELOCALITY
  OBJECT = LOCALITYVALUE
   NUM_VAL = 1
   VALUE = "Global"
  END_OBJECT = LOCALITYVALUE
 END_GROUP = GRANULELOCALITY
END_GROUP = SPATIALDOMAINCONTAINER
GROUP = RANGEDATETIME
 OBJECT = RANGEENDINGDATE
  NUM_VAL = 1
  VALUE = "2013-09-13"
 END_OBJECT = RANGEENDINGDATE
 OBJECT = RANGEENDINGTIME
 NUM_VAL = 1
  VALUE = "23:59:59"
```

```
    END_OBJECT = RANGEENDINGTIME
    OBJECT = RANGEBEGINNINGDATE
    NUM_VAL = 1
     VALUE = "2013-09-13"
    END_OBJECT = RANGEBEGINNINGDATE
    OBJECT = RANGEBEGINNINGTIME
    NUM_VAL = 1
     VALUE = "00:00:00"
    END_OBJECT = RANGEBEGINNINGTIME
  END_GROUP = RANGEDATETIME
  GROUP = PGEVERSIONCLASS
    OBJECT = PGEVERSION
     NUM_VAL = 1
     VALUE = "4.1"
    END_OBJECT = PGEVERSION
  END_GROUP = PGEVERSIONCLASS
END_GROUP = INVENTORYMETADATA
END
GROUP = ARCHIVEDMETADATA
 GROUPTYPE = MASTERGROUP
 OBJECT = ALGORITHMPACKAGENAME
 NUM_VAL = 1
   VALUE = "VIPLAB_CONT_V4"
 END_OBJECT = ALGORITHMPACKAGENAME
 OBJECT = ALGORITHMPACKAGEVERSION
 NUM_VAL = 1
   VALUE = "4"
 END_OBJECT = ALGORITHMPACKAGEVERSION
 OBJECT = LONGNAME
   NUM_VAL = 1
   VALUE = "VIP Vegetation Indices Daily Global 0.05 Deg CMG"
 END_OBJECT = LONGNAME
 OBJECT = PROCESSINGCENTER
   NUM_VAL = 1
   VALUE = "VIPLAB"
 END_OBJECT = PROCESSINGCENTER
 OBJECT = SEAPROCESSED
 NUM_VAL = 1
   VALUE = "No"
 END_OBJECT = SEAPROCESSED
 OBJECT = PROCESSINGENVIRONMENT
 NUM_VAL = 1
   VALUE = "Linux version 3.10.0-327.4.5.el7.x86_64 (Red Hat 4.8.5-4) Intel(R) Xeon(R) CPU X5690 @
3.47GHz"
 END_OBJECT = PROCESSINGENVIRONMENT
 OBJECT = DESCRREVISION
   NUM_VAL = 1
   VALUE = "4.0"
 END_OBJECT = DESCRREVISION
 OBJECT = VIPSCIENCEQUALITYFLAGEXPLANATION
  NUM_VAL = 1
  VALUE = "See https://vip.arizona.edu/documentation/esdr/?VIP01v04"
 END_OBJECT = VIPSCIENCEQUALITYFLAGEXPLANATION
```

```
OBJECT = DATACOLUMNS
 NUM_VAL = 1
 VALUE = 7200
END_OBJECT = DATACOLUMNS
OBJECT = DATAROWS
NUM_VAL = 1
 VALUE = 3600
END_OBJECT = DATAROWS
OBJECT = GLOBALGRIDCOLUMNS
 NUM_VAL = 1
 VALUE = 7200
END_OBJECT = GLOBALGRIDCOLUMNS
OBJECT = GLOBALGRIDROWS
NUM_VAL = 1
 VALUE = 3600
END_OBJECT = GLOBALGRIDROWS
OBJECT = NUMBEROFDAYS
NUM_VAL = 1
 VALUE = "1"
END_OBJECT = NUMBEROFDAYS
OBJECT = DAYSPROCESSED
NUM_VAL = "1"
 VALUE = "2013256"
END_OBJECT = DAYSPROCESSED
OBJECT = GEOANYABNORMAL
NUM_VAL = 1
 VALUE = "False"
END_OBJECT = GEOANYABNORMAL
OBJECT = GEOESTMAXRMSERROR
NUM_VAL = 1
 VALUE = -1.0
END_OBJECT = GEOESTMAXRMSERROR
GROUP = QASTATS


 END_GROUP = QASTATS
END_GROUP = ARCHIVEDMETADATA
END
```

4.5. Quality Assurance

The quality of each VIP product is stored in Quality Assessment (QA) metadata objects and QA science data sets (SDS). The QA metadata objects summarize global level file or product quality with single words and numeric values, and thus are useful for data ordering and screening. The QA SDS, on the other hand, documents the product quality on a pixel-by-pixel basis and thus is useful for data analyses and application.

Table 6. VIP01 quality assurance description

Bit	Description	Values
13-15	Land Water Flag	000 shallow ocean
		001 land
		010 ocean coastlines and lake shorelines
		011 shallow inland water

		100 ephemeral water
		101 deep inland water
		110 continental/moderate ocean
		111 deep ocean
12	View Angle > 30	1 yes
		0 no
11	Sun Zenith Angle > 85	1 yes
		0 no
10	Sun Zenith Angle > 75	1 yes
		0 no
8-9	Gap Fill Status	00 not gap filled
		01 interpolated
		10 long term average
		11 not set
7	Snow/Ice estimated	1 yes
		0 no
6	Snow/Ice flag	1 yes
		0 no
5	Aerosol estimated	1 yes
		0 no
3-4	Aerosol Quantity	00 climatology
		01 low
		10 average
		11 high
2	Cloud Shadow	1 yes
		0 no
0-1	Cloud State	00 clear
		01 cloudy
		10 mixed
		11 not set, assumed clear

Bits are listed from MSB (bit 15) to the LSB (bit 0)

5. VIP Composited Products

5.1. Product Series

Several composited datasets of different period length are produced from the daily filtered VIP01 dataset. The compositing is applied to pre-continuity and pre-gap filling to minimize bias and allow for a composite product specific continuity and gap filling approach. Compositing is capable of gap filing multi-day data and hence needed to be applied to pre continuity and gap filling data. The file names are labeled with the first day of the composite period. In addition, there a Composite Day of the Year SDS layer that stores the day of the year selected for each pixel (Table 7).

5.1.1. VIP07 (0.05-deg) VI Product Series

This multi-day product ingests 7 daily images as input and produces a single composited output image. The annual time series consists of 52 images per year.

5.1.2.   VIP15 (0.05-deg) VI Product Series

Although labeled as a 15-day composited product, the VIP15 does not strictly follow a 15-day periodicity, rather it creates two composites per month. The first composite always covers day 1 to 15 of each month, and the second composite covers the remaining days of the month. The VIP15 annual time series consists of 24 images per year.
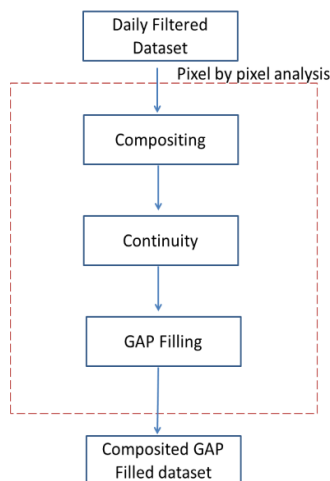
Figure 22. VIP composited products processing flow diagram.

5.1.3.   VIP30 (0.05-deg) VI Product Series

VIP30 is a monthly product that ingests daily images covering the month and the resulting annual time series consists of 12 images. The product does not use 30 days rather the full number of day for the considered month.

For consistency and convenience these composited VIP products are created by the same science algorithm (Constrained View Maximum Value Composite, Huete et al. 2002) and share the same HDF-EOS SDS structures, attributes, and global metadata. The only difference is in the number of days used as input. The following algorithm description, product specific metadata, and global metadata sections apply to all the VIP Composited Product series. For reference XX represents 07, 15 or 30 to indicate the composite period.

5.2.   Algorithm Description

The VIPXX compositing algorithm operates on a per-pixel basis and requires multiple observations (daily) to generate a single composited VI value that will represent the period. The algorithm uses a strict filtering approach based on the per-pixel reliability Rank from the 2nd stage filtered VIP01 daily product suite. Since the inputs are already pre-filtered the compositing reduces to discarding less than ideal data, and selecting the day with the highest NDVI value. The day with the lowest view angle from the top two NDVI values is finally selected to represent the period.

Cloud-contaminated pixels and pixels with extreme off-nadir sensor views are considered lower quality and avoided. A cloud-free, close to nadir view pixel with no residual atmospheric contamination represents the best observation. Because of the way the Rank SDS was created, it is easy to identify pixels free of cloud and with acceptable viewing geometry. If no good quality pixels are found in a given compositing period, a NoData value (Fill Value) is assigned to the output and the pixel is later gap filled using the same gap filling algorithm described earlier (similarly to the VIP01 product).
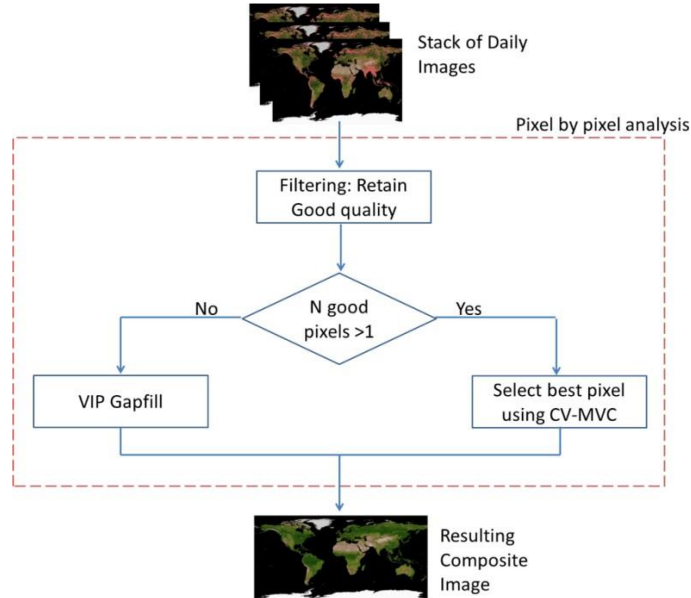
Figure 23. VIP VI compositing algorithm flow diagram

Once the composited product is generated, the continuity algorithm and gap filling process are applied in the same way as for the daily data, thus resulting in sensor independent gap-filled composited products series.

5.3. VIPXX SDS Structure

The CMG 0.05 Deg VIPXX VI product has the following layers:

Table 7. VIP composited products SDS structure

Science Data set	Units	DataType	Valid Range	Fill[a]	Scale
CMG 0.05 Deg NDVI	NDVI	INT16	-10,000 10,000	-15,000	10,000
CMG 0.05 Deg EVI2	EVI2	INT16	-10,000 10,000	-15,000	10,000
CMG 0.05 Deg VI Quality	bits	UINT16	0,65535	65535	N/A
CMG 0.05 Deg Pixel Reliability	Rank	INT8	-4 - 11	-4	1
CMG 0.05 Deg Composite Day of the Year	Day	INT16	1 366	-1	1
CMG 0.05 Deg RED reflectance	Reflectance	INT16	-100 16,000	-28,672	10,000
CMG 0.05 Deg NIR reflectance	Reflectance	INT16	-100 16,000	-28,672	10,000
CMG 0.05 Deg BLUE reflectance[b]	Reflectance	INT16	-100 16,000	-28,672	10,000
CMG 0.05 Deg MIR reflectance[b]	Reflectance	INT16	-100 16,000	-28,672	10,000
CMG 0.05 Deg Solar Zenith Angle	Degree	INT16	0 18,000	0	100
CMG 0.05 Deg View Zenith Angle	Degree	INT16	0 18,000	0	100
CMG 0.05 Deg Relative Azimuth Angle[c]	Degree	INT16	-18,000 18,000	0	100

[a] [b] [b] see notes in Table 5

5.4. Product Specific Metadata

The metadata fields used for QA evaluation of the VIPXX VI products are listed below:

- QAPERCENTEXCELLENT
- QAPERCENTGOOD

- QAPERCENTACCEPTABLE
- QAPERCENTMARGINAL
- QAPERCENTPASS
- QAPERCENTQUESTIONABLE
- QAPERCENTPOOR
- QAPERCENTCLOUDSHADOW
- QAPERCENTSNOW
- QAPERCENTCLOUD
- QAPERCENTESTIMATED
- QAPERCENTNODATA
- RANGEDATETIME
- RANGEBEGINNINGDATE
- RANGEBEGINNINGTIME
- RANGEENDINGDATE
- RANGEENDINGTIME

5.5. Global and Local Metadata Attributes

Following the MODIS product suite metadata approach, the global metadata for the composited VIP products is written to the output file during the generation process. This metadata is used during the search/order process of archives. A listing of relevant metadata is provided below (slightly abridged):

```
Global attributes: 4
HDFEOSVersion: HDFEOS_V2.19GROUP=SwathStructure
END_GROUP=SwathStructure
GROUP=GridStructure
    GROUP=GRID_1
        GridName="VIP_CMG_GRID"
        XDim=7200
        YDim=3600
        UpperLeftPointMtrs=(-180000000.000000,90000000.000000)
        LowerRightMtrs=(180000000.000000,-90000000.000000)
        Projection=GCTP_GEO
        GROUP=Dimension
        END_GROUP=Dimension
        GROUP=DataField
            OBJECT=DataField_1
                DataFieldName="CMG 0.05 Deg MONTHLY NDVI"
                DataType=DFNT_INT16
                DimList=("YDim","XDim")
            END_OBJECT=DataField_1
            OBJECT=DataField_2
                DataFieldName="CMG 0.05 Deg MONTHLY EVI2"
                DataType=DFNT_INT16
                DimList=("YDim","XDim")
            END_OBJECT=DataField_2
            OBJECT=DataField_3
                DataFieldName="CMG 0.05 Deg MONTHLY VI Quality"
                DataType=DFNT_UINT16
```

```
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_3
                OBJECT=DataField_4
                    DataFieldName="CMG 0.05 Deg MONTHLY Pixel Reliability"
                    DataType=DFNT_INT8
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_4
                OBJECT=DataField_5
                    DataFieldName="CMG 0.05 Deg MONTHLY Composite Day of the Year"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_5
                OBJECT=DataField_6
                    DataFieldName="CMG 0.05 Deg MONTHLY RED reflectance"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_6
                OBJECT=DataField_7
                    DataFieldName="CMG 0.05 Deg MONTHLY NIR reflectance"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_7
                OBJECT=DataField_8
                    DataFieldName="CMG 0.05 Deg MONTHLY BLUE reflectance"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_8
                OBJECT=DataField_9
                    DataFieldName="CMG 0.05 Deg MONTHLY MIR reflectance"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_9
                OBJECT=DataField_10
                    DataFieldName="CMG 0.05 Deg MONTHLY Solar Zenith Angle"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_10
                OBJECT=DataField_11
                    DataFieldName="CMG 0.05 Deg MONTHLY View Zenith Angle"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_11
                OBJECT=DataField_12
                    DataFieldName="CMG 0.05 Deg MONTHLY Relative Azimuth Angle"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_12
            END_GROUP=DataField
            GROUP=MergedFields
            END_GROUP=MergedFields
        END_GROUP=GRID_1
    END_GROUP=GridStructure
    GROUP=PointStructure
```

END_GROUP=PointStructure
END
GROUP = INVENTORYMETADATA
 GROUPTYPE = MASTERGROUP
 GROUP = ECSDATAGRANULE

  OBJECT = LOCALGRANULEID
   NUM_VAL = 1
   VALUE = "VIP30.A2013001.004.2016071163558.hdf"
  END_OBJECT = LOCALGRANULEID

  OBJECT = PRODUCTIONDATETIME
   NUM_VAL = 1
   VALUE = "2016-03-11T16:35:58.000Z"
  END_OBJECT = PRODUCTIONDATETIME

  OBJECT = REPROCESSINGACTUAL
   NUM_VAL = 1
   VALUE = "reprocessed"
  END_OBJECT = REPROCESSINGACTUAL

  OBJECT = LOCALVERSIONID
   NUM_VAL = 1
   VALUE = "4.1.0"
  END_OBJECT = LOCALVERSIONID
 END_GROUP = ECSDATAGRANULE

 GROUP = COLLECTIONDESCRIPTIONCLASS
  OBJECT = VERSIONID
   NUM_VAL = 1
   VALUE = 4
  END_OBJECT = VERSIONID

  OBJECT = SHORTNAME
   NUM_VAL = 1
   VALUE = "VIP30"
  END_OBJECT = SHORTNAME
 END_GROUP = COLLECTIONDESCRIPTIONCLASS

 GROUP = INPUTGRANULE
  OBJECT = INPUTPOINTER
  NUM_VAL = 31
   VALUE   =
("MOD09CMG.A2013001.005.2013003045426.hdf","MOD09CMG.A2013002.005.2013005004722.hdf","
MOD09CMG.A2013003.005.2013008004946.hdf","MOD09CMG.A2013004.005.2013011210405.hdf",
"MOD09CMG.A2013005.005.2013011211022.hdf","MOD09CMG.A2013006.005.2013008055744.hdf","
MOD09CMG.A2013007.005.2013009051804.hdf","MOD09CMG.A2013008.005.2013010065750.hdf",
"MOD09CMG.A2013009.005.2013011051959.hdf","MOD09CMG.A2013010.005.2013012085924.hdf","
MOD09CMG.A2013011.005.2013013053756.hdf","MOD09CMG.A2013012.005.2013014060335.hdf",
"MOD09CMG.A2013013.005.2013015055800.hdf","MOD09CMG.A2013014.005.2013016054135.hdf","
MOD09CMG.A2013015.005.2013017052419.hdf","MOD09CMG.A2013016.005.2013018061305.hdf",
"MOD09CMG.A2013017.005.2013019044513.hdf","MOD09CMG.A2013018.005.2013020053652.hdf","
MOD09CMG.A2013019.005.2013021052208.hdf","MOD09CMG.A2013020.005.2013022050329.hdf",

"MOD09CMG.A2013021.005.2013023060824.hdf","MOD09CMG.A2013022.005.2013024053947.hdf","
MOD09CMG.A2013023.005.2013025052000.hdf","MOD09CMG.A2013024.005.2013026052533.hdf",
"MOD09CMG.A2013025.005.2013027051232.hdf","MOD09CMG.A2013026.005.2013028053420.hdf","
MOD09CMG.A2013027.005.2013029052844.hdf","MOD09CMG.A2013028.005.2013030052513.hdf",
"MOD09CMG.A2013029.005.2013031055303.hdf","MOD09CMG.A2013030.005.2013032050117.hdf","
MOD09CMG.A2013031.005.2013040021203.hdf")
  END_OBJECT = INPUTPOINTER
 END_GROUP = INPUTGRANULE
 GROUP = SPATIALDOMAINCONTAINER
  GROUP = HORIZONTALSPATIALDOMAINCONTAINER
   GROUP = BOUNDINGRECTANGLE
    OBJECT = EASTBOUNDINGCOORDINATE
     NUM_VAL = 1
     VALUE = 180.0
    END_OBJECT = EASTBOUNDINGCOORDINATE
    OBJECT = WESTBOUNDINGCOORDINATE
    NUM_VAL = 1
     VALUE = -180.0
    END_OBJECT = WESTBOUNDINGCOORDINATE
    OBJECT = SOUTHBOUNDINGCOORDINATE
    NUM_VAL = 1
     VALUE = -90.0
    END_OBJECT = SOUTHBOUNDINGCOORDINATE
    OBJECT = NORTHBOUNDINGCOORDINATE
    NUM_VAL = 1
     VALUE = 90.0
    END_OBJECT = NORTHBOUNDINGCOORDINATE
   END_GROUP = BOUNDINGRECTANGLE
  END_GROUP = HORIZONTALSPATIALDOMAINCONTAINER
  GROUP = GRANULELOCALITY
   OBJECT = LOCALITYVALUE
    NUM_VAL = 1
    VALUE = "Global"
   END_OBJECT = LOCALITYVALUE
  END_GROUP = GRANULELOCALITY
 END_GROUP = SPATIALDOMAINCONTAINER
 GROUP = RANGEDATETIME
  OBJECT = RANGEENDINGDATE
   NUM_VAL = 1
   VALUE = "2013-01-31"
  END_OBJECT = RANGEENDINGDATE
  OBJECT = RANGEENDINGTIME
  NUM_VAL = 1
   VALUE = "23:59:59"
  END_OBJECT = RANGEENDINGTIME
  OBJECT = RANGEBEGINNINGDATE
  NUM_VAL = 1
   VALUE = "2013-01-01"
  END_OBJECT = RANGEBEGINNINGDATE
  OBJECT = RANGEBEGINNINGTIME
  NUM_VAL = 1
   VALUE = "00:00:00"
  END_OBJECT = RANGEBEGINNINGTIME

```
   END_GROUP = RANGEDATETIME
  GROUP = PGEVERSIONCLASS
  OBJECT = PGEVERSION
     NUM_VAL = 1
     VALUE = "4.1"
   END_OBJECT = PGEVERSION
  END_GROUP = PGEVERSIONCLASS
 END_GROUP = INVENTORYMETADATA
END
GROUP = ARCHIVEDMETADATA
 GROUPTYPE = MASTERGROUP
 OBJECT = ALGORITHMPACKAGENAME
 NUM_VAL = 1
   VALUE = "VIPLAB_CONT_V4"
 END_OBJECT = ALGORITHMPACKAGENAME
 OBJECT = ALGORITHMPACKAGEVERSION
 NUM_VAL = 1
   VALUE = "4"
 END_OBJECT = ALGORITHMPACKAGEVERSION
 OBJECT = LONGNAME
   NUM_VAL = 1
   VALUE = "VIP Vegetation Indices Monthly Global 0.05 Deg CMG"
 END_OBJECT = LONGNAME
 OBJECT = PROCESSINGCENTER
   NUM_VAL = 1
   VALUE = "VIPLAB"
 END_OBJECT = PROCESSINGCENTER
 OBJECT = SEAPROCESSED
 NUM_VAL = 1
   VALUE = "No"
 END_OBJECT = SEAPROCESSED
 OBJECT = PROCESSINGENVIRONMENT
 NUM_VAL = 1
   VALUE = "Linux version 3.10.0-327.4.5.el7.x86_64 (Red Hat 4.8.5-4) Intel(R) Xeon(R) CPU X5690 @
3.47GHz"
 END_OBJECT = PROCESSINGENVIRONMENT
 OBJECT = DESCRREVISION
   NUM_VAL = 1
   VALUE = "4.0"
 END_OBJECT = DESCRREVISION
 OBJECT = VIPSCIENCEQUALITYFLAGEXPLANATION
  NUM_VAL = 1
  VALUE = "See https://vip.arizona.edu/documentation/esdr/?VIP30v04"
 END_OBJECT = VIPSCIENCEQUALITYFLAGEXPLANATION
 OBJECT = DATACOLUMNS
   NUM_VAL = 1
   VALUE = 7200
 END_OBJECT = DATACOLUMNS
 OBJECT = DATAROWS
 NUM_VAL = 1
   VALUE = 3600
 END_OBJECT = DATAROWS
 OBJECT = GLOBALGRIDCOLUMNS
```

```
    NUM_VAL = 1
     VALUE = 7200
   END_OBJECT = GLOBALGRIDCOLUMNS
   OBJECT = GLOBALGRIDROWS
  NUM_VAL = 1
     VALUE = 3600
   END_OBJECT = GLOBALGRIDROWS
   OBJECT = NUMBEROFDAYS
   NUM_VAL = 1
    VALUE = "31"
   END_OBJECT = NUMBEROFDAYS
   OBJECT = DAYSPROCESSED
  NUM_VAL = "31"
    VALUE                                                                      =
("2013001","2013002","2013003","2013004","2013005","2013006","2013007","2013008","2013009",""
2013010","2013011",
"2013012","2013013","2013014","2013015","2013016","2013017","2013018","2013019","2013020","2
013021","2013022",
"2013023","2013024","2013025","2013026","2013027","2013028","2013029","2013030","2013031")
   END_OBJECT = DAYSPROCESSED
   OBJECT = GEOANYABNORMAL
    NUM_VAL = 1
    VALUE = "False"
   END_OBJECT = GEOANYABNORMAL
   OBJECT = GEOESTMAXRMSERROR
   NUM_VAL = 1
    VALUE = -1.0
   END_OBJECT = GEOESTMAXRMSERROR
   GROUP = QASTATS

   END_GROUP = QASTATS
  END_GROUP = ARCHIVEDMETADATA
  END
```

5.6. Quality Assurance

Similar to the daily product series the multi-day composited product series stores quality assurance information useful for post-processing. A Quality SDS layer 'VI Quality' contains multi-bit flags that describe various aspects of the pixel/observation (see Table 6).

6. VIPPHEN (0.05-deg) Phenology ESDR

The VIPPHEN product series is a seamless global 3600x7200 pixel data product with 26 SDSs. Each file is approximately 100 MB (internally compressed using the HDF-EOS compression API). This product is based on the daily VIP product series and uses a 3-year moving window average to smooth out any remaining noise in the data. It contains a Pixel Reliability (Rank) layer that provides information about the quality of the input time series used to derive the phenology metrics and reflects to a degree the accuracy of the resulting phenology metric estimation, i.e. poor data may lead to erroneous estimation and vice versa high quality data should support accurate estimation. However, since most phenology metrics are time/date based they tend to be less prone to noise and error in the input.

6.1. Algorithm Description

Phenology metrics are estimated using a modified Half-Max method (White et al. 2009). A 35% delay is used instead of the standard/original 50% (White, 1997), resulting in a longer season, with earlier start and later end dates. The 35% was found to be more accurate especially in regions with a protracted slow emerging growing season. Default thresholds of 0.12 for NDVI and 0.08 for EVI2 were set as minimum global vegetation index values above which consistent and significant vegetation activities are present. Any signal below these numbers corresponds to inactivity or no vegetated areas.

In prior collections (V1 and V2) a seasonal VI dynamic signal was reported in areas with hardly any vegetation activity, such as deserts, this was later found to be related to the sun angle and BRDF effects and soil color. To address these artifacts in V4, a VI Minimum, Maximum, and Averages global maps were created to assist in identifying and masking those areas. Clustering the VI values in maps helps create a mask that is used to filter these regions out by requiring that a minimum ΔVI of 0.05 be required when the max VI is below 0.2 for NDVI. This guarantees a certain level of VI dynamic/activity in order for the algorithm to consider a growing season. The same NDVI based mask is applied to EVI2 also.

The algorithm was further enhanced by analyzing and establishing a set of typical phenological curves to work with, in order to correctly estimate the metrics across the same range of ecosystems and biomes. These typical growing season profiles are discussed below.

6.1.1. Perpetual Snow/Ice Cover with no Growing Season

VI values are negative or close to zero. Land surface is perpetually covered by snow, i.e. Greenland, Icefields, and Glaciers over high mountains.

Typical annual profile

Typical long term time series
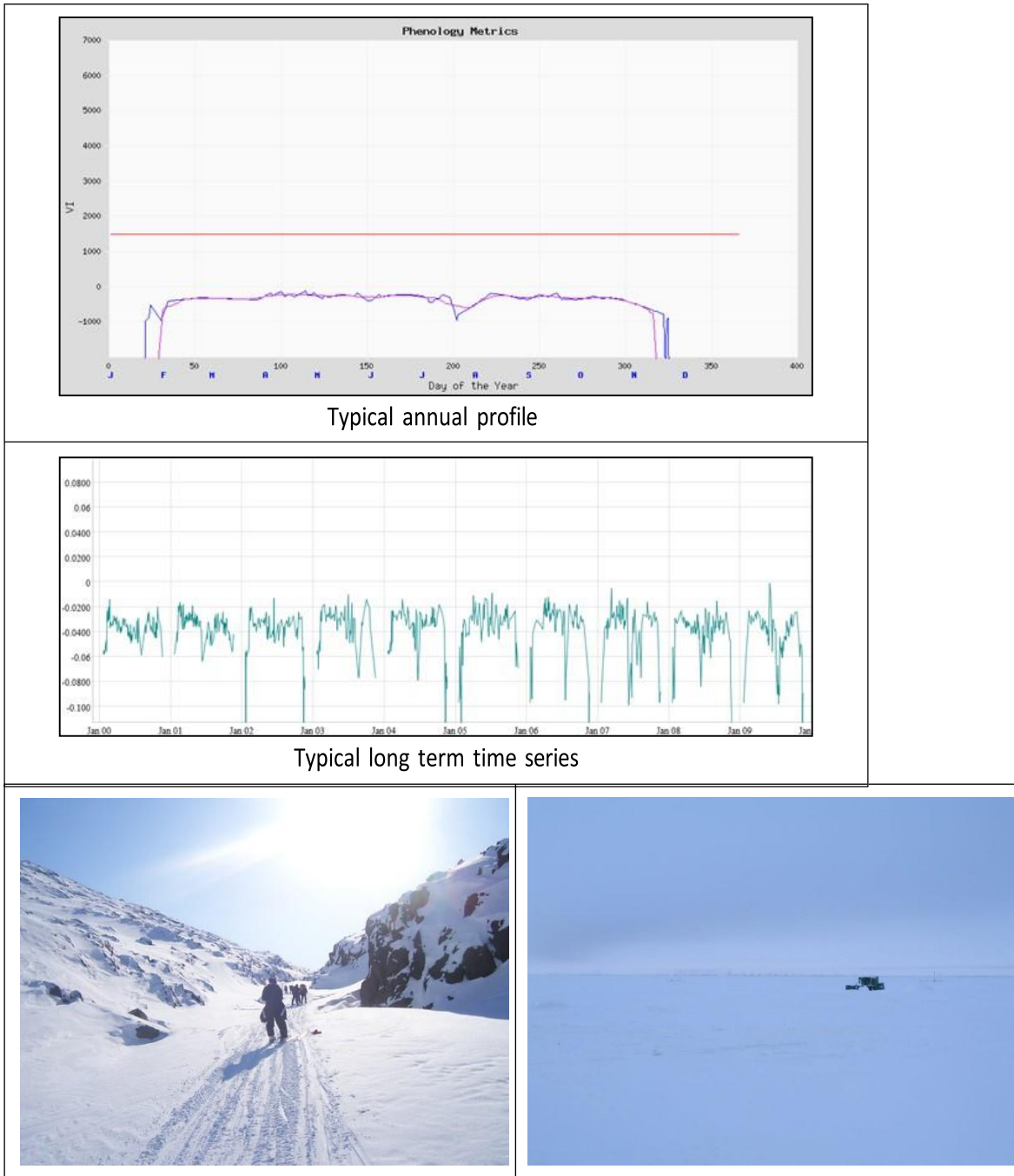
Figure 24. Perpetual snow/ice cover and no seasonality

6.1.2.   Desert and Barren Soil with no Growing Season

Non-vegetated with VI values below the established global minimum VI value. Any changes seen in the time series profiles are mostly due to sun angle, shadowing, or soil color change.
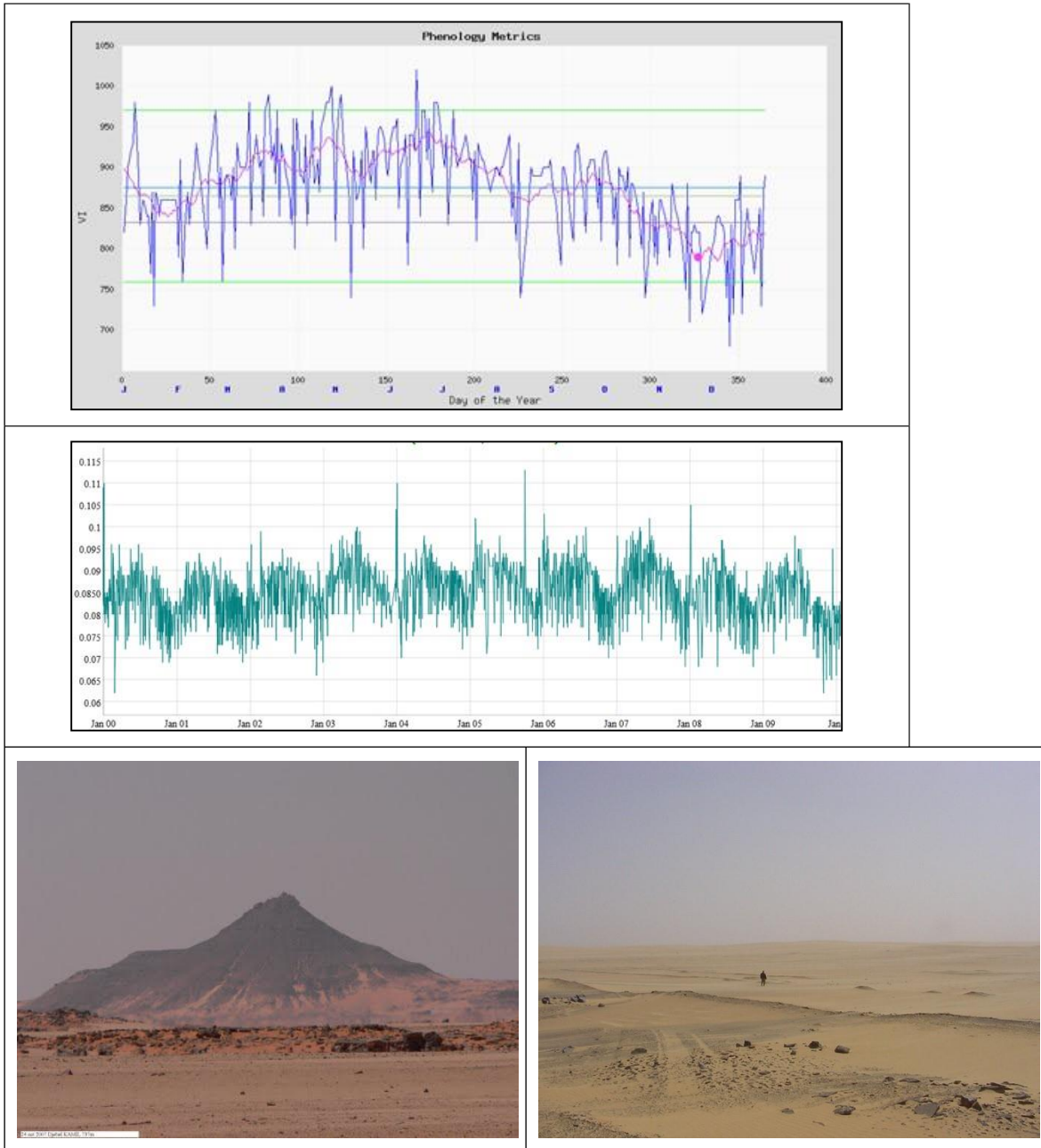
Figure 25. Barren no vegetated and no growing season.

6.1.3. Semi-Arid Regions with Single Minor/Short Growing Season

Sparsely vegetated with generally low VI values. The landscape is covered by sparse bushes, trees, and occasional grasses that may show greening during the occasional rainy season.
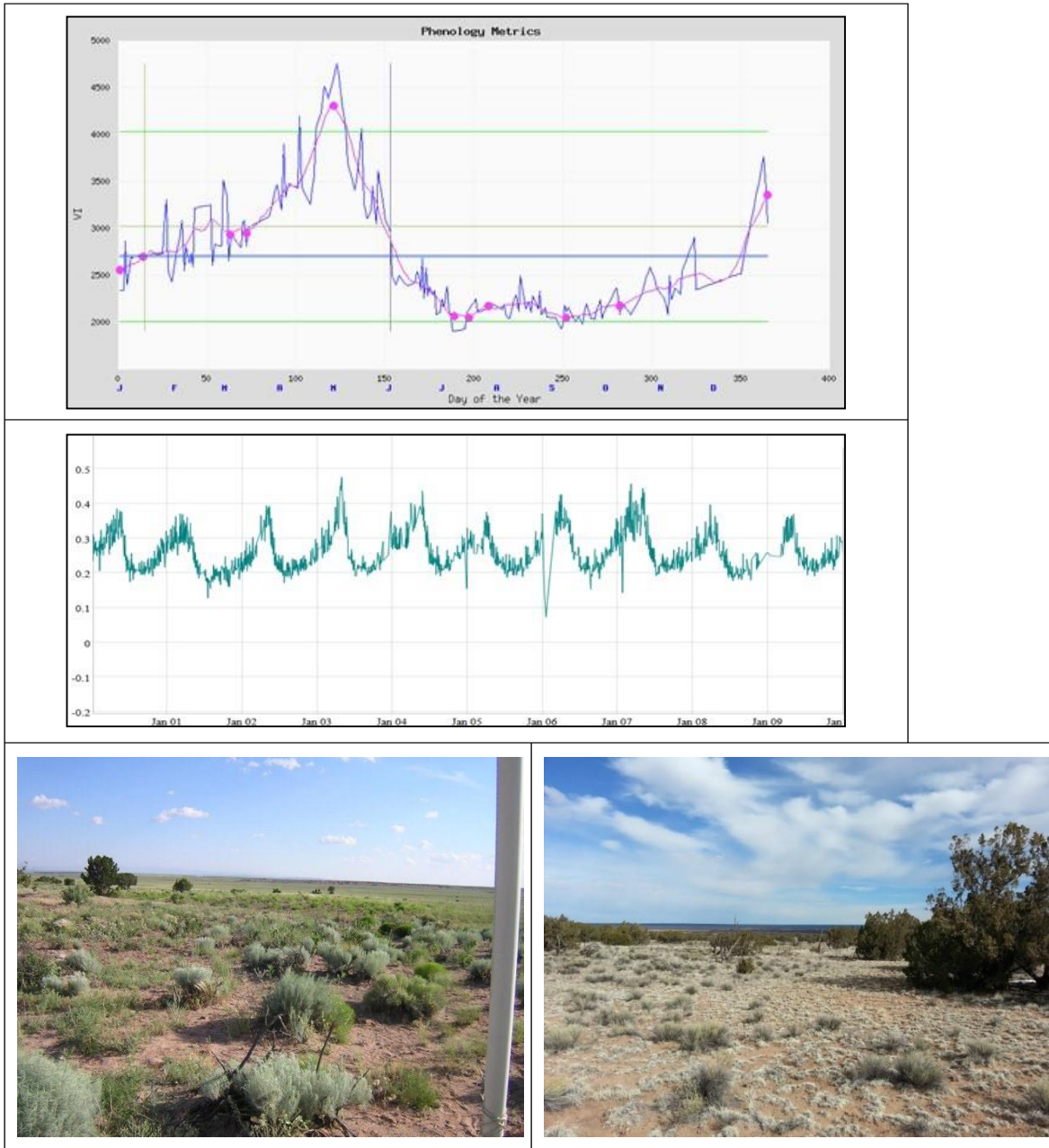
Figure 26. One growing season, semiarid

6.1.4. Tundra with Single Well Defined Growing Season

Very high VI values with a well-defined phenology curve. Snow/Ice dictates the start and end of the growing season. These regions are usually temperature and sunlight limited.
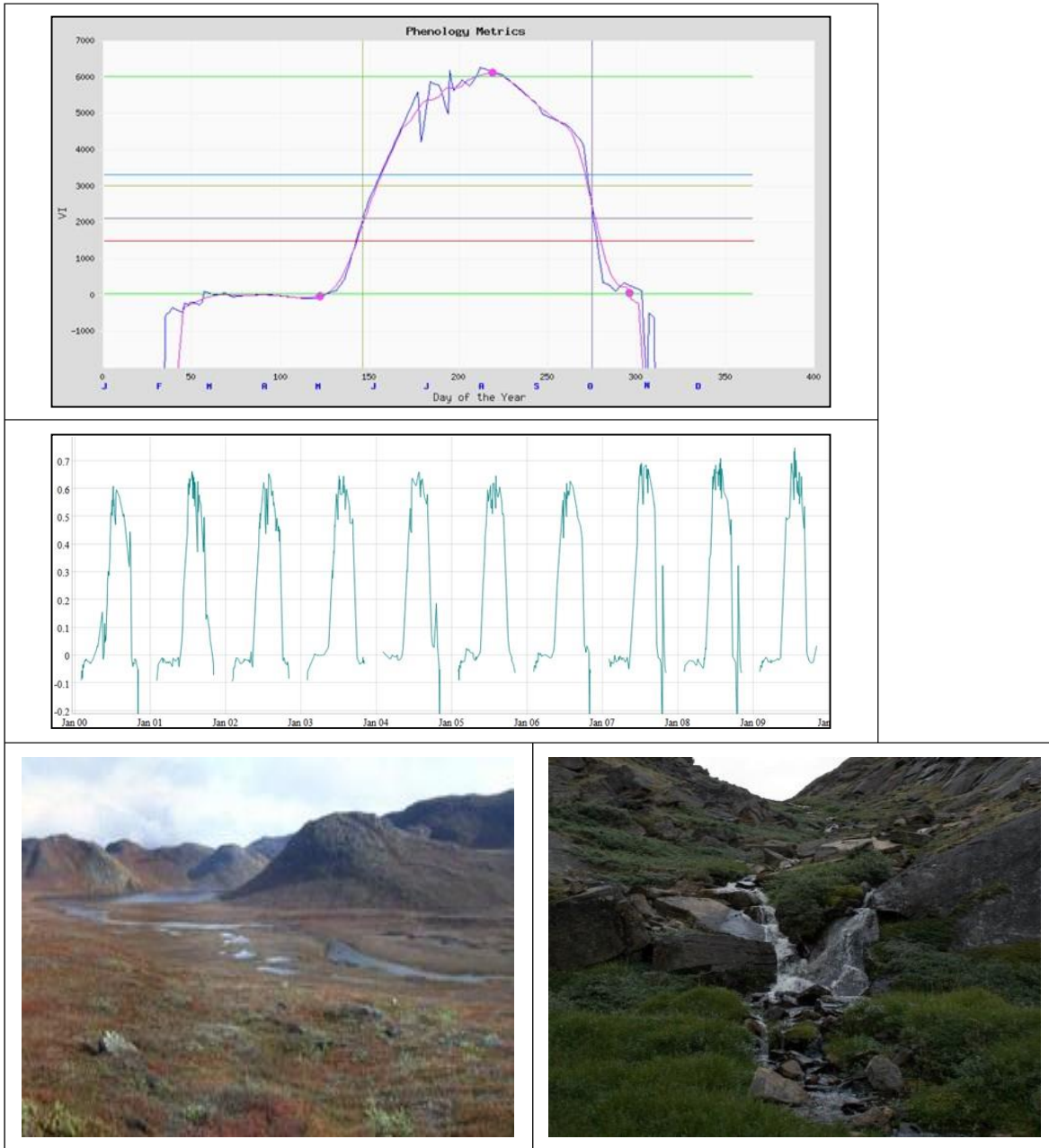
Figure 27. Tundra land cover with one growing season

6.1.5. Taiga/Boreal Forest with Single Fast Emerging Growing Season

Very high VI values at the peak of the curve, high rate of greening and senescence creates this characteristic curve representative of, homogeneous and dense taiga (needle leaf) vegetation.
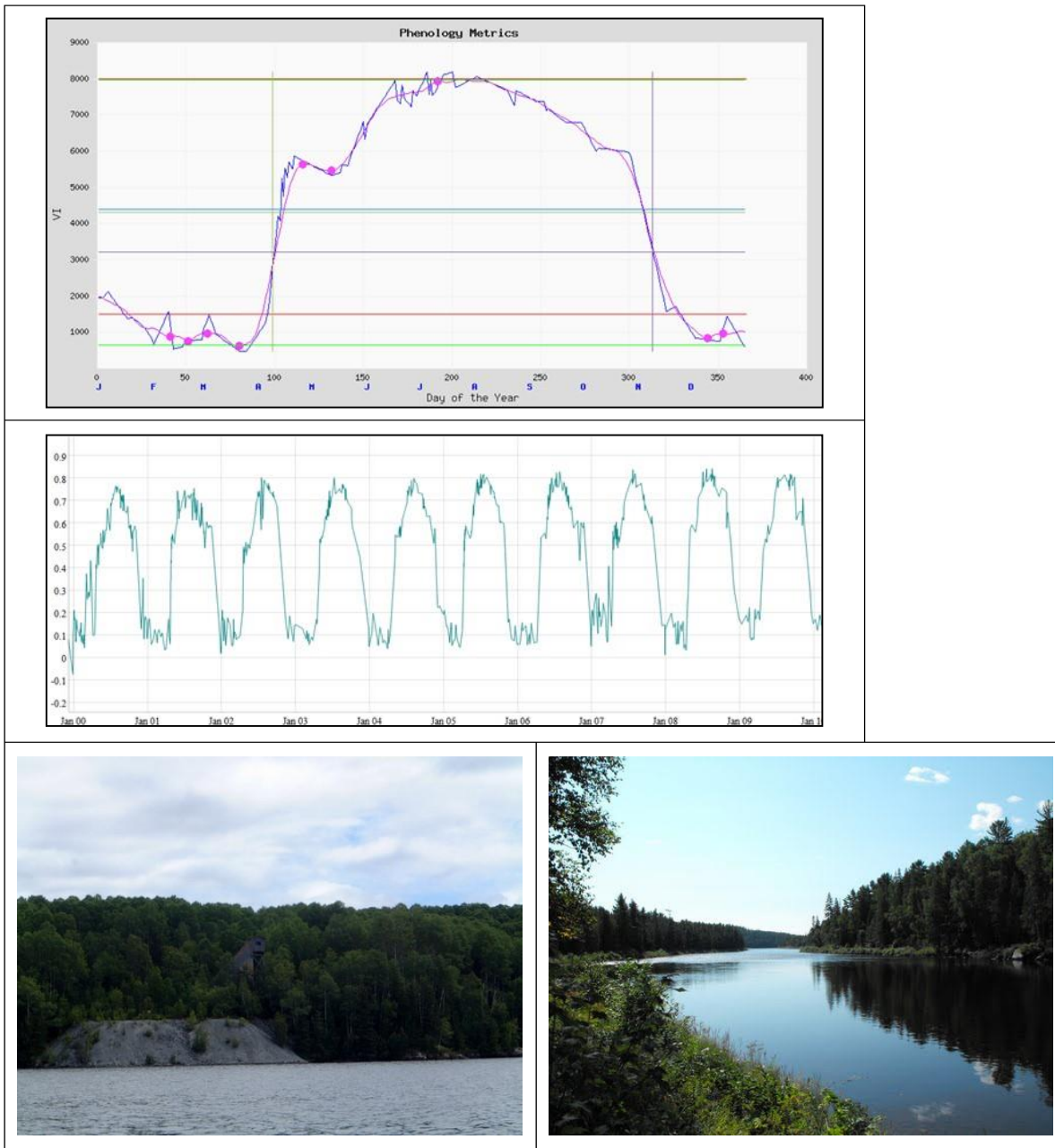
Figure 28. Taiga and boreal forests with one growing season

6.1.6.　Mosaic Landscape with Single Season and High Background Signal

The growing season is well defined by a clear and strong increase in activity. However, the VI signal remains quite high outside this well-defined start and end of the growing season. This usually results from a mixed and heterogeneous landscape with patches of intact forest and managed agriculture. As a result, the VI signal is a composite of the strong agriculture seasonality and the fairly stable forest VI signal (background). The algorithm will indicate the well-defined growing season but will also provide some information about the background vegetation signal.
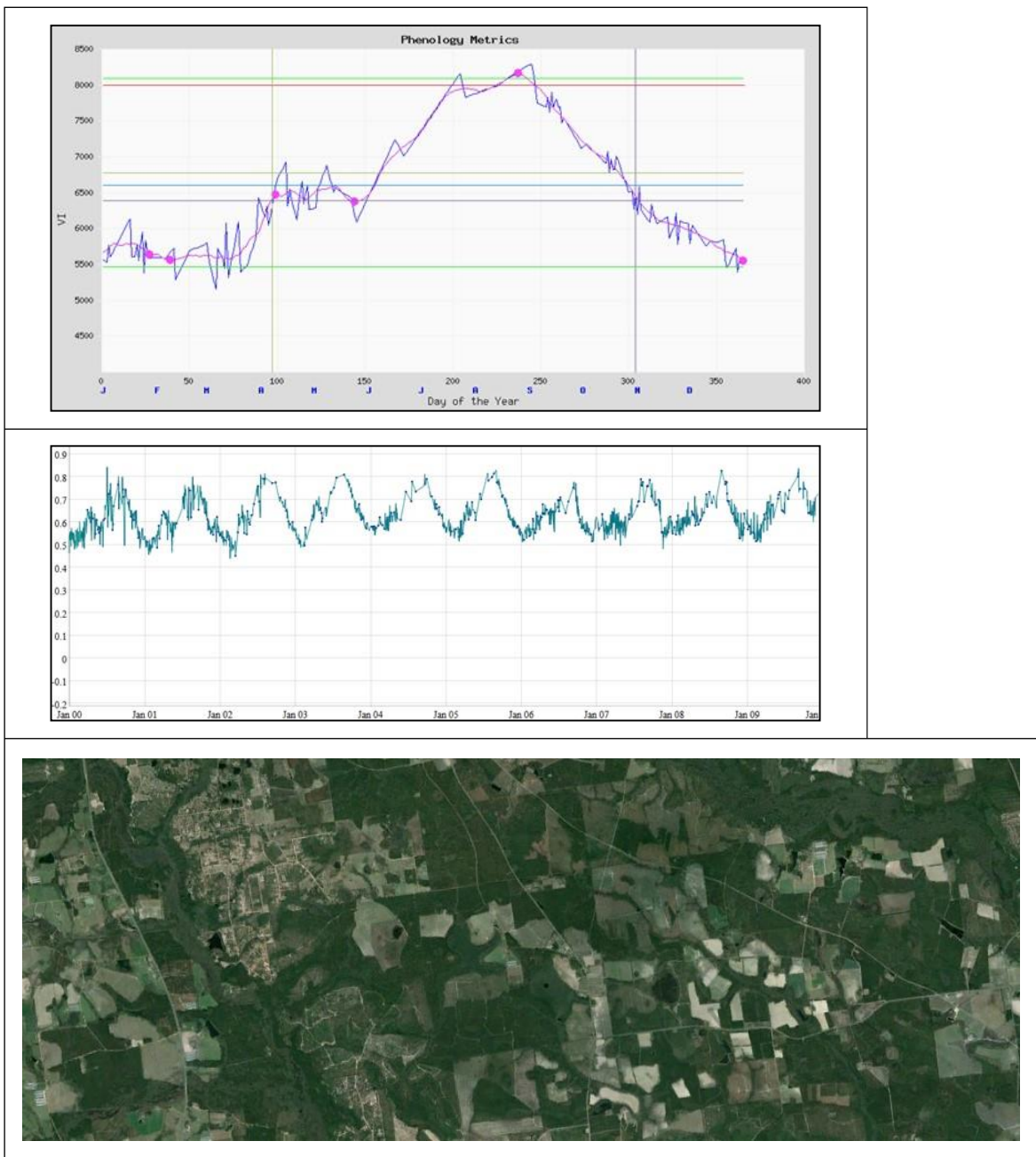
Figure 29. Single growing season with strong background signal.

6.1.7.   Tropical Forest with Single Yearlong Growing Season

The tropics are characterized by a very high and flat VI time series profile yearlong. Low VI values are usually the result of noise, unscreened and sub-pixel clouds, or strong aerosols that artificially lower the VI signal.
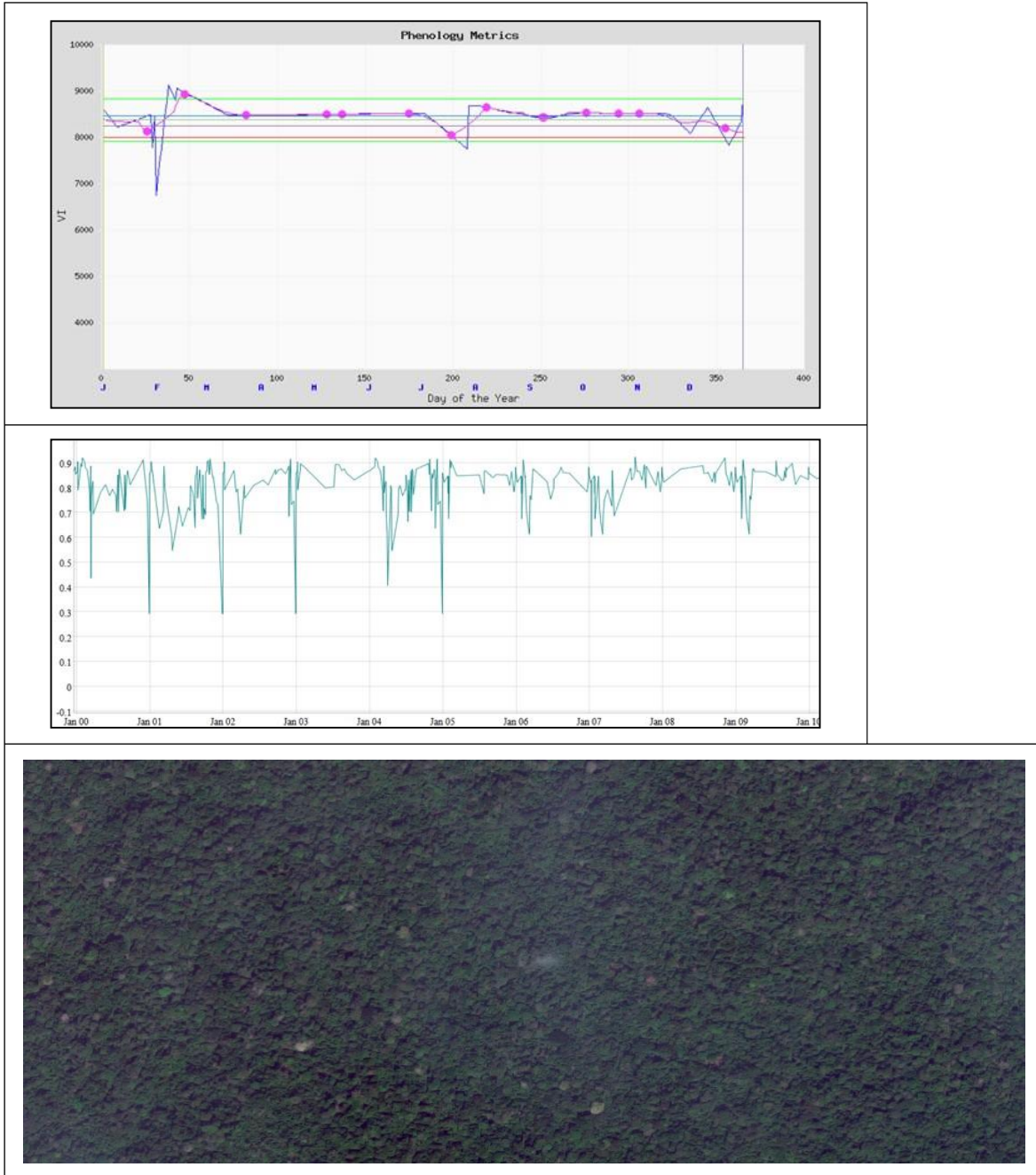
Figure 30. Single season with no breaks

6.1.8.   Landscapes with Multiple Growing Seasons

6.1.8.1. Two growing Seasons

While natural vegetation with two separate and strong rainy seasons may exhibit bimodal VI profiles (Ethiopian highlands, Himalayan foothills, etc...), generally intensely cultivated areas are characterized by multiple seasons. The depression in the curve may well be due to the difference in timing of the various crops and crops rotation of the different parcels, and not necessarily the end of one season and the beginning of another.
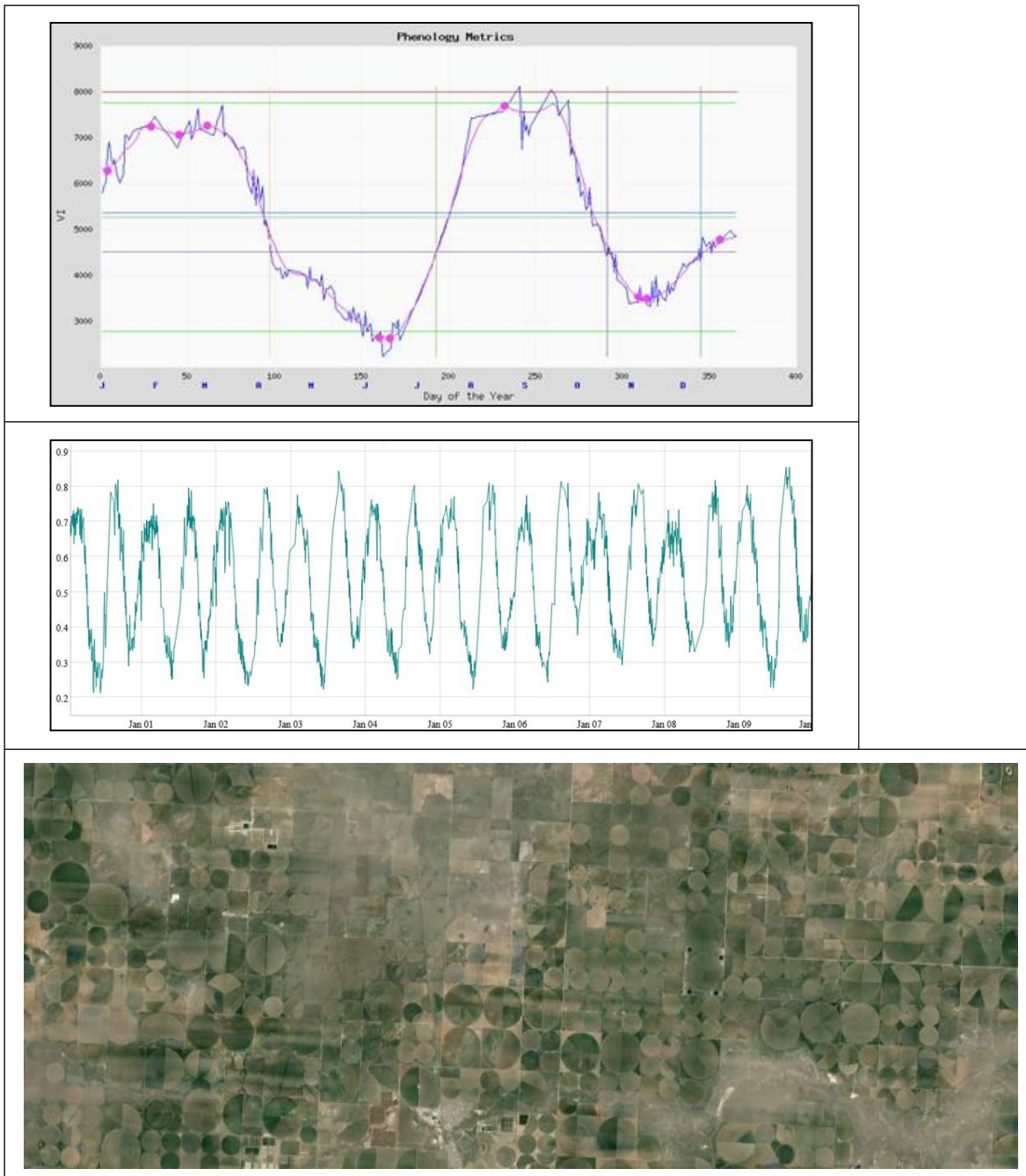
Figure 31. Two growing seasons

6.1.8.2. Three growing Seasons

Regions with intensive agriculture activities may show up to three different growing seasons (even more in some limited areas). These areas are mostly found in parts of China (North East). This multiple season characteristic may also be the result of mixing of active crops and bare soils due to the parcels size and different crops rotation in each season. This mixing results in a composite signal from crops at different phenological stages and bare soil.
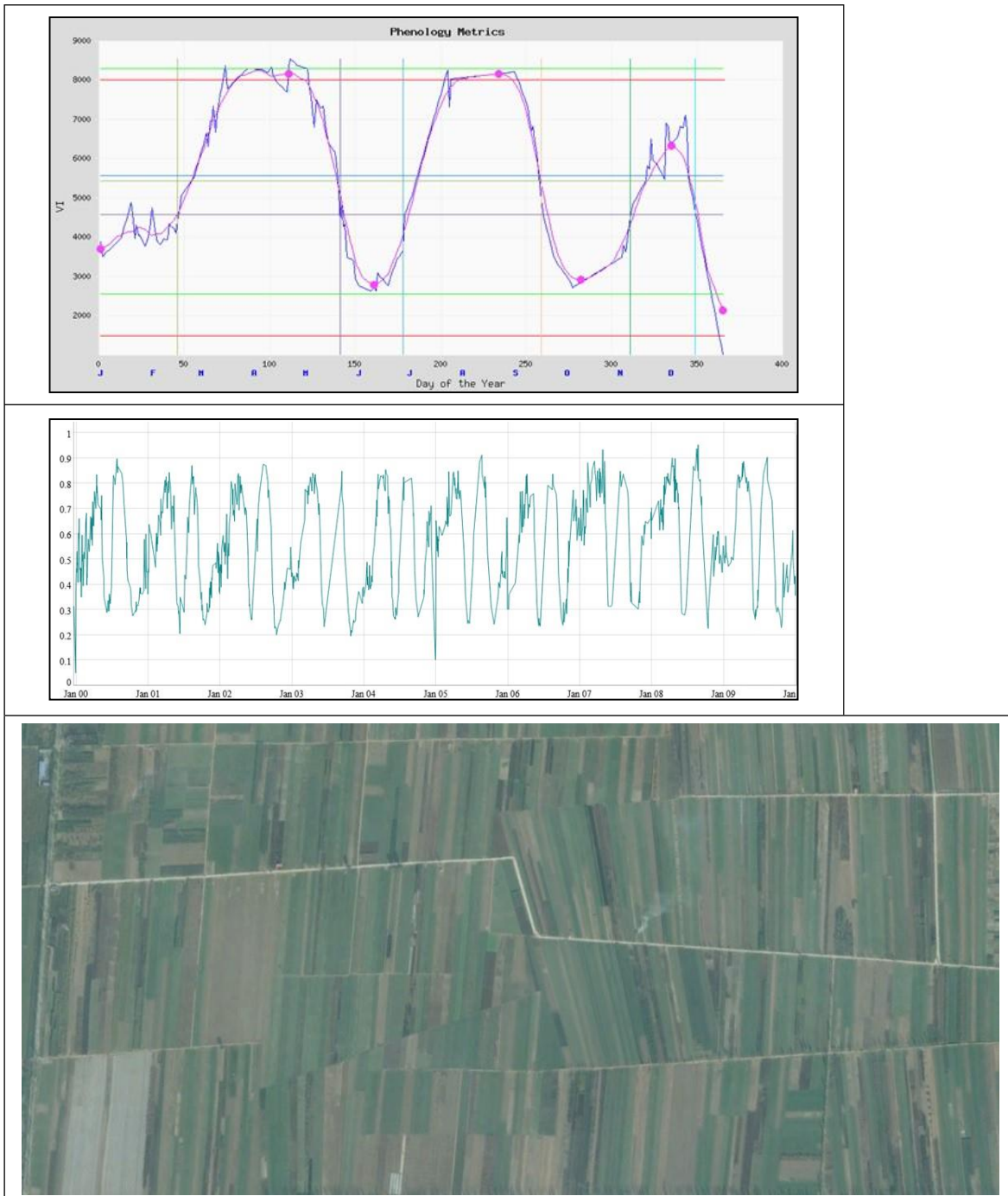
Figure 32. Three growing seasons

6.2. Algorithm Parameterization

Given these VI time series profiles variations a series of thresholds and rules were defined to aid the phenology algorithm identify the proper VI profile and correct phenology model. And while the algorithm parameterization helps with the proper characterization of the growing season on a global base, local and regional conditions may still be different or complex leading to errors. The algorithm uses the following set of parameters and rules (Fig. 33):

- The daily data were smoothed using a 14-day moving window to remove persistently poor quality data and residual and sub-pixel noise.
- Phenology is not considered below an NDVI < 0.12 or EVI2 < 0.08
- A Minimum seasonal $\Delta VI = 0.05$ (NDVI) and 0.03 (EVI2) is required, otherwise a growing season cannot take hold (the 0.05/0.03 VI thresholds represent the level of error/noise in the data and any change need to be above this noise level to be considered).
- The VI change must be consistent and sustained for a minimum sequence of days or it is considered only noise (45 days of consistent trend)
- To further mitigate noise in the data the Minimum and Maximum VI values of the annual profiles are estimated using the top and bottom 5% percentiles. This aggregates the data to insure a consistent Min/Max values and avoids outliers.

Figure 33. Phenology algorithm parametrization

6.3. Algorithm Performance

The phenology algorithm was applied globally and resulted in a consistent 33-year long Phenology ESDR. A spatial and quantitative analysis and evaluation of the phenology metrics revealed some interesting behavior and issues that may have resulted from the residual noise in the input time series, natural phenomenon, or climate change drivers. Below is a description of some of these observations that may aid in understanding the algorithm global performance and ESDR.
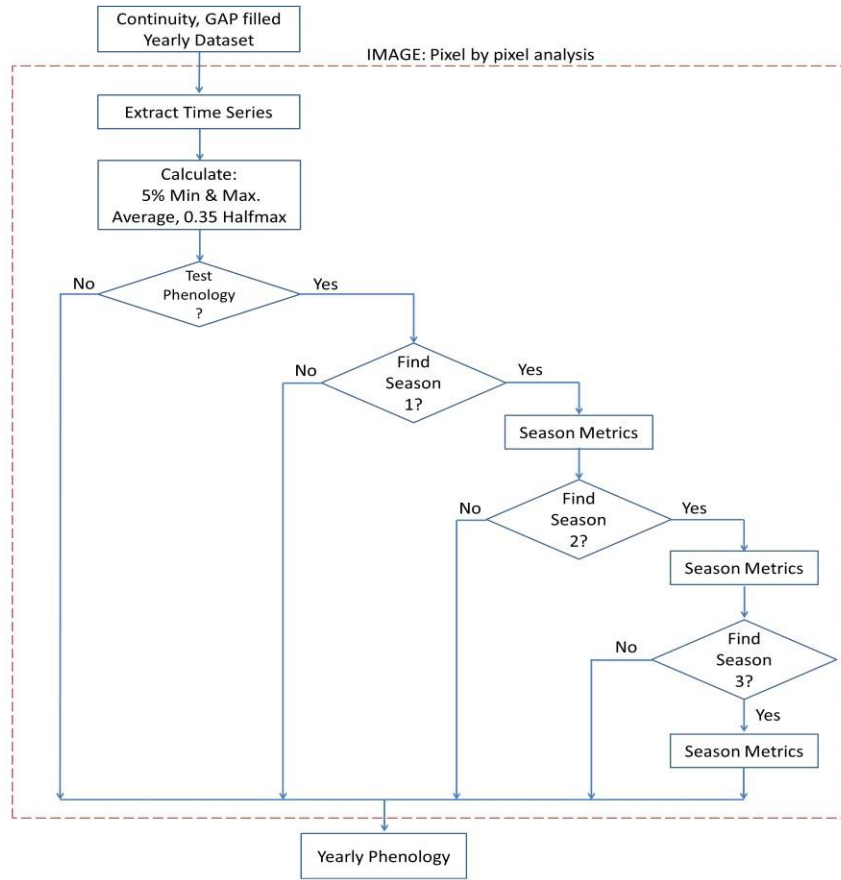
Figure 34. Phenology algorithm flow diagram

6.3.1. Forested Areas with Seasonal Snow/Ice Cover

When the snow/ice season begins to retreat trees and the ground are usually still covered with snow which is noticeable by the low VI signal. Depending on how sustained the start of the growing season, trees sometimes shed completely or partially their snow revealing their leaves to the sensor. This initial break in snow cover only concerns the trees since the snow cover on the ground takes longer to melt away resulting in the start of the gradual increase in the VI signal. This gradual increase may relapse and disappear due to a late snow storm that once again can cover the canopy and ground. This temporal increase, while having all the characteristics of a growing season start is usually ignored by the algorithm due to the VI signal dropping close to where it was during the dormancy period due to a late season snow storm. This may look like a small phenological curve (well defined greening, peak, senescence), however it can neither be considered a standalone

growing season nor a start of the regular season. The landscape vegetation dynamic is typical of Eurasia and North America (Canada/Alaska). The algorithm handles this artifact by requiring the growing season be longer than a preset threshold value (Fig. 35).
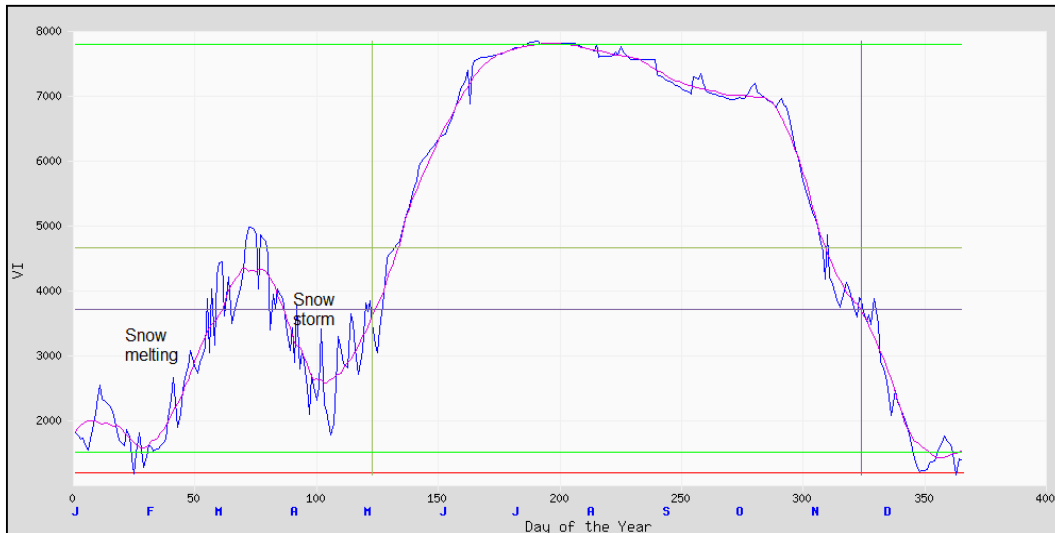
Figure 35. Late snow storms impact the time series profiles. The small growing season is in fact not real and results from a combination of the landscape snow dynamic and canopy structure.

6.3.2. Tropical Evergreen Forests with Persistent Clouds

The persistence of clouds and aerosols make it sometimes hard to identify cloud free and good quality data to construct a reasonable time series profile. These long term cloudy observations are typical of regions like Guyana, Suriname, French Guiana, parts of Colombia, Ecuador, Peru in South America and the regions of Sierra Leone, Liberia, Ghana, Nigeria, Cameroon, Gabon in West Africa, and the Indonesian Islands. These areas are characterized by annual profiles that have a consistent depression similar to senescence but resulting from the presence of sub-pixel clouds during the long cloudy season.
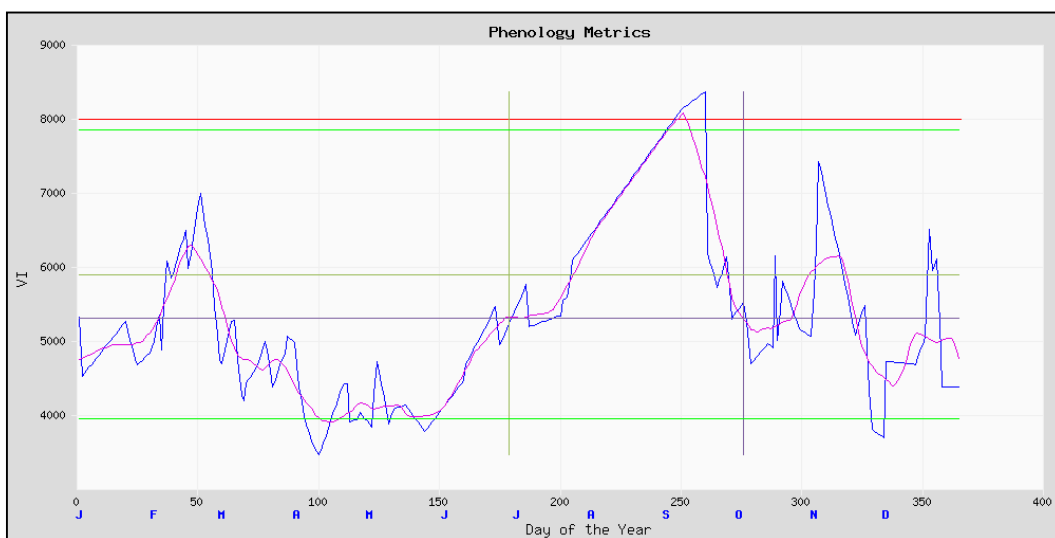
Figure 36. NDVI time series from Guyana. The profile shows a false growing season.

The algorithm will indicate and extract growing season metrics from these profiles that most

56

likely are not a true reflection of vegetation activities. The algorithm reports up to 3 seasons for some of the pixels in these areas, however this is purely the result of noise in the time series. And while the pixel Reliability/Rank will inform the users that the input data is quite poor it is still up to the user to filter out these observations or use consider these limitations.

6.3.3.   Agriculture Areas

Agriculture areas with parcels smaller than the CMG pixel size (5.6 x 5.6 km) will result in mixed spectral signatures from this diverse composition. These areas are usually a mix of crops, bare soil, different growing season stages, or different crops altogether, consequently, the pixel temporal profile may not reflect an accurate depiction of the actual vegetation dynamic or phenology of a particular crop. The peaks and valleys are an average signal of the dominant landscape cover (Zhang et al., 2009) at the time and hence the metrics may be inaccurate if not viewed in that context. Nevertheless, even with these small plots regional practices and cropping patterns are quite homogenous for wide areas and the impact should be minimal.
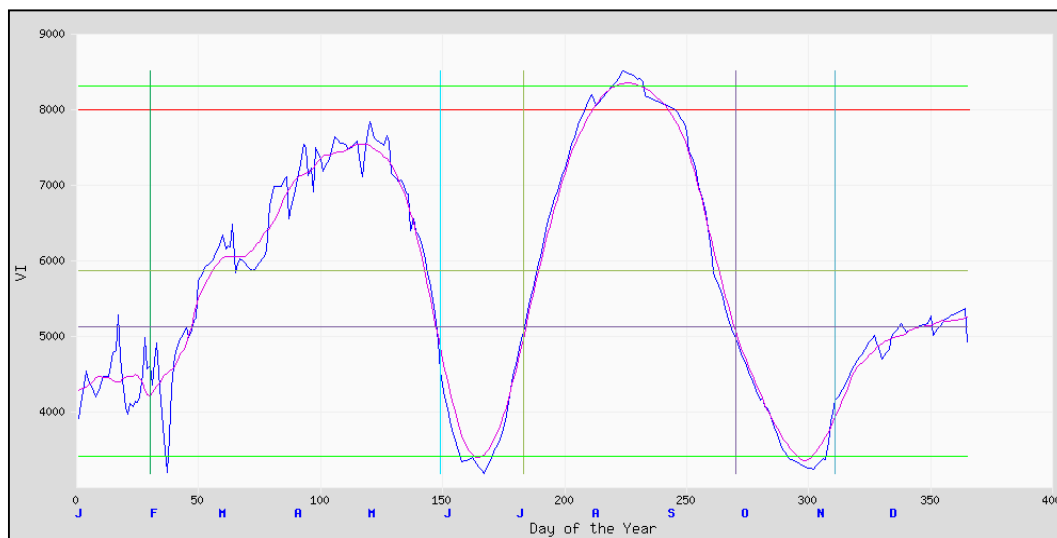
Figure 37. Agricultural areas with 3 distinct growing seasons (Henan, China). While, this is quantitatively accurate from a landscape perspective, conditions on the ground may not be that distinct.

Although, the phenology algorithm could theoretically combine these partial seasons into a single long growing season we determined that reporting them in separate seasons with start and end almost side by side is more accurate. The phenology algorithm looks at the landscape and dominant signal and is not specie or land cover type dependent.

6.3.4.   Background Season

Mixed growing season profiles are also observed in areas where agricultural fields are mixed and surrounded by intact forests. The VI profile is a composite signal from the crops and forest. Once the crop is harvested the forest remains green maintaining a high VI signal. Meanwhile, the signatures cannot be separated/unmixed, and the effect of the crop strong growing season is interpreted as the only change in the profile. This obviously disguises any change in the forest phenology. For these reasons the algorithm was adjusted to estimate and report a background season/signal and help identify these landscapes.

Figure 38. Agriculture fields surrounded by forest (Georgia, USA)

6.3.5. Deserts and Arid Lands with Occasional Precipitation

In desert areas where normally vegetation is extremely sparse and hardly active from a sensor view perspective (signal too small to be recorded), the occurrence of rainfall will sometimes cause an immediate and noticeable vegetation response with emergence of grasses and sustained greening that is recorded by the sensor. This increase is usually short and sporadic and not necessarily annual. The phenology algorithm is designed to detect and estimate the growing season properly over these landscapes with some years having no reported growing season as a result of the lack of enough rain or when the VI signal is too small to pass the algorithm required change rate thresholds. Moreover, if the maximum value is below the global minimum threshold set for the algorithm, no phenology will be estimated even though some vegetation activity and greening may have taken place. These landscape and profiles are typical of the Australian, Saudi Arabian, and other deserts.
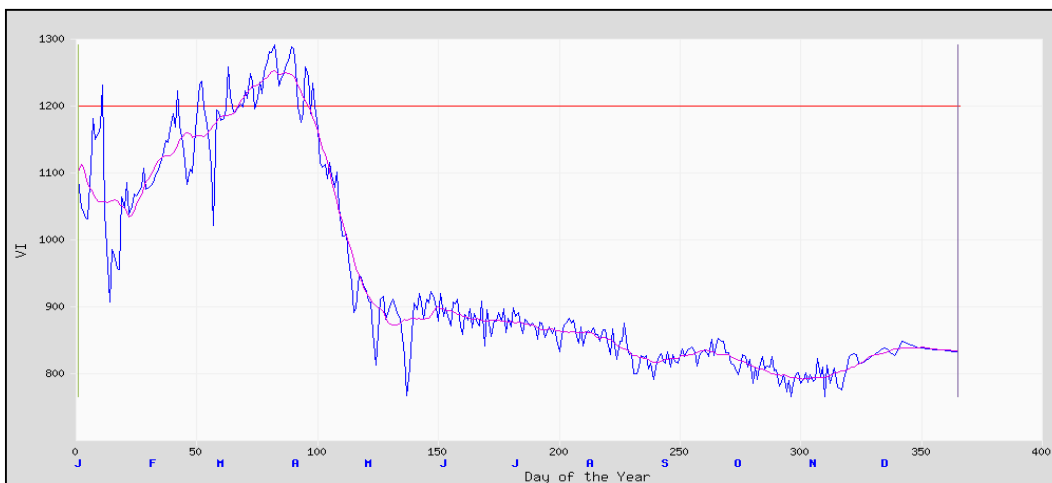
Figure 39. Precipitation effect on deserts and arid lands

6.4. Scientific Data Sets

Each annual 0.05-deg VIPPHEN product contains 26 SDSs, listed in Table 8. There are two separate products, an NDVI based and an EVI2 based phenology product.

58

Table 8. VIP phenology SDS list

Science data set	Units	Data Type	Valid Range	Fill*	Scale Factor
Start of Season 1	Days	INT16	1-366	-2	1
End of Season 1	Days	INT16	1-366	-2	1
Length of Season 1	Days	INT16	1-366	-2	1
Day of Peak Season 1	Days	INT16	1-366	-2	1
Rate of Greening Season 1	VI/Day	INT16	> 0	-15,000	100
Rate of Senescence Season 1	VI/Day	INT16	> 0	-15,000	100
Max VI Season 1	VI	INT16	-10,000 − 10,000	-15,000	10,000
Start of Season 2	Days	INT16	1-366	-2	1
End of Season 2	Days	INT16	1-366	-2	1
Length of Season 2	Days	INT16	1-366	-2	1
Day of Peak Season 2	Days	INT16	1-366	-2	1
Rate of Greening Season 2	VI/Day	INT16	> 0	-15,000	100
Rate of Senescence Season 2	VI/Day	INT16	> 0	-15,000	100
Max VI Season 2	VI	INT16	-10,000 – 10,000	-15,000	10,000
Start of Season 3	Days	INT16	1-366	-2	1
End of Season 3	Days	INT16	1-366	-2	1
Length of Season 3	Days	INT16	1-366	-2	1
Day of Peak Season 3	Days	INT16	1-366	-2	1
Rate of Greening Season 3	VI/Day	INT16	> 0	-15,000	100
Rate of Senescence Season 3	VI/Day	INT16	> 0	-15,000	100
Max VI Season 3	VI	INT16	-10,000 - 10,000	-15,000	10,000
Cumulative VI	VI	INT16	0 – 10,000	-15,000	1,000
Average VI	VI	INT16	-10,000 – 10,000	-15,000	10,000
Background VI	VI	INT16	-10,000 – 10,000	-15,000	10,000
Number of Seasons	N/A	INT8	1-3	-2	1
Reliability	N/A	INT8	-1 6	-1	1

*: While we are listing a single value for fill the products may contain multiple Fill_Values to separate between the locations of the pixels and why they are missing.

For all Date related SDS:

- Fill Value = 1 over land,
- Fill_Value = -2 over water

For Greening rates and VI values

- Fill Value =-13,000 over land,
- Fill_Value = -15,000 over water

For number of seasons:

- Fill Value = 1 over land,
- Fill_Value = -2 over water

For Reliability:

- Fill Value = 1 over land,
- Fill_Value = -2 over water

6.5. Product Specific Metadata

A list of the metadata fields used for QA evaluation of the Phenology ESDR files is below:
- QAPERCENTEXCELLENT
- QAPERCENTGOOD
- QAPERCENTACCEPTABLE
- QAPERCENTMARGINAL
- QAPERCENTPOOR
- QAPERCENTUNRELIABLE
- QAPERCENTLONGSEASON
- QAPERCENTNODATA
- RANGEDATETIME
- RANGEBEGINNINGDATE
- RANGEBEGINNINGTIME
- RANGEENDINGDATE
- RANGEENDINGTIME

6.6. Global and Local Metadata Attributes

The global metadata is written to the output file during the product generation process and could be used during search/order from the archives. A listing of the most relevant metadata is provided below (slightly abridged and only the first and last input filenames are listed):

```
Global attributes: 4
HDFEOSVersion: HDFEOS_V2.19GROUP=SwathStructure
END_GROUP=SwathStructure
GROUP=GridStructure
    GROUP=GRID_1
        GridName="VIP_CMG_GRID"
        XDim=7200
        YDim=3600
        UpperLeftPointMtrs=(-180000000.000000,90000000.000000)
        LowerRightMtrs=(180000000.000000,-90000000.000000)
        Projection=GCTP_GEO
        GROUP=Dimension
        END_GROUP=Dimension
        GROUP=DataField
            OBJECT=DataField_1
                DataFieldName="Start of Season 1"
                DataType=DFNT_INT16
                DimList=("YDim","XDim")
            END_OBJECT=DataField_1
            OBJECT=DataField_2
                DataFieldName="End of Season 1"
                DataType=DFNT_INT16
                DimList=("YDim","XDim")
            END_OBJECT=DataField_2
            OBJECT=DataField_3
                DataFieldName="Length of Season 1"
                DataType=DFNT_INT16
                DimList=("YDim","XDim")
            END_OBJECT=DataField_3
```

```
OBJECT=DataField_4
    DataFieldName="Day of Peak Season 1"
    DataType=DFNT_INT16
    DimList=("YDim","XDim")
END_OBJECT=DataField_4
OBJECT=DataField_5
    DataFieldName="Rate of Greening Season 1"
    DataType=DFNT_INT16
    DimList=("YDim","XDim")
END_OBJECT=DataField_5
OBJECT=DataField_6
    DataFieldName="Rate of Senescence Season 1"
    DataType=DFNT_INT16
    DimList=("YDim","XDim")
END_OBJECT=DataField_6
OBJECT=DataField_7
    DataFieldName="Max VI Season 1"
    DataType=DFNT_INT16
    DimList=("YDim","XDim")
END_OBJECT=DataField_7
OBJECT=DataField_8
    DataFieldName="Start of Season 2"
    DataType=DFNT_INT16
    DimList=("YDim","XDim")
END_OBJECT=DataField_8
OBJECT=DataField_9
    DataFieldName="End of Season 2"
    DataType=DFNT_INT16
    DimList=("YDim","XDim")
END_OBJECT=DataField_9
OBJECT=DataField_10
    DataFieldName="Length of Season 2"
    DataType=DFNT_INT16
    DimList=("YDim","XDim")
END_OBJECT=DataField_10
OBJECT=DataField_11
    DataFieldName="Day of Peak Season 2"
    DataType=DFNT_INT16
    DimList=("YDim","XDim")
END_OBJECT=DataField_11
OBJECT=DataField_12
    DataFieldName="Rate of Greening Season 2"
    DataType=DFNT_INT16
    DimList=("YDim","XDim")
END_OBJECT=DataField_12
OBJECT=DataField_13
    DataFieldName="Rate of Senescence Season 2"
    DataType=DFNT_INT16
    DimList=("YDim","XDim")
END_OBJECT=DataField_13
OBJECT=DataField_14
    DataFieldName="Max VI Season 2"
    DataType=DFNT_INT16
```

```
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_14
                OBJECT=DataField_15
                    DataFieldName="Start of Season 3"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_15
                OBJECT=DataField_16
                    DataFieldName="End of Season 3"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_16
                OBJECT=DataField_17
                    DataFieldName="Length of Season 3"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_17
                OBJECT=DataField_18
                    DataFieldName="Day of Peak Season 3"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_18
                OBJECT=DataField_19
                    DataFieldName="Rate of Greening Season 3"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_19
                OBJECT=DataField_20
                    DataFieldName="Rate of Senescence Season 3"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_20
                OBJECT=DataField_21
                    DataFieldName="Max VI Season 3"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_21
                OBJECT=DataField_22
                    DataFieldName="Cumulative VI"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_22
                OBJECT=DataField_23
                    DataFieldName="Average VI"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_23
                OBJECT=DataField_24
                    DataFieldName="Background VI"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_24
                OBJECT=DataField_25
```

```
                    DataFieldName="Number of Seasons"
                    DataType=DFNT_INT16
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_25
                OBJECT=DataField_26
                    DataFieldName="Reliability"
                    DataType=DFNT_INT8
                    DimList=("YDim","XDim")
                END_OBJECT=DataField_26
            END_GROUP=DataField
            GROUP=MergedFields
            END_GROUP=MergedFields
        END_GROUP=GRID_1
END_GROUP=GridStructure
GROUP=PointStructure
END_GROUP=PointStructure
END
GROUP = INVENTORYMETADATA
 GROUPTYPE = MASTERGROUP
 GROUP  = ECSDATAGRANULE
 OBJECT = LOCALGRANULEID
   NUM_VAL = 1
   VALUE = "VIPPHEN_NDVI.A2013.004.2016075191527.hdf"
  END_OBJECT = LOCALGRANULEID
  OBJECT = PRODUCTIONDATETIME
   NUM_VAL = 1
   VALUE = "2016-03-15T19:15:27.000Z"
  END_OBJECT = PRODUCTIONDATETIME
  OBJECT = REPROCESSINGACTUAL
  NUM_VAL = 1
   VALUE = "reprocessed"
  END_OBJECT = REPROCESSINGACTUAL
  OBJECT = LOCALVERSIONID
   NUM_VAL = 1
   VALUE = "4.1"
  END_OBJECT = LOCALVERSIONID
 END_GROUP = ECSDATAGRANULE
 GROUP = COLLECTIONDESCRIPTIONCLASS
 OBJECT = VERSIONID
   NUM_VAL = 1
   VALUE = 4
  END_OBJECT = VERSIONID
  OBJECT = SHORTNAME
  NUM_VAL = 1
   VALUE = "VIPPHEN"
  END_OBJECT = SHORTNAME
 END_GROUP = COLLECTIONDESCRIPTIONCLASS
 GROUP = INPUTGRANULE
  OBJECT = INPUTPOINTER
   NUM_VAL = "1096"
   VALUE = ("VIP01.A2012001.004.2016074185016.hdf"..."VIP01.A2014365.004.2016075115204.hdf")
  END_OBJECT = INPUTPOINTER
 END_GROUP = INPUTGRANULE
```

```
GROUP = SPATIALDOMAINCONTAINER
  GROUP = HORIZONTALSPATIALDOMAINCONTAINER
    GROUP = BOUNDINGRECTANGLE
      OBJECT = EASTBOUNDINGCOORDINATE
        NUM_VAL = 1
        VALUE = 180.0
      END_OBJECT = EASTBOUNDINGCOORDINATE
      OBJECT = WESTBOUNDINGCOORDINATE
      NUM_VAL = 1
        VALUE = -180.0
      END_OBJECT = WESTBOUNDINGCOORDINATE
      OBJECT = SOUTHBOUNDINGCOORDINATE
      NUM_VAL = 1
        VALUE = -90.0
      END_OBJECT = SOUTHBOUNDINGCOORDINATE
      OBJECT = NORTHBOUNDINGCOORDINATE
      NUM_VAL = 1
        VALUE = 90.0
      END_OBJECT = NORTHBOUNDINGCOORDINATE
    END_GROUP = BOUNDINGRECTANGLE
  END_GROUP = HORIZONTALSPATIALDOMAINCONTAINER
  GROUP = GRANULELOCALITY
    OBJECT = LOCALITYVALUE
      NUM_VAL = 1
      VALUE = "Global"
    END_OBJECT = LOCALITYVALUE
  END_GROUP = GRANULELOCALITY
END_GROUP = SPATIALDOMAINCONTAINER
GROUP = RANGEDATETIME
  OBJECT = RANGEENDINGDATE
    NUM_VAL = 1
    VALUE = "2013-12-31"
  END_OBJECT = RANGEENDINGDATE
  OBJECT = RANGEENDINGTIME
  NUM_VAL = 1
    VALUE = "23:59:59"
  END_OBJECT = RANGEENDINGTIME
  OBJECT = RANGEBEGINNINGDATE
  NUM_VAL = 1
    VALUE = "2013-01-01"
  END_OBJECT = RANGEBEGINNINGDATE
  OBJECT = RANGEBEGINNINGTIME
  NUM_VAL = 1
    VALUE = "00:00:00"
  END_OBJECT = RANGEBEGINNINGTIME
END_GROUP = RANGEDATETIME
GROUP = PGEVERSIONCLASS
  OBJECT = PGEVERSION
    NUM_VAL = 1
    VALUE = "4.1"
  END_OBJECT = PGEVERSION
  END_GROUP
END_GROUP = PGEVERSIONCLASS
```

```
END_GROUP = INVENTORYMETADATA
END
GROUP = ARCHIVEDMETADATA
  GROUPTYPE = MASTERGROUP
  OBJECT = ALGORITHMPACKAGENAME
  NUM_VAL = 1
    VALUE = "VIPLAB_PHEN_V4"
  END_OBJECT = ALGORITHMPACKAGENAME
  OBJECT = ALGORITHMPACKAGEVERSION
  NUM_VAL = 1
    VALUE = "4.1"
  END_OBJECT = ALGORITHMPACKAGEVERSION
  OBJECT = PROCESSINGCENTER
    NUM_VAL = 1
    VALUE = "VIPLAB"
  END_OBJECT = PROCESSINGCENTER
  OBJECT = SEAPROCESSED
  NUM_VAL = 1
    VALUE = "No"
  END_OBJECT = SEAPROCESSED
  OBJECT = PROCESSINGENVIRONMENT
  NUM_VAL = 1
    VALUE = "Linux version 3.10.0-327.4.5.el7.x86_64 (Red Hat 4.8.5-4) Intel(R) Xeon(R) CPU X5690 @
3.47GHz"
  END_OBJECT = PROCESSINGENVIRONMENT
  OBJECT = DESCRREVISION
    NUM_VAL = 1
    VALUE = "4.1"
  END_OBJECT = DESCRREVISION
  OBJECT = VIPSCIENCEQUALITYFLAGEXPLANATION
  NUM_VAL = 1
  VALUE = "See http://vip.arizona.edu/documentation/esdr/?VIPPHENv04"
  END_OBJECT = VIPSCIENCEQUALITYFLAGEXPLANATION
  OBJECT = DATACOLUMNS
    NUM_VAL = 1
    VALUE = 7200
  END_OBJECT = DATACOLUMNS
  OBJECT = DATAROWS
  NUM_VAL = 1
    VALUE = 3600
  END_OBJECT = DATAROWS
  OBJECT = GLOBALGRIDCOLUMNS
    NUM_VAL = 1
    VALUE = 7200
  END_OBJECT = GLOBALGRIDCOLUMNS
  OBJECT = GLOBALGRIDROWS
  NUM_VAL = 1
    VALUE = 3600
  END_OBJECT = GLOBALGRIDROWS
  OBJECT = NUMBEROFDAYS
  NUM_VAL = 1
    VALUE = "1096"
  END_OBJECT = NUMBEROFDAYS
```

```
 OBJECT = DAYSPROCESSED
   NUM_VAL = "1096"
   VALUE = ("2012001"..."2014365")
 END_OBJECT = DAYSPROCESSED
 GROUP = QASTATS
...........
END_GROUP = QASTATS
END_GROUP = ARCHIVEDMETADATA
END
```

6.7. Quality Assurance

While quality assurance bit fields similar to the VI product suite cannot be associated with the phenology products, a special pixel reliability/Rank measure similar to the VI product was created to aid with data filtering and quality assessment. The rank SDS is a derivative of the input and its quality on a pixel-by-pixel basis and thus is useful for data automated data analyses and applications. This rank information is summarized by quality classes and captures and describes the quality of the input data used to derive the phenology metrics. The rank provides the user with a simple measure of confidence in the derived phenology metric values.

Table 9. Phenology reliability index rank

| QA Rank	Description	QA Rank	Description
0	Excellent	5	Unreliable
1	Good	6	Long Season Stable
2	Acceptable	-1	Fill value over land
3	Marginal	-2	Fill value over water
4	Poor		
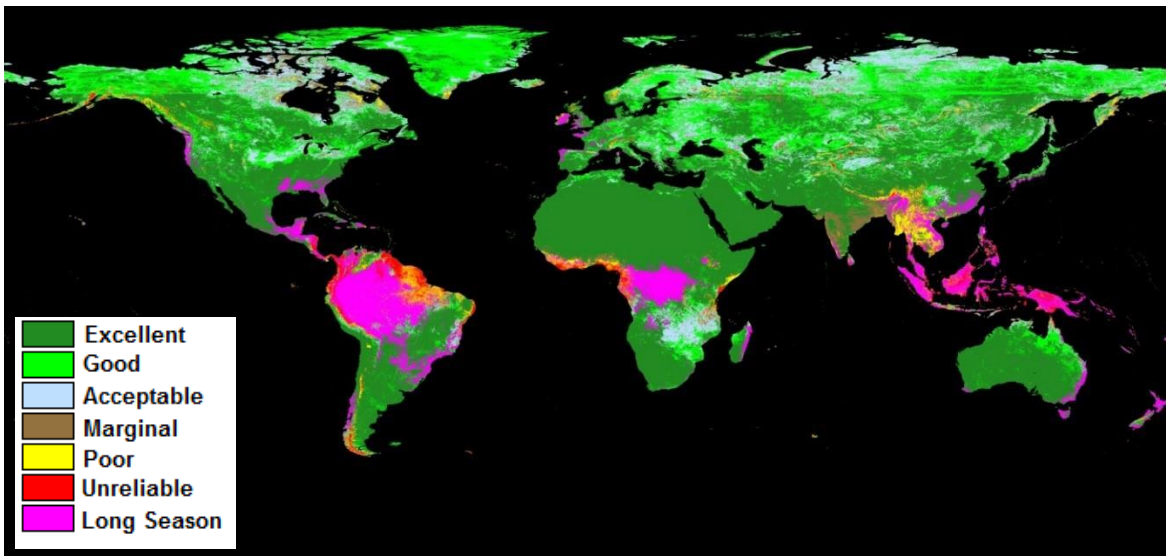
Figure 40. Phenology QA Rank/Pixel Reliability global distribution

7.   ESDRs Error and Uncertainty

All V4 VIP ESDRs are suitable for the study of land surface vegetation long term patterns, trends and anomalies and should support various ecosystem and climate modeling efforts. However, the error and uncertainty associated with these Vegetation Index and Phenology ESDRs are sometimes quite large, complex, and needs to be well characterized in order to promote accurate and proper use of these records. It is practically impossible to directly quantify the error in a vegetation index value due to the nature of the index itself and the underlying remote sensing data, and its spatial and temporal context, but we can still explore other quasi-quantitative methods that aim to simply characterizing these records and elicit their general error and uncertainty.

In support of these ESDRs, we have developed a simple framework (Fig. 41) and a set of metrics to characterize and capture the error and uncertainty in these data records.
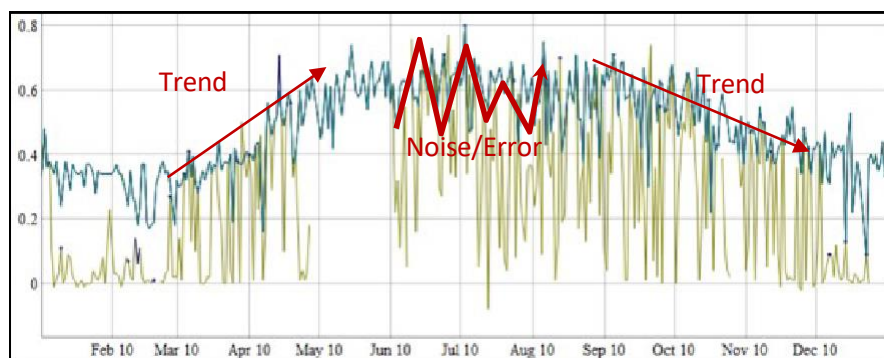
Figure 41. Error and Uncertainty model framework

This framework is spatially and temporally explicit, and is designed to capture key features of these data records:

- **Quality of the atmosphere correction and impact:** This is a key characteristic, and while current science algorithms are capable of addressing and correcting a host of atmosphere issues (water vapor, ozone, Rayleigh scattering, and light aerosols, viewing geometry) their performance is always an issue. In many situations, the correction actually exacerbates the problems due to ingesting poor quality ancillary data needed to drive the atmosphere correction algorithm.
- **Departure from the long term average:** This departure although can result from natural factors, is in many cases the result of noise and error in the data. The departure will be measured by the absolute distance from the long term average. To separate the noise from natural change, the error related departures are identified by examining their persistence. A sustained departure is most likely the result of a natural change/disturbance and not an error in the data.
- **Number of times the profile changes direction over short periods of time:** Although, these artifacts capture and reflect the natural cycle of vegetation dynamic, they can also result from noise and error in the underlying input data. Random and large oscillation about the long term average are most likely the result of error in the data and cannot result from natural and gradual vegetation dynamic.

Using this framework, the error is portioned into three categories (Fig. 42):
1. Error related to input (1): Could be estimated from accurate atmospherically corrected data over validation sites (ex: sunphotometers, Holben et al. 2006).
2. Departure from normal/expected profile (2): Based on long term standard deviations and

statical analyses of the records.

3. Temporal profile stability (3): Based on change about normal/mean. This indicates noise in data records and inhibits the profile characterization especially phenology metrics.
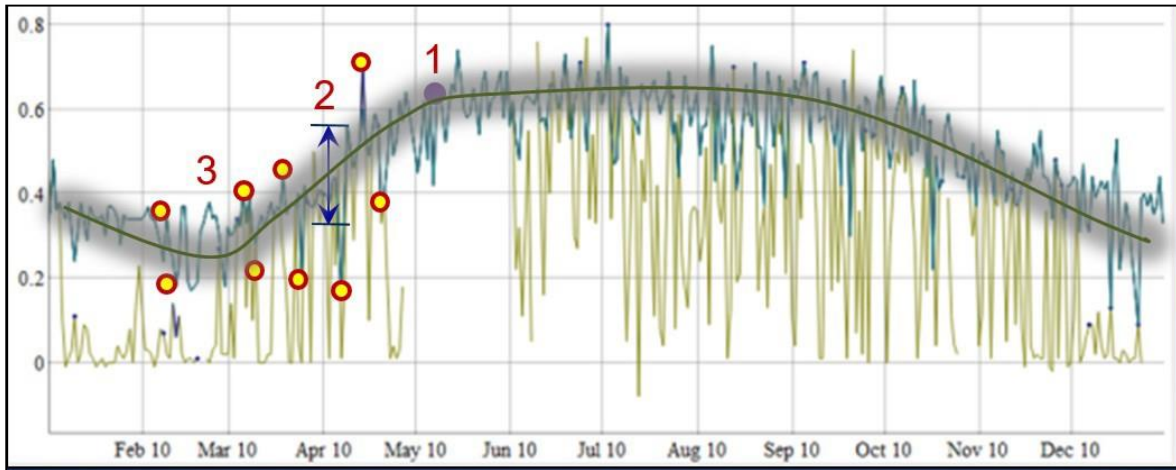

Figure 42. Error model

7.1.   LSR Input Related Error

Here we're not concerned with the VI formulation, but the error in the input to the VI equation and how it translates into a VI error envelope. We can measure how close we are to the Top of Canopy (TOC) reflectance (ability to remove all atmosphere contamination). To estimate this error, we used Surface Reflectance data from validation sites (ex: sunphotometers measurements over EOS Core sites, LPV). The LSR error is reported (EOS MODIS LPV website, http://landval.gsfc.nasa.gov/) to average about 2-5% in the Red & NIR for high quality data, (i.e. data with no residual clouds and no to minimum aerosol). Using an Ad-hoc approach we can model a maximum error, assuming the presence of residual aerosols and clouds, of about ±10% in the Red/NIR reflectance. The LSR error is simply transferred to the VI using the VI equations (Fig. 43).
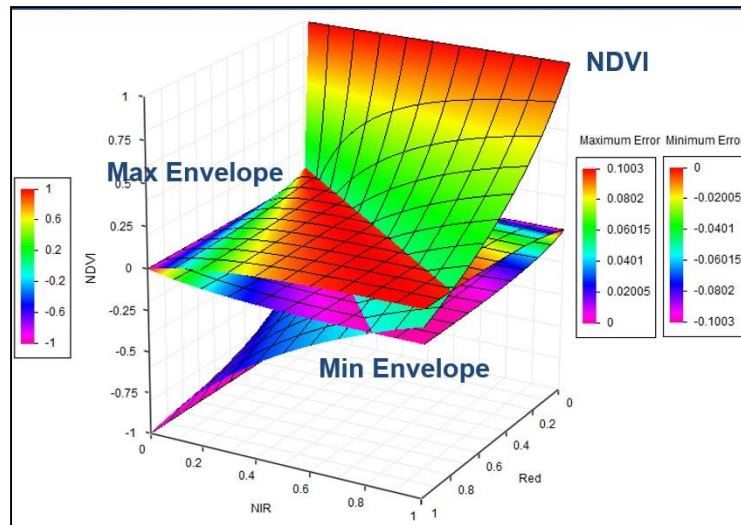
Figure 43. VI Error model resulting from input

Using this simple transfer model, the resulting maximum VI error is:
- For vegetated areas ~= 0.04-0.05 VI Units (~1-5-% relative)

- For sparsely and non-vegetated areas ~= 0.11 VI Units (~100% relative)

7.2.   Spatial Error Modeling

To model the input error spatially and characterize its impact on the resulting VI data records we can translate the VI error values from the model in Fig. 43 to a long term annual average map of Red and NIR. Each pair of (Red, NIR) will correspond to VI error envelope and a spatial error distribution map is easily constructed (Figs. 44 &45).
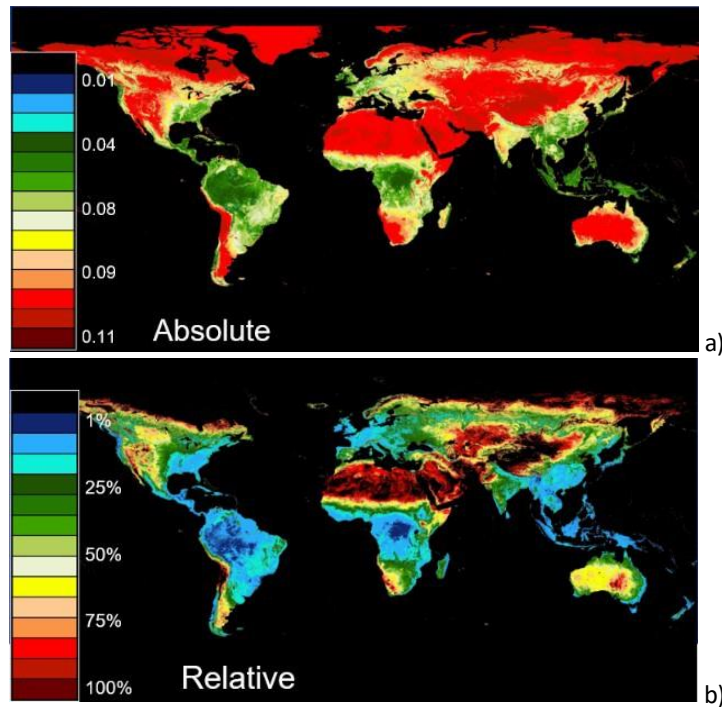
Figure 44. Spatial map of the absolute and Relative NDVI Error using the model from Fig. 43.
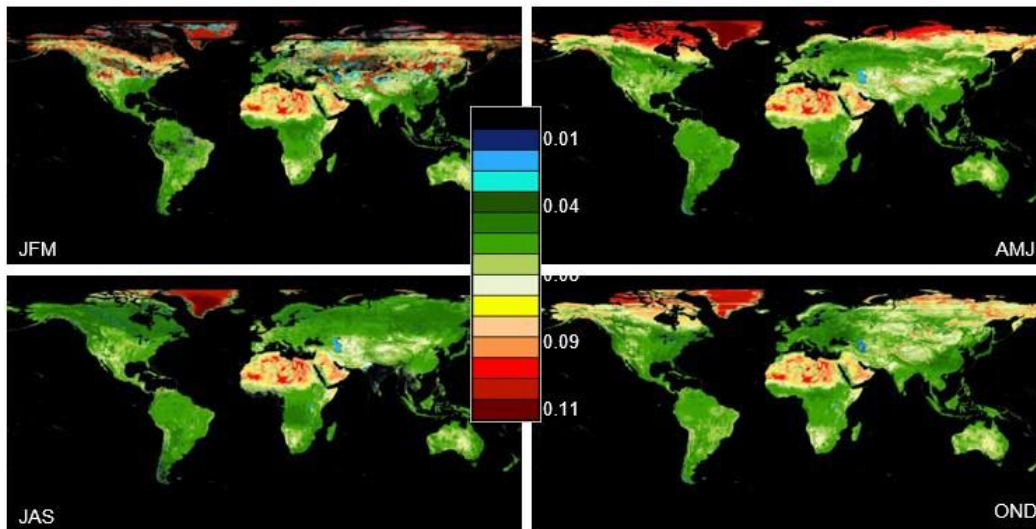
Figure 45. EVI2 Seasonal Error spatial distribution

7.3. Standard Deviation of the VI ESDRs

In statistics the standard deviation depicts the general departure from normal, or how spread a data about its average. In the context of these data records the standard deviation captures the impact of both natural processes and noise related error in the time series. This error is estimated by the statistical analysis of the long term standard deviation of VI profiles and is readily spatially explicit (Fig. 46).

Figure 46. VI Standard deviation (Error) or departure from long term mean. (a) During the winter and due to snow related noise this error is highest. (b) EVI2 standard deviation showing a smaller error ranges.

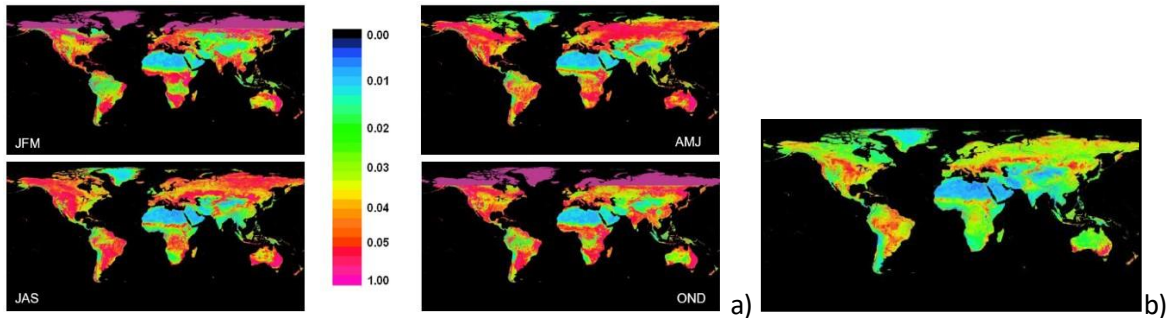Using this statistical approach, the average general error was:

- Error = 0.05 (VI Units) for vegetated and 0.005 (VI Units) for sparse or non-vegetated
- Largest error observed over vegetated areas and during spring/summer (peak growing season)

With this framework we computed spatially explicit annual metrics for each pixel. The resulting global maps elucidate the spatial coherency and error in these ESDRs. The results can aid end users assess these records and associate a spatial and seasonal per-pixel estimate of error and uncertainty. Post analysis results using these records can be constrained and their significance established. When transferring these errors into the phenology algorithm date related metrics the (95% percentile) error was estimated to around 5-30 days with a ±15-day average.

Overall the error due to the surface reflectance input uncertainty was rather small with an absolute max error = ~0.05, except for sparsely vegetated areas where it becomes relatively high up to 100% due to the signal (VI<=0.1) being small and overwhelmed by noise. The impact on EVI2 was larger than that of NDVI. The error was also found to be largest during the winter due to the presence of snow/ice which introduces more noise.

8.   Summary and Conclusions

Continuous acquisition of global satellite imagery over the years has contributed to the creation of long term data records from AVHRR, MODIS, TM, SPOT-VGT and other sensors. These records now account for more than 40 years of synoptic Earth surface observation. As these archives grow they become an invaluable tool for environmental monitoring, resources management, and climate studies from local, to regional, to global scales.

The Vegetation Index and Phenology Lab. (vip.arizona.edu) developed a series of state of the art science and processing algorithms to generate a global multi-sensor Earth Science Data Record of NDVI, EVI2, and Phenology metrics. Data from AVHRR, MODIS, and SPOT-VGT, covering the period 1981 to 2014, were processed into a seamless and sensor independent data records using a suite of science algorithms for data filtering, translation and continuity, Vegetation Index (NDVI and EVI2), land surface Phenology, and spatial and temporal gap filling. With V4 reprocessing these ESDRs are now fully capable of supporting the study of land surface vegetation processes, dynamics, long term patterns, trends and anomalies and should support various ecosystem and climate modeling efforts.

While adapting the various science algorithms to ingest and process these data records many challenges emerged, ranging from the excessive clouds and poor quality data for extended periods of time, to complex spatially and temporally dependent divergences across the different platforms/sensors making continuity quite difficult. Our proposed algorithm suite implemented solutions ranging from strict and low tolerance to noise data filters (double filter), where the input data quality is used along with the long term expected dynamic range to screen for low quality data. Additionally, a sophisticated and spatially explicit per-pixel and seasonally dependent (monthly) continuity algorithm was developed to translate the data more accurately taking into account its temporal dynamic.

To generate the land surface phenology, we modified various community algorithms to work with and take advantage of this new multi-sensor daily record. A modified hybrid and rule based Half-Max method was designed specifically for this seamless data record and operates separately on NDVI and EVI2. We've also characterized the error and uncertainty of these records using statistical means. The VI error was in the range of 5-10% VI units for NDVI and EVI2, and the date dependent phenology parameters error was 5-30 days with an average error of ±15days.

The overall result of this effort is a high quality and fourth generation (V4) data record capable of supporting accurate change, trend, and land surface vegetation dynamic studies and provide for the parameterization of various ecosystem and climate related models. All data are public and are now available via an interactive online tool, the VIP Data Explorer (vip.arizona.edu/viplab_data_explorer.php) and concurrently via the LP-DAAC (lpdaac.usgs.gov)

References

Ahl, D. E., Gower, S. T., Burrows, S. N., Shabanow, N. V., Myneni, R. B., & Knyazikhim, Y. (2006). Monitoring spring canopy phenology of a deciduous broadleaf forest using MODIS. *Remote Sensing of Environment*, *104*, 88-95.

Asner G, Townsend A, Braswell B (2000). Satellite observations of El Nino effects on Amazon forest phenology and productivity. Geophysical Research Letters, 27, 981-984.

Asrar G, Fuchs M, Kanemasu ET, Hatfield JL. Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat. Agronomy journal. 1984;76(2):300-6.

Badhwar GB. Automatic corn-soybean classification using Landsat MSS data. II. Early season crop proportion estimation. Remote Sensing of Environment. 1984 Jan 31;14(1):31-7.

Badhwar GD. Classification of corn and soybeans using multitemporal thematic mapper data. Remote Sensing of Environment. 1984 Oct 31;16(2):175-81.

Baldocchi D, Falge E, Gu L, Olson R, Hollinger D, Running S, Anthoni P, Bernhofer C, Davis K, Evans R, Fuentes J. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bulletin of the American Meteorological Society. 2001 Nov;82(11):2415-34.

Barreto A. M. and K. Didan. (2015). VI-PHEN: an online vegetation index and phenology analysis tool. 8th PROSE Symposium, Oct. 16th, 2015. The University of Arizona (2nd Prize)

Barreto A., Didan K., Riveracamcho J., Yitayew M., (2010), A New Hybrid Method for Remote Sensing Time Series Reconstruction in Support of Land Surface Phenology, Abstract IN21C-1345 presented at 2010 Fall Meeting, AGU, San Francisco, Calif., 13-17 Dec.

Beaubien, E.G., and Hall-Beyer, M. (2003) Plant phenology in Western Canada: Trends and links to the view from space. *Environmental Monitoring and Assessment* 88(1-3): 419-429

Bogaert J, Zhou L, Tucker CJ, Myneni RB, Ceulemans R. Evidence for a persistent and extensive greening trend in Eurasia inferred from satellite vegetation index data. Journal of Geophysical Research: Atmospheres (1984–2012). 2002 Jun 16;107(D11):ACL-4.

Bonan, G.B., Pollard, D. & Thompson, S.L. (1992) ffects of boreal forest vegetation on global climate. *Nature*, 359, 716–718.

Braswell, B. H., Schimel, D. S., Privette, J. L., Moore, B., Emery, W. J., Sulzman, E. W., & Hudak, A. T. (1996). Extracting ecological and biophysical information from AVHRR optical data: An integrated algorithm based on inverse modeling. *Journal of Geophysical Research-Atmospheres*, 101, 23335-23348.

Brown ME, Pinzón JE, Didan K, Morisette JT, Tucker CJ. Evaluation of the consistency of long-term NDVI time series derived from AVHRR, SPOT-vegetation, SeaWiFS, MODIS, and Landsat ETM+ sensors. Geoscience and Remote Sensing, IEEE Transactions on. 2006 Jul;44(7):1787-93.

Chen F, Dudhia J. Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. Monthly Weather Review. 2001 Apr;129(4):569-85.

Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B. and Eklundh, L. (2004): A simple method for reconstructing a high-quality NDVI time-series dataset based on the Savitzky-Golay filter.* Remote Sensing of Environment, 91 332-344.

Churkina G, Schimel D, Braswell BH, Xiao X. Spatial analysis of growing season length control over net ecosystem exchange. Global Change Biology. 2005 Oct 1;11(10):1777-87.

Claussen, M. (1994) On coupling global biome models with climate models. *Climate Research*, 4,

203–221.

Cochrane MA, Alencar A, Schulze MD, Souza CM, Nepstad DC, Lefebvre P, Davidson EA. Positive feedbacks in the fire dynamic of closed canopy tropical forests. Science. 1999 Jun 11;284(5421):1832-5.

Cooke JE, Weih M. Nitrogen storage and seasonal nitrogen cycling in Populus: bridging molecular physiology and ecophysiology. New Phytologist. 2005 Jul 1;167(1):19-30.

de Beurs, K.M., and Henebry, G.M. (2005) Land surface phenology and temperature variation in the IGBP high-latitude transects. *Global Change Biology* 11, 779-790.

DeFelice, T. P., Lloyd, D., Meyer, D. J., Baltzer, T. T., & Piraino, P. (2003). Water vapour correction of the daily 1 km AVHRR global land dataset: part I - validation and use of the Water Vapour input field. *International Journal of Remote Sensing*, 24, 2365-2375.

Dickinson, R.E. & Henderson-Sellers, A. (1988) Modelling tropical deforestation: a study of GCM land-surface parameterizations. *Quarterly Journal of theRoyal Meteorological Society*, 114, 439– 462.

Didan K. A. R., Huete. 2004. Analysis of the global vegetation dynamic metrics using MODIS Vegetation Index and land cover products. Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International. Volume 3, 2004 Page(s):2058 - 2061 vol.3

Didan K., (2010). Multi-Satellite Earth Science Data Record For Studying Global Vegetation Trends And Changes: Phase I Vegetation Index Continuity. *2010 IEEE International Geoscience and Remote Sensing Symposium, 25-30 July 2010, Honolulu, Hawaii, USA*.

Didan K., A. R . Huete (2005). Drought Induced Impacts on Vegetation Phenology Derived from Satellite-based MODIS Vegetation Index Observations. The 8th Biennial Conference of Research on the Colorado Plateau Northern Arizona University, Flagstaff, AZ. Nov. 7-10, 2005

Didan K., A. R. Huete (2005). Global Scale 1km Phenological Metrics Derived From the Moderate Resolution Imaging Spectroradiometer Vegetation Index Product: Data Characteristics and Potential Applications. AGU Fall Meeting, 2005. San Francisco.

Didan K., Huete R. A., 2006. MODIS VI Product suite, Collection 4.0 to Collection 5.0 changes. On LDOPE. Web (http://landweb.nascom.nasa.gov/QA_WWW/forPage/MOD13_VI_C5_Changes_Document_0 6_28_06.pdf)

Donahue, D.R., Sapper, J., Heidinger, A., Jelenak, A., and Kapoor, V., 2005. CLAVR-x initial operational capability at NESDIS. *International Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, 21st, San Diego, CA, 8-13 January 2005*, American Meteorological Society, Boston, MA, P10.2

Eidenshink, J. (2006). A 16-year time series of 1 km AVHRR satellite data of the conterminous United States and Alaska. *Photogrammetric Engineering and Remote Sensing*, 72, 1027-1035.

Fischer, A., A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters, Rem. Sens. Environ., 48,220-230, 1994. 1996.

Friedl M., et. al., Land Surface Phenology White Paper, 2006, lcluc.umd.edu/products/Land_ESDR/index.asp

Gallo K, Ji L, Reed B, Eidenshink J, Dwyer J. Multi-platform comparisons of MODIS and AVHRR normalized difference vegetation index data. Remote Sensing of Environment. 2005 Nov 30;99(3):221-31.

Gedney N, Valdes PJ. The effect of Amazonian deforestation on the northern hemisphere circulation and climate. Geophysical Research Letters. 2000 Oct 1;27(19):3053-6.

Gitelson AA. Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. Journal of plant physiology. 2004 Dec 31;161(2):165-73.

Goward SN, Markham B, Dye DG, Dulaney W, Yang J. Normalized difference vegetation index measurements from the Advanced Very High Resolution Radiometer. Remote sensing of environment. 1991 Mar 31;35(2):257-77.

Gutman, G., D. Tarpley, A. Ignatov and S. Olson, 1995, The enhanced NOAA Global Land datasets from the Advanced Very High Resolution Radiometer. *Bull. Amer. Meteorol. Soc.*, 76: 1141-1156.

Gutman, G., Ignatov, A. and Olson, S., 1994, Towards better quality of AVHRR composite images over land: Reduction of cloud contamination, *Remote Sens. Env.*, 50: 134-148.

Heimann M, Esser G, Haxeltine A, Kaduk J, Kicklighter DW, Knorr W, Kohlmaier GH, McGuire AD, Melillo J, Moore B, Otto RD. Evaluation of terrestrial carbon cycle models through simulations of the seasonal cycle of atmospheric CO2: First results of a model intercomparison study. Global Biogeochemical Cycles. 1998 Mar 1;12(1):1-24.

Henderson-Sellers, A. (1993) Continental vegetation as a dynamic component of a global climate model: a preliminary assessment. *Climatic Change*, 23, 337–377.

Henderson-Sellers, A. (1995) Global climate models and 'dynamic' vegetation changes. *Global Change Biology*, 1, 63–75.

Herwitz SR. Interception storage capacities of tropical rainforest canopy trees. Journal of Hydrology. 1985 Apr 25;77(1):237-52.

Hogg EH, Price DT, Black TA. Postulated feedbacks of deciduous forest phenology on seasonal climate patterns in the western Canadian interior. Journal of Climate. 2000 Dec;13(24):4229-43.

Holben, B. N., T. F. Eck, I. Slutsker, A. Smirnov, A. Sinyuk, J. Schafer, D. Giles, O. Dubovik (2006),AERONET's Version 2.0 quality assurance criteria, in Remote Sensing of Atmosphere and Clouds, edited by S.-C. Tsay et al. *Proc. SPIE,* 6408, 64080Q, doi:10.1117/12.706524.

Holben, N.B. 1996, Characteristics of maximum value compositing imagines for AVHRR data. International Journal of Remote Sensing, 7 (1986), pp. 1417–1437

Houborg R, Soegaard H, Boegh E. Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using Terra and Aqua MODIS reflectance data. Remote Sensing of Environment. 2007 Jan 15;106(1):39-58.

Huete A.R., T. Miura, Y. Kim, K. Didan, J. Privette, 2006, Assessments of multisensor vegetation index dependencies with hyperspectral and tower flux data, SPIE 6298-45, Proceedings on Remote Sensing and Modeling of Ecosystems for Sustainability III.

Huete AR, Liu HQ. An error and sensitivity analysis of the atmospheric-and soil-correcting variants of the NDVI for the MODIS-EOS. Geoscience and Remote Sensing, IEEE Transactions on. 1994 Jul;32(4):897-905.

Huete AR, Restrepo-Coupe N, Ratana P, Didan K, Saleska SR, Ichii K, Panuthai S, Gamo M. Multiple site tower flux and remote sensing comparisons of tropical forest dynamics in Monsoon Asia. Agricultural and Forest Meteorology. 2008 May 15;148(5):748-60.

Huete AR. A soil-adjusted vegetation index (SAVI). Remote sensing of environment. 1988 Aug 31;25(3):295-309.

Huete, A. R., Liu, H. Q., Batchily, K., & vanLeeuwen, W. (1997). A comparison of vegetation indices global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, 59, 440-451.

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

Huete, A.R., et al., 2006, VI White Paper, lcluc.umd.edu/products/Land_ESDR/index.asp

IPCC Reports, 2006. Solomon S, editor. Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC. Cambridge University Press; 2007 Sep 10.

Jacobson, A. K. Didan, A. Huete, 2004. "Gap filling the MODIS VI CMG Product". MODIS Vegetation Workshop, Missoula Montana. July 2004.

Jia GJ, Epstein HE, Walker DA. Greening of arctic Alaska, 1981–2001. Geophysical Research Letters. 2003 Oct 1;30(20).

Jiang Z, Huete AR, Didan K, Miura T. Development of a two-band enhanced vegetation index without a blue band. Remote Sensing of Environment. 2008 Oct 15;112(10):3833-45.

Jönsson, P. and Eklundh, L. (2002): Seasonality extraction by function fitting to time-series of satellite sensor data. IEEE Transactions on Geoscience and Remote Sensing, 40 1824-1832

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

Justice, C.O., Townshend, J.R.G., Vermote, E.F., Sohlberg, R., Descloitres, J., Roy, D., Hall, D., Salomonson, V., Riggs, G., Huete, A., Didan, K., Miura, T., Wan, Z., Strahler, A., Schaaf, C., Myneni, R., Running, S., Glassy, J., Nemani, R., El Saleous, N.Z., Wolfe, R., Preliminary land surface products from the NASA Moderate Resolution Imaging Spectroradiometer(MODIS). Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS 00), Honolulu, Hawaii, July 2000.

Kaduk, J., and M. Heimann, A prognostic phenology model for global terrestrial carbon cycle models, Clim. Research, 6, 1-19, 1996.

Kaufman YJ, Tanre D. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. Geoscience and Remote Sensing, IEEE Transactions on. 1992 Mar;30(2):261-70.

Keeling C. D., Chin J.F.S, Whorf T.P. (1996). Increased activity of northern vegetation inferred from atmospheric CO2 measurements. Nature, 382, 146-149.

Keeling RF, Piper SC, Heimann M. Global and hemispheric CO2 sinks deduced from changes in atmospheric O2 concentration. Nature. 1996 May 16;381(6579):218-21.

Kidwell, K. B. (1998). *NOAA Polar Orbiter Data User's Guide*: National Oceanic and Atmospheric Administration, Washington, D.C.

Lambin EF, Geist HJ, Lepers E. Dynamics of land-use and land-cover change in tropical regions. Annual review of environment and resources. 2003 Nov;28(1):205-41.

Lyapustin AI, Wang Y, Laszlo I, Hilker T, Hall FG, Sellers PJ, Tucker CJ, Korkin SV. Multi-angle implementation of atmospheric correction for MODIS (MAIAC): 3. Atmospheric correction. Remote Sensing of Environment. 2012 Dec 31;127:385-93.

Melillo, J.M., Prentice, I.C., Farquhar, G.D., Schulze, E.-D. & Sala, O.E. (1996) Terrestrial biotic responses to environmental change and feedbacks to climate. *Climate change 1995. Contribution of WG I to the second assessment report of the intergovernmental panel on climate change* (ed. By J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg & K. Maskell), pp. 445– 481. Cambridge University Press, Cambridge, UK.

Menzel, A. & Fabian, P. Growing season extended in Europe. *Nature* 397, 659 (1999).

Mintz, Y. (1984) The sensitivity of numerically simulated climates to land–surface boundary conditions. *The global climate* (ed. by J.T. Houghton), pp. 79–105. Cambridge University Press, Cambridge, UK.

Miura, T., Huete, A., & Yoshioka, H. (2006). An empirical investigation of cross-sensor relationships of NDVI and red/near-infrared reflectance using EO-1 hyperion data. *Remote Sensing of Environment*, 100, 223-236.

Moore KE, Fitzjarrald DR, Sakai RK, Goulden ML, Munger JW, Wofsy SC. Seasonal variation in radiative and turbulent exchange at a deciduous forest in central Massachusetts. Journal of Applied Meteorology. 1996 Jan;35(1):122-34.

Morisette, J., Nickeson, J.E., S. Garrigues, F. Baret, A. Huete, K. Didan, T. Miura, W.van Leeuwen, M. Friedl, 2006, Report from the CEOS Land Product ValidationTopical Workshop on the Validation of GlobalVegetation Indices and their Time Series, *The Earth Observer*, Vol. 18(6), 34-37.

Myneni, R.B., C.D Keeling, C.J Tucker, G Asrar, R.R Nemani Increased plant growth in northern high latitudes from 1981–1991. Nature, 386 (1997), pp. 698–702

Nagler PL, Scott RL, Westenburg C, Cleverly JR, Glenn EP, Huete AR. Evapotranspiration on western US rivers estimated using the Enhanced Vegetation Index from MODIS and data from eddy covariance and Bowen ratio flux towers. Remote sensing of environment. 2005 Aug 15;97(3):337-51.

Ollinger SV, Richardson AD, Martin ME, Hollinger DY, Frolking SE, Reich PB, Plourde LC, Katul GG, Munger JW, Oren R, Smith ML. Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. Proceedings of the National Academy of Sciences. 2008 Dec 9;105(49):19336-41.

Prince SD. Satellite remote sensing of primary production: comparison of results for Sahelian grasslands 1981-1988. International Journal of Remote Sensing. 1991 Jun 1;12(6):1301-11.

Rahman AF, Sims DA, Cordova VD, El-Masri BZ. Potential of MODIS EVI and surface temperature for directly estimating per-pixel ecosystem C fluxes. Geophysical Research Letters. 2005 Oct 1;32(19).

Randerson J, Field C, Fung I et al., (1999). Increase in early season ecosystem uptake explain recent changes in the seasonal cycle of atmospheric CO2 at high northern latitudes. Geophysical Research Letters, 26, 2765-2768.

Rao, C.R.N. and J. Chen, 1994: Post-launch calibration of the visible and near-IR channels of AVHRR on NOAA-7, -9, and -11 spacecraft. *NOAA Tech. Rpt. NESDIS 78*, US Dept. of Commerce, NOAA, 22 pp.

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W. & Ohlen, D.O. (1994) Measuring phenological variability from satellite imagery. *Journal of Vegetation Sci*ence, 5, 703-714.

Reich PB, Borchert R. Changes with leaf age in stomatal function and water status of several tropical tree species. Biotropica. 1988 Mar 1:60-9.

Richardson AD, Braswell BH, Hollinger DY, Jenkins JP, Ollinger SV. Near-surface remote sensing of spatial and temporal variation in canopy phenology. Ecological Applications. 2009 Sep;19(6):1417-28.

Rowntree, P.R. (1988) Review of general circulation models as a basis for predicting the effects of vegetation change on climate. *Forests, climate, and hydrology: regional impacts* (ed. by E.R.C. Reynolds & F.B. Thompson), pp. 162–196. The United Nations University, Tokyo.

Running SW, Justice CO, Salomonson V, Hall D, Barker J, Kaufmann YJ, Strahler AH, Huete AR, Muller

JP, Vanderbilt V, Wan ZM. Terrestrial remote sensing science and algorithms planned for EOS/MODIS. International journal of remote sensing. 1994 Nov 1;15(17):3587-620.

Running SW, Nemani RR. Relating seasonal patterns of the AVHRR vegetation index to simulated photosynthesis and transpiration of forests in different climates. Remote Sensing of Environment. 1988 Mar 31;24(2):347-67.

Saleska SR, Didan K, Huete AR, Da Rocha HR. Amazon forests green-up during 2005 drought. Science. 2007 Oct 26;318(5850):612-.

Schwartz MD, Reed BC. Surface phenology and satellite sensor-derived onset of greenness: an initial comparison. International Journal of Remote Sensing. 1999 Jan 1;20(17):3451-7.

Shabanov NV, Zhou L, Knyazikhin Y, Myneni RB, Tucker CJ. Analysis of interannual changes in northern vegetation activity observed in AVHRR data from 1981 to 1994. Geoscience and Remote Sensing, IEEE Transactions on. 2002 Jan;40(1):115-30.

Sims DA, Rahman AF, Cordova VD, El-Masri BZ, Baldocchi DD, Flanagan LB, Goldstein AH, Hollinger DY, Misson L, Monson RK, Oechel WC. On the use of MODIS EVI to assess gross primary productivity of North American ecosystems. Journal of Geophysical Research: Biogeosciences (2005–2012). 2006 Dec 1;111(G4).

Steven, M. D., Malthus, T. J., Baret, F., Xu, H., & Chopping, M. J. (2003). Intercalibration of vegetation indices from different sensor systems. *Remote Sensing of Environment*, *88*(4), 412-422.

Stowe, L. L., Davis, P. A., & McClain, E. P. (1999). Scientific basis and initial evaluation of the CLAVR-1 global clear cloud classification algorithm for the advanced very high resolution radiometer. *Journal of Atmospheric and Oceanic Technology*, 16, 656-681.

Trishchenko AP, Cihlar J, Li Z. Effects of spectral response function on surface reflectance and NDVI measured with moderate resolution satellite sensors. Remote Sensing of Environment. 2002 Jul 31;81(1):1-8.

Trishchenko, A. P., Cihlar, J., & Li, Z. (2002). Effects of spectral response function on surface reflectance and NDVI measured with moderate resolution satellite sensors. *Remote Sensing of Environment*, 81, 1-18.

Tsend-Ayush, J., Miura, T., Didan, K., and Barreto-Munoz, A. (2010). Development of a multi-sensor vegetation index translation algorithm. *2010 IEEE International Geoscience and Remote Sensing Symposium, 25-30 July 2010, Honolulu, Hawaii, USA*.

Tsend-Ayush, J., Miura, T., Didan, K., and Barreto-Munoz, A. (2010). Deriving multi-sensor translation equations for generation of a long-term vegetation index data record. *2010 Association of American Geographers Annual Meeting, 14-18 April 2010, Washington, DC, USA.*

Tucker CJ, Pinzon JE, Brown ME, Slayback DA, Pak EW, Mahoney R, Vermote EF, El Saleous N. An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. International Journal of Remote Sensing. 2005 Oct 20;26(20):4485-98.

Tucker CJ. Red and photographic infrared linear combinations for monitoring vegetation. Remote sensing of Environment. 1979 May 31;8(2):127-50.

Tucker, C.J., P.J. Seller. Satellite remote sensing for primary production. Remote Sensing, 7 (1986), pp. 1395–1416

Ünsalan C, Boyer KL. A system to detect houses and residential street networks in multispectral satellite images. Computer Vision and Image Understanding. 2005 Jun 30;98(3):423-61.

Vaiopoulos D, Skianis GA, Nikolakopoulos K. The contribution of probability theory in assessing the efficiency of two frequently used vegetation indices. International Journal of Remote Sensing. 2004 Oct 1;25(20):4219-36.

Vermote EF, El Saleous NZ, Justice CO. Atmospheric correction of MODIS data in the visible to middle infrared: first results. Remote Sensing of Environment. 2002 Nov 30;83(1):97-111.

Vermote, E. and Saleous, N.Z. (2006). Calibration of NOAA16 AVHRR over a desert site using MODIS data, Remote Sens.Environ., 105, 214-220.

Vermote, E., & Kaufman, Y. J. (1995). Absolute calibration of AVHRR visible and near-infrared channels using ocean and cloud views. *International Journal of Remote Sensing*, 16, 2317-2340.

Wardlow BD, Egbert SL, Kastens JH. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains. Remote Sensing of Environment. 2007 Jun 15;108(3):290-310.

Waring RH, Coops NC, Fan W, Nightingale JM. MODIS enhanced vegetation index predicts tree species richness across forested ecoregions in the contiguous USA. Remote Sensing of Environment. 2006 Jul 30;103(2):218-26.

Watson RT, Zinyowera MC, Moss RH. Climate Change 1995 impacts, adaptations and mitigation of climate change: Scientific-technical analysis. Cambridge University Press; 1996.

Watson, R.T., Zinyowera, M.C. & Moss, R.H., eds (1996) *Climate change 1995*. *Contribution of WG II to the second assessment report of the intergovernmental panel on climate change*. 878 pp. Cambridge University Press, Cambridge, UK.

Welsch C, Swenson H, Cota S, DeLuccia F, Haas JM, Schueler C, Durham RM, Clement JE, Ardanuy PE. VIIRS (Visible Infrared Imager Radiometer Suite): a next-generation operational environmental sensor for NPOESS. InGeoscience and Remote Sensing Symposium, 2001. IGARSS'01. IEEE 2001 International 2001 (Vol. 3, pp. 1020-1022). IEEE.

White MA, BEURS D, Kirsten M, DIDAN K, INOUYE DW, RICHARDSON AD, JENSEN OP, O'KEEFE JO, ZHANG G, NEMANI RR, LEEUWEN V. Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. Global Change Biology. 2009 Oct 1;15(10):2335-59.

White MA, Hoffman F, Hargrove WW, Nemani RR. A global framework for monitoring phenological responses to climate change. Geophysical Research Letters. 2005 Feb 1;32(4).

White MA, Nemani RR, Thornton PE, Running SW. Satellite evidence of phenological differences between urbanized and rural areas of the eastern United States deciduous broadleaf forest. Ecosystems. 2002 Apr 1;5(3):260-73.

White MA, Thornton PE, Running SW. A continental phenology model for monitoring vegetation responses to interannual climatic variability. Global biogeochemical cycles. 1997 Jun 1;11(2):217-34.

Xiao X, Hagen S, Zhang Q, Keller M, Moore B. Detecting leaf phenology of seasonally moist tropical forests in South America with multi-temporal MODIS images. Remote Sensing of Environment. 2006 Aug 30;103(4):465-73.

Xiao X, Zhang Q, Braswell B, Urbanski S, Boles S, Wofsy S, Moore B, Ojima D. Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data. Remote Sensing of Environment. 2004 May 30;91(2):256-70.

Xiao X, Zhang Q, Saleska S, Hutyra L, De Camargo P, Wofsy S, Frolking S, Boles S, Keller M, Moore B. Satellite-based modeling of gross primary production in a seasonally moist tropical evergreen forest. Remote Sensing of Environment. 2005 Jan 15;94(1):105-22.

Yang J, Ding Y, Chen R. Spatial and temporal of variations of alpine vegetation cover in the source regions of the Yangtze and Yellow Rivers of the Tibetan Plateau from 1982 to 2001. Environmental Geology. 2006 Jun 1;50(3):313-22.

Yoshioka, H., Miura, T., & Yamamoto, H. (2005). Relationships of spectral vegetation indices for continuity and compatibility of satellite data products. *Multispectral and Hyperspectral Remote Sensing Instruments and Applications II (A. M. Larar, M. Suzuki, Q. Tong, Eds.)*, Proc. SPIE Vol. 5655, 233-240.

Zhang X, Friedl MA, Schaaf CB, Strahler AH, Hodges JC, Gao F, Reed BC, Huete A. Monitoring vegetation phenology using MODIS. Remote sensing of environment. 2003 Mar 31;84(3):471-5.

Zhang X, Friedl MA, Schaaf CB. Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. Journal of Geophysical Research: Biogeosciences (2005–2012). 2006 Dec 1;111(G4).

Zhang X, Friedl MA, Schaaf CB. Sensitivity of vegetation phenology detection to the temporal resolution of satellite data. International Journal of Remote Sensing. 2009 Apr 20;30(8):2061-74.