Vegetation Index greenness global data set
White Paper for NASA ESDR/CDR (April 2006)

Alfredo R. Huete[1], Karl F. Huemmrich[2], Tomoaki Miura[3], Xiangming Xiao[4], Kamel Didan[1], Willem van Leeuwen[1], Forrest Hall[2], Compton J. Tucker[2]
[1]Univ. of Arizona; [2]NASA/GSFC; [3]Univ. of Hawaii; [4]Univ. of New Hampshire

Description of Earth System Data Record (ESDR) and Climate Data Record (CDR)

Vegetation Indices (VI) are optical measures of vegetation canopy 'greenness', a direct measure of photosynthetic potential resulting from the *composite* property of total leaf chlorophyll, leaf area, canopy cover, and structure. VI's are widely used as proxies in estimating canopy state variables (leaf area index, fraction of absorbed photosynthetically-active radiation, chlorophyll content, vegetation fraction) and canopy biophysical processes (photosynthesis, transpiration, net primary production). The Vegetation Index ESDR/CDR includes the Normalized Difference Vegetation Index (NDVI) and an Enhanced Vegetation Index (EVI) at coarse (~5-8km) and moderate (~250m -1km) resolutions to effectively characterize ecosystem states and processes for long-term climate change and near real-time operational applications.

Scientific rationale and importance of measurement and expected end uses

The assessment of vegetation condition, cover, processes, and changes are major components of national and international global change research programs and are topics of considerable societal relevance. Spectral vegetation indices are among the most widely used satellite data products providing key measurements for climate, hydrologic, and biogeochemical studies, land cover and land cover change detection, and natural resource management and sustainable development. There is currently a consistent NDVI vegetation record extending more than 2 decades from the NOAA Advanced Very High Resolution Radiometer (AVHRR), which have contributed significantly to the recent advancement of Earth System Science, in particular to global, biome, and agricultural primary production; interannual fluctuations and impacts of ENSO on primary production; phenology; and climate change and variability, e.g., Myneni *et al.* (1997) showed evidence of a lengthening of the plant growing season at northern latitudes in response to global temperature increases. Over 3,500 papers have been published using vegetation indices and these papers have been referenced by >60,000 other scientific papers.

More recent applications of highly calibrated VI's at moderate spatial resolution (250m to 1 km) and weekly (8-16 day) time intervals from the MODIS and SPOT VEGETATION (VGT) sensors have demonstrated their utility in characterizing the structure, metabolism, and functioning of ecosystems which affect net ecosystem exchange of CO_2 and water between the land and the atmosphere. MODIS and VGT EVI data have been incorporated into the Vegetation Photosynthesis Model (VPM) to produce tower-calibrated predictions of gross primary productivity (GPP) flux across a series of biomes (Xiao *et al.* 2005) and EVI has depicted phenology cycles in dense Amazon rainforests with close couplings to tower-calibrated GPP measurements of carbon fluxes in both intact rainforest and forest conversion sites.

The VI data record is important to several NASA Earth Science Focus Areas in providing one of the longest, consistent measurement records for Earth System studies. NASA's Global Inventory Modeling and Mapping Studies (GIMMS) group has produced a global NDVI data set at 8-km resolution from 1981 to 2004, maximizing the length and quality of the AVHRR data set (Tucker *et al.* 2005). The USGS EROS Data Center provides finer resolution, 1 km AVHRR-NDVI for the conterminous U.S. from 1989. The Sea-viewing Wide Field-of-view Sensor (SeaWiFs) from Orbital Imaging Corporation generates an NDVI product and foreign institutions have also developed 1 km NDVI time series from AVHRR, VGT, and MODIS (e.g., Canadian Centre for Remote Sensing).

Due to their simplicity, VI's are readily fused across sensor systems facilitating an underlying need to ensure continuity of critical data sets to study climate-related processes. A wide range of the Earth science and applications user community groups are using VI time-series, particularly at 250 m resolution, in natural resource management, agriculture, invasive species, public health, and in many climate, hydrology, and biogeochemical models that ingest the VI data record. Lastly, crucial canopy biophysical properties such as fAPAR and LAI are not currently listed as ESDR's for the NPOESS era and on-demand biophysical products may need to be derived from the VI data record.

Scientific requirements for the measurement

The main requirement of a VI measurement is to combine the chlorophyll-absorbing red spectral region with the non-absorbing, leaf reflectance signal in the near-infrared (NIR) to provide a consistent and *robust* measure of area-averaged canopy photosynthetic capacity. The VI algorithm is designed to extract the active *greenness* signal from terrestrial land covers which are complex structural and biochemical assemblages of chlorophyll, leaf and canopy layers, non-photosynthetic vegetation components, and underlain by substrate layers (soil, rock, & litter) that may periodically be covered with snow or standing water.

The VI product is a *composite* of frequent measurements over set time intervals to enable the removal of cloud and cloud shadow contamination and constrain sensor view angles. The VI is produced for local solar zenith angle conditions and computed from surface directional reflectances that have been atmospherically-corrected to remove the effect of absorbing gases, molecular scattering, aerosols, and water vapor. Measurements are needed at various spatial and temporal scales to capture the important spatial and temporal complexities associated with vegetation dynamics, biogeochemical and hydrologic processes, and for quantifying ecosystem responses to climate variability.

- Temporal resolutions between 8- and 16-day are normally needed to reduce sources of noise and variability and provide a gap-free and cloud-free product.
- Desired spatial resolutions include a coarse (5 - 8 km, climate modeling grid, CMG) resolution and a moderate (0.25 -1km) resolution VI data record.

However, since climate model grid size is likely to change over time with advancements in algorithms, the VI product should retain sensor measurements at their finest spatial and temporal resolutions.

The accuracy of the VI product vary in space and time, due to geographic and seasonal variations in cloud cover persistence and unresolved clouds; the quality of upstream surface reflectance retrievals and cross band (red, NIR) spectral integrity; systematic biases introduced by sun angle (seasonal, and latitudinal), topography, canopy background (soil, water, snow); and sensor calibration and geolocation accuracies. Current estimates of uncertainty in the AVHRR-

NDVI record are approximately 0.04 precision; +/-0.05 accuracy, and 0.04 long term stability (NRC, 2003). The desired uncertainties for MODIS, which are not fully assessed, are for much improved values of 0.01 for both precision and accuracy. VI accuracies further vary in their measurement sensitivity to *greenness* and ability to capture essential biophysical phenomena with adequate fidelity. For a CDR, the VI should be able to detect a <5% change in GPP and <5% foliage density change.

Overlapping satellite observations among mission-based VI records are required to generate a seamless VI ESDR/CDR. This provides opportunities for simultaneous calibration of old and new sensors against each other and over a common set of targets on the Earth. 1 to 2 year overlap periods across current and future sensors (e.g., MODIS, VGT, AVHRR/3, MERIS, and VIIRS) are desirable (Tucker *et al.*, 2005). The VI product and input data sets that fulfill the traceability requirement should be uniformly archived and continuously available.

A quick turn-around-time between data acquisition and VI product delivery greatly enhances the benefit for many user groups that need to respond quickly to observed changes in vegetation activity due to drought, fire and extreme weather events (famine response, cattle quota, fire danger etc.). For the VI-CDR, a 50-year (>2030) time series is needed to study important climate-related processes. The time frames needed for management and decision making applications vary greatly and can approach 50-100 years (e.g., forest succession, disturbance, and recovery).

Approach to generating the measurement

Vegetation Indices are seamless data products that are computed from the same mathematic formulae across all pixels in time and space, without prior assumptions of biome type, land cover condition, or soil type, and thus represent actual, long-term measurements of the land surface. The VI product works optimally with cloud filtering, radiometric calibration, precise geolocation, and a snow mask. In addition, the product performs best using top-of-canopy reflectance inputs, corrected for atmospheric ozone, molecular scattering, aerosol, and water vapor. The NDVI is computed as,

$$\mathrm{NDVI} = [\rho_{NIR} - \rho_{red}] \ / \ [\rho_{NIR} + \rho_{red}] \tag{1}$$

The EVI gains its heritage primarily from the soil adjusted vegetation index (SAVI) and is an optimized combination of red and NIR bands, designed to extract canopy *greenness*, independent of biases from differences in soil background (Huete *et al.* 2002). This is accomplished using a 2-layer soil-vegetation radiant transfer equation, $\rho_C = \rho_v + \tau^2_C\, \rho_s$, where canopy reflectance (ρ_C) is the sum of the vegetation layer reflectance (ρ_v) and the canopy transmitted (τ_C) soil reflectance (ρ_s). The EVI also incorporates an 'aerosol resistance' term developed in the ARVI and SARVI equations that utilize the blue reflectance to stabilize aerosol influences and mis-corrections in red reflectances,

$$\mathrm{EVI} = 2.5\, [\rho_{NIR} - \rho_{red}] \ / \ [L + \rho_{NIR} + C_1\, \rho_{red} - C_2\, \rho_{blue}], \tag{2}$$

where L is the canopy background adjustment factor, and C_1 and C_2 are the aerosol resistance weights. The coefficients of the EVI equation are L=1; C_1=6 and C_2 =7.5. If the aerosol resistance term cannot be performed (no blue band available), C_1 and C_2 may be collapsed to values of 1 and 0, respectively, for continuity purposes (e.g., backward continuity in AVHRR).

The composited VI product is generated with a constrained-view angle – maximum value compositing (CV-MVC) methodology to produce a gap-free VI product in space and time that minimizes residual cloud and atmosphere effects and constrains view angles. This is modified from the heritage AVHRR-MVC methodology, which selects the highest NDVI value over a compositing cycle. With advancements in atmosphere correction, the highest NDVI is increasingly biased toward off-nadir pixel selection. The CV-MVC method selects the closest to nadir NDVI value from the 2 or 3 highest NDVI values remaining after QA filtering of the data, and retains actual VI measurements with minimal view angle biases. Additional issues in re-processing of VI data include end-of-year processing, a gap-filled product, using recognized noise removal and smoothing algorithms, and referencing to multi- sensor calibration/reference sites.

Intended sources for the measurement

Vegetation Index land surface observations have been made for the entire globe from orbiting satellites at approximately daily intervals since 1981. This includes the operational NOAA- AVHRR, the Terra and Aqua MODIS sensors, SPOT-VEGETATION (VGT), SeaWiFS, MERIS, and the Visible Infrared Imaging Radiometer Suite (VIIRS) on a new series of NPOESS satellites that will replace the AVHRR series of instruments. Gaps in the VI data record could be filled with higher resolution sensors, such SPOT-HRV and Landsat or other moderate resolution sensors such as SAC-C, CBERS, IRS-1,2, and other international sensor platforms.

Necessary supporting activities, tasks

Because the compositing process limits the temporal resolution of the VI data and may introduce bias, independent surface measurements of VIs and plant activity should be made on the ground. The seasonal and interannual patterns from the satellite VI data record can be validated and cross-calibrated with 3 sets of independent surface measurements;

(1) Globally distributed eddy covariance flux tower measurements of CO_2 fluxes. NDVI-like hemispherical sensors installed at some flux tower sites offer promising validation and calibration opportunities,
(2) Evolving phenology networks with budbreak, greening, and browning information; and
(3) Biophysical and optical samplings at leaf, canopy, and landscape scales in large field campaigns (Bigfoot, HAPEX, BOREAS, SAFARI, LBA, etc.) provide validation data of through an array of vegetation types, soils, and climates.

Intercalibration and data continuity across multiple sensors are critical components of the VI CDR/ESDR. Long-term VI observations from multiple satellites require translation equations to ensure continuity and compatibility across different sensor characteristics (e.g., band pass filters, spatial resolution, signal to noise ratio) and product algorithms (e.g., atmospheric correction and compositing schemes) and remove any systematic biases (e.g. data acquisition times).

Fine resolution VI data (e.g. AVIRIS, ASTER, Ikonos, etc.) will provide periodic checks of multi-sensor compatibility and information to resolve issues of surface heterogeneity, translation, and leaf to landscape scaling in the VI product. An important benefit of inter-calibration is having multiple VI data sources in case one or more satellites or instruments fail. Baseline stability measures are needed over hyperarid deserts to monitor multi-instrument characteristics/ calibration and ensure the accuracy, integrity, and usefulness of long term the VI data record.

Relationships to other products and programs

The MODIS VI product is an important input dataset to MODIS land cover/ land cover dynamics product and is used in backup algorithms for MODIS LAI, fAPAR, and NPP products, which were previously derived directly from AVHRR-NDVI. VI data are routinely assimilated in many different applications, ranging from agricultural, invasive species and natural resource monitoring and forecasting systems to wildfire, disease, pest and famine early warning systems, in support of scientific and societal purposes. Several national application areas, including Agricultural Efficiency, Carbon Management, Ecological Forecasting, Homeland Security, Invasive Species, Public Health and Water Management, involve agencies that assimilate VIs as part of their decision support tools of value and benefit to resource management, policy decisions and resource exploration.

MODIS NDVI data are used in Famine Early Warning Systems (FEWS); assimilated into decision support systems for crop assessments by the Foreign Agricultural Service (FAS); and are part of the Invasive Species Forecasting System (ISFS), co-developed by the USGS and NASA. AVHRR and MODIS VI contributions to carbon sequestration are found in partnerships between NASA, USDA, Department of Energy, and the Environmental Protection Agency. MODIS EVI data is used in operational deforestation monitoring in the Amazon (SIAD-Amazônia) and in paddy rice agriculture mapping, methane emissions, and risk assessment of avian flu in Asia.

Key references

Huete, A., Didan, K., Miura, T., et al., 2002, Overview of the radiometric and biophysical performance of the MODIS vegetation indices, *Remote Sens. Environ*. 83:195-213.

Myneni, R. B., Keeling, C.D., Tucker, C.J., Asrar, G., and Nemani, R.R., 1997, Increased plant growth in the northern high latitudes from 1981 to 1991, *Nature*, *386*, 698.

National Research Council, 2003, Issues in the integration of research and operational satellite systems for climate research, I. Science and Design, Committee on Earth Studies, Space Studies Board, National Academy Press, Washington D.C.

Tucker, C. J., Pinzon, J. E., Brown, et al., 2005, An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data, *Int. J. of Remote Sensing*, 26, 4485-4498.

Xiao, X., Zhang, Q., Saleska, et al., 2005, Satellite-based modeling of gross primary production in a seasonally moist tropical evergreen forest, *Remote Sens. Environ. 94*, 105-122.