Vegetation Phenology and Vegetation Index Products from Multiple Long Term Satellite Data Records

Project Overview

Kamel Didan
VIP Research Group, The University of Arizona

NASA MEASURES #NNX08AT05A
Context

• There are three general frameworks to this project
  – Vegetation makes up most of the Earth land cover and plays a key role in the biosphere and atmosphere functioning
  – Programmatically NASA, & other entities, invested indifferent RS platforms with most derived data records being sensor dependent leading to very short consistent temporal coverage
  – The value of the Vegetation index data record (30+) to science and research is supported by a 30+year of publications, major science findings, easy access, and reliability
context - 1

- Planet earth is distinguished (so far) from all other celestial objects by two very special things:
  - Water (which is becoming not so special)
  - And vegetation

- The amount of vegetation in the ocean is huge and central to the food chain, but for the most part land provides most of the vegetation in the human diet

- Vegetation couples the biosphere to the atmosphere, and as such plays a major role in carbon storage, cycling, and moderates the weather/climate
Context - 2

Official context: One of the stated goals of NASA Making Earth Science Data Records for Use in Research Environments (MEaSUREs) program is the support of the Earth Science research community by providing reliable Earth System (Science Data Records (ESDR).

Practical meanings: These products are expected not only to be of high quality but should also combine data from multiple sources to form the long and coherent measurements required for studying climate change impact on the Earth system.

What is an ESDR
(NASA) An Earth Science/System Data Records (ESDRs) are observations of a parameter of the Earth system optimized to meet requirements to address Earth science questions and to provide for applications. With explicit attention to error, uncertainty, and precision.

CDR - Climate Data Record(NOAA)
ECV - Essential Climate Variable (European)

In practice this means: Support of Earth Science Research in the context of climate change with long term multi-sensor reliable and coherent data records.
Vegetation indices, which capture the aggregate functioning of a canopy, are:

- Robust, simple and make no assumptions, and are fully traceable to what the sensor acquired
  - Key characteristics for accurate long term change research
  - VI could be applied at any level of data refinement (TOA, TOC, partial or full correction)
- The most widely used RS measurement for:
  - Studying vegetation
  - Large-scale ecosystem processes, and
  - Extracting phenology information (impact of climate change)

Changes in vegetation phenology depict an integrated response to change in environmental factors and provide valuable information to global change research.

Knowledge of phenologic variability and the environmental conditions at play are prerequisite to inter-annual studies and predictive modeling of land surface responses to climate change.
Motivations

• The following 2 statistics should justify monitoring our planet land surface vegetation
  – 54% increase in global population (from 4.45B in 1980 to 6.89B in 2010)
  – 14% increase in atm. CO2 (340ppmv in 1980 to 389ppmv as of Nov. 2010)
  – And this is not abating
  – We’re having serious impact on the Earth and its resources that require monitoring
Recent trends and observations

<table>
<thead>
<tr>
<th>Biota</th>
<th>Location</th>
<th>Change</th>
<th>Climate Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treeline</td>
<td>Europe, New Zealand</td>
<td>Upward shift</td>
<td>Warming</td>
</tr>
<tr>
<td>Arctic Shrub Tundra</td>
<td>Alaska</td>
<td>Spread</td>
<td>Warming</td>
</tr>
<tr>
<td>Alpine Plants</td>
<td>Alps</td>
<td>Upward shift</td>
<td>Warming</td>
</tr>
<tr>
<td>Biota</td>
<td>Antarctica</td>
<td>Spread</td>
<td>Melting</td>
</tr>
<tr>
<td>Zooplankton</td>
<td>Calif. &amp; N. Atlantic</td>
<td>Population increase</td>
<td>Warming</td>
</tr>
<tr>
<td>Butterflies (39 spp.)</td>
<td>N Amer. &amp; Europe</td>
<td>Northward shift, 200 km</td>
<td>Warming</td>
</tr>
<tr>
<td>Birds, lowland</td>
<td>Costa Rica</td>
<td>Expansion upward</td>
<td>Dry-season mists</td>
</tr>
<tr>
<td>Birds, migratory</td>
<td>England</td>
<td>Northward shift, 20 km</td>
<td>Winter Warming</td>
</tr>
<tr>
<td>Birds, migratory</td>
<td>England</td>
<td>Earlier but, too late for food</td>
<td>Spring Warming</td>
</tr>
<tr>
<td>Foxes, red, white</td>
<td>Canada</td>
<td>Northward boundary shift</td>
<td>Warming</td>
</tr>
</tbody>
</table>

Role of remote sensing

- Remote sensing emerged as an ideal technology/tool for studying the earth system repetitively, at large scale and ‘quite’ accurately
- The earth science community must focus on climate science and satellites are crucial for this (many major earth related findings are satellite based)
- AVHRR-MODIS-[VIIRS] data are the most important climate science measurements (land-ocean-atmosphere) - [Look at the publication record].
- Landsat observations are the next most important land climate science observation
- Canopy Lidar is the 3rd most important land climate science observation
  - We must also support systematic ground measurements that provide key validation information

Tucker C. (Global Veg. Wrkshp, Missoula, MT, 2006)
Role of Vegetation indices

• By far the RS based VI data record (Time Series) is the most successful (publication record, application record, science & research, ease of use, accessibility to non RS people, etc...)
• It permits the study of change that is at the core of Earth Science and climate change research
• But it is still only a proxy
  • New research suggests that it may in fact be capturing the complexity of photosynthesis (complex function that is controlled by many parameters)
• The VI record however lacks continuity, and that is a serious obstacle
  • Missions, specification, and objectives change
  • Algorithms change
Value of VI Record

• Historical value of AVHRR (30+ yrs).
  – And now MODIS (~11 years) + other sensors

• Wealth of findings (science research & applications)

• Simplicity
  – Widely used (Agriculture yield, phenology, to vector borne disease, etc...)
  – VI ability to proxy many land surface physical/biophysical characteristics quite nicely

• Major science claims based on this data record
  – Phenology – lengthening of snow-free season in the arctic
  – Increased NPP in N America (now it is decreasing)
  – Rain forest seasonality
  – Fire frequencies and Land cover changes
Vegetation Indices

- NDVI – widely used but with saturation issues
- EVI – a complimentary Index that works better over dense canopies/vegetation
  - But requires the blue band (not available on AVHRR), and when available has a poor S/N ratio.
  - EVI also shows poor performance over areas with:
    - Residual cloud
    - Sub-pixel snow
    - And due to blue band noise
EVI

- Resistance to background and atmosphere noise
- Does not saturate over high biomass (High-C)
- Strong correlation with canopy structure (Phenology)
- Time series measures of MODIS EVI have been shown highly correlated with flux tower photosynthesis (GPP) in both tropical and temperate ecosystems at seasonal scales
- Widely and successfully used and adopted

EVI can greatly complement existing NDVI time series in forest (high-C) canopies

Actual EVI2 from MODIS, SPOT-VGT, and AVHRR

\[ y = 8282x - 2118 \]

\[ R^2 = 0.74 \]
CDR Definition

• A climate data record is a time series of measurements of sufficient length, consistency, and continuity to determine climate variability and change. (NRC, 2004). The NRC further segmented satellite-based CDRs into:
  – Fundamental CDRs (FCDRs), which are calibrated and quality-controlled sensor data that have been improved over time, and
  – Thematic CDRs (TCDRs), which are geophysical variables derived from the FCDRs, such as sea surface temperature and cloud fraction.

• These definitions still need refining
• This work will contribute to this discourse
ESDR Definition

- Earth System Data Records (ESDRs) are observations of a parameter of the Earth system optimized to meet requirements to address Earth science questions and to provide for applications. With explicit attention to error, uncertainty, and precision.
  - Low level and high level products are involved in ESDR's
    - Higher level products depend on products such as reflectance and vegetation index.
    - Hierarchical organization is useful
  - Need to derive priority from the importance of the end use
  - Explicit attention to error, uncertainty, and precision is required in definition and production.
  - Issue of consistency between user sub-groups and ESDR’s – important for the modeling community
  - Need to consider what will be needed to create the retrospective data record
- These definitions still need refining
- This work will contribute to this discourse
Prior to ‘00 there was only AVHRR. For the near future long term records will need to build on AVHRR.

Post ‘00 many ‘newer’ sensors to choose from but with different characteristics.
Challenges

- **Sensor platforms**
  - Bandpass
  - Spatial resolution
  - Spectral/Radiometric resolution
  - Calibration

- **Algorithm**
  - Compositing
  - Spatial aggregation
  - VI formulation
Challenges

• With virtually no exceptions all ecosystem research has at some point used time series data from AVHRR, MODIS, SeaWiFS, and/or other sensors.
• In most cases a single data source was used (AVHRR only, MODIS only, etc...)
  • Limiting the value and validity of our work and findings
• Dealing with climate change requires breaking that artificial barrier, i.e. mixing and fusing data sources.
• To properly study change you need long time series (WMO: at least 30 years).
• There is a need for consistent, continuous and well characterized earth data records
• Until continuity is built into the RS system(s) we’ll have to mix data sources
Project hypotheses

• Our hypotheses are then
  – There is so much to gain by creating ‘consistent’ and long term records useful for global climate change research?
    • To properly study change you need long time series (WMO: at least 30 years)
  – VI are the most traceable, easily replicated, and best behaved RS measurements
  – RS based phenology is an emerging science that captures the integrated impact of climate and environmental changes
  – All RS based Phenology algorithms are exclusively driven by VI time series.
  – There is a wealth of information to help characterize these two measurements, establish their error and uncertainty. Data anomalies are then easily traced and their impact known
**Project Objectives**

- In order to study climate change impact and trends we must use long term data (medium to short term are not sufficient and could be misleading). The WMO suggests a 30 year record at a minimum.
- With the current available records there is still strong need for the historic AVHRR data (at least in the near future)
- Create a long term data record
  - Backward compatible EVI to take advantage of EVI improvements
  - Independent of sensor and focused on consistency to enable change detection
  - Derive a phenology algorithm/ESDR
Project Objectives

• The goal of this project is then to generate a VI + ‘Land Surface’ Phenology ESDR/CDR product suite using three different data sources [AVHRR+MODIS+’VIIRS’]
• Distribute & support these products
• Reach out to the user/science/modeling communities
  – Mature science algorithms
  – Well characterized ESDRs with known error and Uncertainty. The aim here is ‘consistency’ to enable change research
Vegetation Index ESDR

• NDVI & EVI
  – AVHRR missing the blue band
  – Development of a 2-band EVI
  – NDVI and EVI are complementary
• Daily data
  – Compositing turns out to be the source of some problems with current records
  – Retain all useful data
    • Eliminate cloudy/poor data
    • Fill the spatial gaps
Phenology

• The word Phenology is derived from the greek word “Phaino” meaning show or appear (but, as a science it is older than its Greek name).

• In its most generic sense Phenology is the study of the timing of recurring natural phenomena (e.g. seasonal events).
  – It is the study of plant and animal life cycle events, which are triggered by environmental changes, especially temperature. Wide ranges of phenomena are included, from first openings of leaf and flower buds, to insect hatchings and return of birds. Each one gives a ready measure of the environment as viewed by the associated organism. Thus, phenological events are ideal indicators of the impact of local and global changes in weather and climate on the earth’s biosphere.

• Phenology is an integrative environmental science
  – Change results not from a specific event (like snowmelt or temperature increase) but from all controlling factors

• Phenophase: A particular phenolgy event. Example budbreak, unfolding of first leaf, flowering, fruiting, turning of leaves, animal migration, emergence, growth stages, breeding, nesting, hibernation, etc.
Phenology an essential component of the biosphere

- Governed by complex interactions and feedbacks between the biosphere and the climate
  - Climate defines phenology (biome distribution, where and what grows)
  - Phenology impacts the climate (surface characteristics)
    - Albedo, boundary layer
    - Mass balance, Energy balance, etc…
The pulse of our planet, and is an essential and critical component of environmental science influencing biodiversity, species interactions, their ecological functioning, and their effects on fluxes of water, energy, and biogeochemical elements at various scales.

Changes in phenology depict an integrated response to environmental change and provide valuable information for global change research, invasive species management, drought monitoring, wildfire risk assessment, and agricultural production.
Remote Sensing and Phenology

- Spatial extent (regional to global scale depiction of phenology) and temporal frequency of RS platforms permits high sampling rates (needed for studying phenology)
- Allow detailed biophysical inputs into a variety of models (move from on/off parameters to smoother time dependent values)
- Facilitate understanding of phenological phenomena, including causes and effects
- However:
  - Not at the specie level but at the landscape level (section or expanse of a scenery, usually extensive, that can be seen from a single viewpoint)
  - Short period-of-record
  - Cloud cover interference
  - Interpretation issues
  - Lack of a solid foundation
    - Based on what particular measurement (NDVI, EVI, LSWI, NDWI, etc...)
    - Validation is an issue
Phenology ESDR Algorithms

- Current MODIS Veg. Dynamic method (Friedl & Zhang)
  - Piece wise logistic function
  - Fit the data, then look for the inflection points

- Phenoregion/Cluster method (Didan variation based on M. White)
  - Phenology depends on environmental control factors:
    - Climate (Temperature, Precip.)
    - Land cover, Soil
    - Elevation gradient, Aspect, etc...
  - We then create clusters of similar characteristics and apply the Half-Max method

- Combination of both
Project data sources

19 yrs of AVHRR data

Project performance period
Backward compatible EVI2 Algorithm

- From a 3-band blue dependent EVI to a 2-band universal EVI
  - Blue band in EVI provides additional atmosphere resistance and no biophysical information
  - But no blue band for backward compatibility

\[
EVI = 2.5 \frac{N - R}{1 + N + 6R - 7.5B}
\]

Under condition

\[
MAD = \sum_{i=1}^{n} |EVI_i - EVI_{2i}|
\]

- Backward compatible EVI algorithm

\[
EVI_2 = 2.5 \frac{N - R}{1 + N + 2.4R}
\]

This linear relation breaks over noisy data and/or over snow/ice, which are filtered.
ERROR AND UNCERTAINTY
Estimating error

• Error sources in RS data records are quite complex:
  – Due to atmosphere, Geolocation, Calibration, Science Algorithms, etc...

• Error also suggest some reference truth that actually does not exist in this context (scale)

• Atmosphere and Surf. Refl. Algorithms are the sources of most error
  – VI are fully traceable, they do not change the data

• We will focus on consistency to indicates the level of noise in the data records
First step to address the error and uncertainty in these records is through the use of pixel QA/Reliability metric.
Data quality considerations

To avoid biasing the continuity algorithm with poor/noisy data, we:

- Looked at the 30+ yr data record (AVHRR + MODIS)
- Constructed a data reliability Index (usefulness) that indicates when, where and how much useful data these records contain

  - Based on Cloud, Aerosol, Shadow, Snow/Ice, View Angle, etc...
  - In the context of how likely science algorithms can lessen their impacts
Noise related uncertainty envelope

Rain Forest NDVI Error Range

MODIS EVI Range/Error
Boreal Forest

<table>
<thead>
<tr>
<th>Date</th>
<th>EVI/Error</th>
<th>Avg</th>
<th>Min</th>
<th>Max</th>
<th>STD Dev</th>
<th>Median</th>
<th>Max_Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/6/1999</td>
<td>0.054</td>
<td>0.069</td>
<td>0.012</td>
<td>0.017</td>
<td>0.022</td>
<td>0.030</td>
<td>0.043</td>
</tr>
<tr>
<td>4/19/2001</td>
<td>0.069</td>
<td>0.087</td>
<td>0.109</td>
<td>0.158</td>
<td>0.892</td>
<td>0.875</td>
<td>0.845</td>
</tr>
<tr>
<td>9/1/2002</td>
<td>0.054</td>
<td>0.069</td>
<td>0.012</td>
<td>0.017</td>
<td>0.022</td>
<td>0.030</td>
<td>0.043</td>
</tr>
<tr>
<td>1/14/2004</td>
<td>0.069</td>
<td>0.087</td>
<td>0.109</td>
<td>0.158</td>
<td>0.892</td>
<td>0.875</td>
<td>0.845</td>
</tr>
<tr>
<td>5/28/2005</td>
<td>0.054</td>
<td>0.069</td>
<td>0.012</td>
<td>0.017</td>
<td>0.022</td>
<td>0.030</td>
<td>0.043</td>
</tr>
</tbody>
</table>

NDVI Statistical Analysis for the MODIS 5+ year record

<table>
<thead>
<tr>
<th>Measure</th>
<th>Avg</th>
<th>Min</th>
<th>Max</th>
<th>STD Dev</th>
<th>Median</th>
<th>Max_Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humid RF</td>
<td>0.880</td>
<td>0.875</td>
<td>0.109</td>
<td>0.158</td>
<td>0.892</td>
<td>0.875</td>
</tr>
</tbody>
</table>

Global Abs Error 0.0715
Global Rel Error 6.2922 %
Approach to VI error

- The concept of error in VI measurement is not quite accurate
- VI error suggest some reference VI truth that actually does not exist
- VI problems/‘errors’ are exclusively the result of SR errors
- Need a better approach to quantify VI ‘error’
Potential applications

• Replicate previous findings
  – Boreal forest lengthening of the growing season?
  – Tropical forest seasonality
  – Vegetation migration to higher elevation, higher latitude
  – Global drought impact
  – Agriculture response to climate change
  – Recent biomes shifts?
  – Input to SVAT, Biogeochemical, GCM, and other models
Distribution

- Via the USGS-EROS LP-DAAC
- And separately via a PI based data distribution system (focused on value addition).
Project Data and Processing flow
Team organization

- Kamel Didan, UA
- Miura Tomoaki, UH, at Manoa
- Van Leeuwen Willem, UA
- Friedl Mark, BU
- Xioyang Zhang, ERT/NOAA/NESDIS/STAR
- Czapla-Myers Jeff, UA
- Jenkerson Calli, ERT/USGS/EROS/LP-DAAC
- Meyer David, USGS/EROS/LP-DAAC
- Tsend-Ayush Javzandulm, UH, at Manoa
- Hufkens Koen, BU
- Barreto Armando, UA
- Javier Riveracamacho, UA
- Ram Sundaresh, UA
## Project timeline

<table>
<thead>
<tr>
<th>Current Status/Time</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organize</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Website</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input Data Acquisition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coordinate with NPN/Community</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alg. Prototyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UWG &amp; ATBD Review</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESDR Production/Distribution - Prototype</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESDR characterization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESDR Refinement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Workshop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reprocessing/Distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second Workshop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESDR Final Reprocessing and long term Archiving</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conclusions

• New backward compatible EVI2 index
  – No need for blue band (AVHRR, ASTER)
  – EVI2 actually outperforms EVI globally
• Continuity algorithm to facilitate the translation of AVHRR data to MOIS like data
• MODIS Vegetation Dynamic (Phenology) Algorithm adjusted to work with this new data record and AVHRR in particular
• Enable the production of a 30+ year long record
• Suitable for the study of climate change impact on vegetation
• Critical to understanding long term change, trends, anomalies
• Provide input and validation means to modeling efforts, and support policy decisions

This work is supported by a NASA MEaSUREs grant

http://measures.arizona.edu
BACKUP SLIDES
Data and methods

AVHRR record, ‘81 – ‘99

MODIS record, ‘00 – ’13

‘98-’02 SPOT- VGT Data to Bridge AVHRR & MODIS records

Data retention filter.
Only retain high quality data

Homogenous clusters map.
Areas with same characteristics

Continuity Transfer Function map

1981 -2013 sensor independent surface reflectance record
ESDR Products and objectives

- Support the Earth Science research community by providing
  - Long term, multi-sensor, reliable, and consistent Earth System Data Records (ESDR) of
    - Sensor independent EVI (2-band version), NDVI, and
    - A set of Remote Sensing based “Land Surface” Phenology parameters

These ESDRs are expected to be well characterized, of high consistency, and with known error, to support studying climate change & impacts on the Earth system.

First generation (Beta) product suite to be released late 2010 early 2011.