Long Term Vegetation Index Products from Multiple Satellite Data Records Overview and Error/Uncertainty

Didan Kamel\textsuperscript{1}, Barreto Armando\textsuperscript{1}, Miura Tomoaki\textsuperscript{2}, Tsend-Ayush Javzandulm\textsuperscript{2}, and Van Leeuwen Wim\textsuperscript{1}

\textsuperscript{1}The University of Arizona, Tucson, AZ
\textsuperscript{2}University of Hawaii at Manoa, Honolulu, Hawaii

VIP Workshop
The University of Arizona, Tucson, AZ. Jan. 23-24, 2013
Context - 1

- There are three general frameworks to this project/Product suite
  - Vegetation makes up most of the Earth land cover and plays a key role in the biosphere and atmosphere functioning.
  - Programmatically NASA and other agencies invested in different 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, science claims & findings, applications, ease of access/use, and reliability.
Context - 2

• 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
Motivations

• The following 2 statistics should justify monitoring our planet land surface vegetation
  – 57% increase in global population (from 4.45B in 1980 to 7+B in 2012)
  – 15% increase in atm. CO$_2$ (340ppmv in 1980 to 391ppmv as of Oct. 2012)
  – And this is not abating
  – We’re having serious impact on the Earth and its resources that require monitoring
### Recent trends and observations

#### BIOTIC RESPONSE TO RECENT CLIMATE CHANGE

<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>

Data are the key!

- Global long term data are key to understanding, assessing the impact, and predicting trends and changes
  - Change is not a trend unless it persist over a long period.
  - Change is the norm, but a trend is a change towards a new state

- This requires consistent long term data with high accuracy and character. Error is a serious issue especially:
  - When ascribing causality, and
  - Assisting policy making, and
  - For planning

- To produce long term data in support of this research we have limited options:
  - The same sensor that can go on forever (!!!???)
  - Multiple back to back identical sensors (possible but still has issues – new technology, hardware issues, etc...). Can’t impose on other nations
  - At the post processing level (our approach)
RS based Land Biophysical products

<table>
<thead>
<tr>
<th>Energy Balance and Surface Radiation Products Suite</th>
<th>Vegetation Biophysical and Ecology Products Suite</th>
<th>Land Cover/Land Use Products Suite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surf. Reflectance</td>
<td><strong>Vegetation Indices</strong></td>
<td>Land Cover/LC Change</td>
</tr>
<tr>
<td>Land Surface Temperature</td>
<td>LAI/FPAR</td>
<td>LC Dynamic</td>
</tr>
<tr>
<td>BRDF-Albedo</td>
<td>NPP/PSN</td>
<td>Vegetation Cover</td>
</tr>
<tr>
<td>Physical assessment of land-surface processes</td>
<td>Ecosystem functioning &amp; characterization</td>
<td>Conversion / Vegetation</td>
</tr>
<tr>
<td></td>
<td>Seasonal productivity patterns</td>
<td>Continuous Field</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thermal Anomalies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improved Carbon Budget</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Anthropogenic” effects</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sustainable development</td>
</tr>
</tbody>
</table>
Remote Sensing of Vegetation

- Vegetation Indices: Not a physical parameter, but can proxy a long list of physical/biophysical parameters (LAI, fPAR, GPP, LC, Biomass, etc...)
  - NDVI: Strong history and following (70+K pubs, 100K+ of active users, via an assessment of published work and data requests).
  - Plus historical value of AVHRR (30+ yrs).
  - EVI: An improved/enhanced version that requires backward compatibility
    - Strong correlation with biomass/Carbon, structure, less background noise
    - Compatibility - EVI2 (2 bands only)
- 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/Decrease NPP in N America
  - Rain forest seasonality
  - Fire frequencies and Land cover changes
- Can Derive Land Surface Phenology (or growing season characterization). An excellent Integrator of climate impact and barometer of change
Many of the recent advancements of Earth System Science can be attributed to the development and consistency of the AVHRR-NDVI time series, since 1981, in particular to global:

- Biome and agricultural primary production;
- Interannual fluctuations and impacts of ENSO on primary production;
- Phenology and climate change and variability
Vegetation Indices

• NDVI – Widely used but has some issues (ex: Saturation)
  
• 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 has also some issues:
    • Small dynamic range
    • Poor performance over areas with residual cloud, sub-pixel snow, and due to blue band noise (prone to noise)
EVI

- Resistance to background and atmosphere noise and 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.

EVI & Phenology

MaeKlong (SE Asia)

GPP (kg C ha\(^{-1}\) mo\(^{-1}\))

MODIS indices

Regression vs tower GPP

(Huete, et al., GRL (2006))
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 not running away from AVHRR (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 Products/ESDRs
Backward compatible EVI$_2$ Algorithm

Jiang, et al., RSE (2008)

- 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
    \[
    \text{EV} = \frac{2.5}{1 + N + 6R - 7.5B} N - R
    \]
    \[
    \text{Under condition}\quad \text{MAD} = \sum_{i=1}^{n} |\text{EV}_i - \text{EV}_2 i|
    \]
  - Backward compatible EVI algorithm

\[
\text{EV}_2 = 2.5 \frac{N - R}{1 + N + 2.4R}
\]

This linear relation breaks over noisy data and/or over snow/ice, but that poor data are filtered.
What is an EVI2

• Uses Red and NIR, but no blue
• Soil-adjusted (as EVI)
• Improved sensitivity in high biomass regions (as EVI)
• But not atmospherically resistant
  – EVI2 ≈ EVI when atmospheric influences insignificant
  – Should not been a ‘serious drawback’ when data are fully or partially corrected
EVI_2 performance

GLOBAL Histograms

Comparisons per Land Cover Type (3)

North America VI per land cover
good data

VI
- Water
- Evergreen needle
- Evergreen broadleaf
- Deciduous needleleaf
- Deciduous broadleaf
- Mixed forest
- Close shrub
- Open shrub
- Woody savannas

EVI

NDVI

EVI

EVI_2
MEaSUREs ESDRs

• Generate long term ESDR
  – Using multiple data sources/sensors
  – Address sensors differences
  – Processing Differences
  – Algorithm differences
Expectations vs. Data Reality

- Serious problems with RS that are extremely hard and practically impossible to quantify and thoroughly characterize.
- This Error/Noise when translated into quantitative analysis (ex: Physical or Biophysical quantities) becomes very large.
- No longer change or trend: A mix between noise/real change/trend.

Error Characterization Framework

- **Framework**: A set of theories/practices accepted to serve as guiding principles of a research or to address a problem

- In the context of these long term data records: The Framework is defined by a set of metrics for characterizing the error and uncertainty in these ESDRs in order to guide research and application.

- Will focus on spatially and temporally quantifying three features of these ESDRs:
  - **Quality of the atmosphere during the observations acquisition**
    This is a key characteristic, and while current science algorithms are capable of addressing and correcting a host of atmosphere issues (water vapor, ozone, and light aerosols, viewing geometry) their performance is always an issue. In many cases, this correction may even exacerbate the problems.
  
  - **Departure from the normal/expected average**
    This departure although can result from natural factors, is in most cases the result of noise and error in the data. In this context, the departure will be measured by the absolute distance from the long term average. To separate 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 and not error in the data.
  
  - **Temporal profile stability (key to phenology and physical/biophysical modeling)**
    Troughs and Valleys are artefacts that capture and reflect the natural cycle of vegetation dynamic, however they can also result from noise and error in the underlying input data. Random and large deviations are most likely the result of error in the data and cannot result from natural vegetation dynamic. Translated into physical and/or biophysical parameters this could result in large errors (ex: Cumulative VI)
Error Model

• Error related to input (1)
  – Based on accurate field validation (sunphotometers)

• Departure from normal/expected profile (2)
  – Based on long term standard deviations

• Temporal profile stability (3)
  – Based on change about normal/mean
LSR Input related error
Land Surface Reflectance

• \( \text{NDVI} = \frac{(N-R)}{(N+R)} \)

• We’re not concerned with the VI formulation but the error in the input to this equation. In other words, how close we are to the TOC (Atm. Corr. Performance)
  
  - Approach: Use Surface Reflectance validation results & Error analysis over sunphotometers sites (ex: EOS Core sites, LPV). Although, the reported average error is about 2-5% in Red & NIR, we will estimate the impact of a 10% error in Red/NIR on the VI estimation.

Max. VI input related Error
Vegetated = 0.04-0.05 VI Unit (~1-5%)
Non-vegetated = 0.11 VI Unit (~100%)
LSR Input Error Global Propagation

- Apply previous results to Long Term Avg. VI map
LSR Input Error Global Propagation

- EVI Seasonal Error

- Largest error During winter and Over barren, sparsely, or snow/ice covered land
CONSIDERING FILTERED DATA ONLY
Filtering to Reduce Error/Noise

- Filtering + Gap filling to remove useless data
Departure from Normal Related Error

- Departure from normal/expected profile (2)
- Approach is to estimate error by long term standard deviation analysis
  - Error = 0.05 for vegetated and 0.005 for non-vegetated
  - Largest over vegetated area and during winter (snow/ice cover)
Departure from Normal Related Error

- Look at EVI2 (small maps are NDVI errors), separate sensors, and for JAS (Peak growing season). Very similar error across sensors (0.005 – 0.02, max 0.05)
- NDVI Error is larger (0.005-0.05, max 0.08) and more widespread (NDVI saturation).
Stability of Time Series Profiles

- Number of peaks and valleys (3)
  - Assess the number of time the profile crosses the normal/expected profile
  - Impact on cumulative VI (proxy of Carbon sequestration)
  - Showing only Start of Growing Season (AMJ) and Start of Senescence (OND)
  - Insets are standard deviation
  - AMJ: Lowest over vegetated areas and highest over non-veg and high latitude.
  - OND: Lowest over vegetated areas and highest over non-veg.
Impact on Physical/Biophysical parameters (phenology)

- Impact on Cumulative VI (proxy of biomass/carbon)
- The errors in the time series profiles translate into an error in physical/biophysical parameters.
- Impact on Cumulative VI (proxy of biomass/carbon)
  - Insets (Cum. VI, 50-300+ VI units)
- Error = estimated by the standard deviation
- Non-Vegetated: 5 VI units (10%)
- Largest over vegetated areas (25 VI unit, +8%)
Summary of Results

Conclusions

• Backward compatible EVI (2-band EVI or EVI2)
• Sensor independent NDVI & EVI2
• Complex and divergent Error
• Error due to input is rather small (absolute max ~0.05)
  — Surprisingly small impact on VI (Both NDVI and EVI)
  — EVI slightly larger error
• Largest error is due to noise (0.1, 100% over sparsely vegetated areas)
• Impact on physical/biophysical parameters
  — 8-10%
  — The error is largest during winter (cloud and snow, VI dependence on Surf. Refl)
• Error in the stability of the time series profiles is largest over non-vegetated areas
• Similar behavior during start and end of growing season
  — Impact on estimating the exact day of event (10-40 days –analysis not shown here)
• Ability to capture/present error in its proper spatiotemporal context
  — As opposed to one number fits all with little use
• The aim is to aid the 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 then be constrained and their significance established.
Acknowledgments

• Support by NASA-MEASURES

• Many thanks especially to Armando Barreto and Abd-Salam El-Vilaly
It seems like a long time ago

2008-Start
2010-Review
2013-Closure
BACKUP SLIDES
Definitions

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.