Landsat Higher-level Product Plans

March 27, 2014

Pat Scaramuzza
SGT, contractor to the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center

John Dwyer
Project Scientist, Land Satellites Data System
U.S. Geological Survey
Earth Resources Observation & Science (EROS) Center

*Work performed under U.S. Geological Survey contract G10PC00044

U.S. Department of the Interior
U.S. Geological Survey
The Land Satellite Data Systems (LSDS) Science Research and Development (LSRD) project at the U.S. Geological Survey's (USGS) Earth Resources Observation and Science (EROS) Center is prototyping systems and software to generate high-level data products from Landsat 5-7 inputs to support the USGS Terrestrial Monitoring activities with Essential Climate Variables (ECV) and Climate Data Records (CDR). These prototypes are now being modified for L8.

The LSRD has already developed on-demand TOA reflectance, surface reflectance (SR), land surface temperature (LST), and 7 spectral indices for Landsat TM and ETM+ data.

LSRD's EROS Science Processing Architecture (ESPA) prototype utilizes Apache Hadoop to move algorithms across a multi-server cluster to process input data files. ESPA technology facilitates the shortest path to data production by normalizing any software languages and enabling multiple servers to join or disconnect from the cluster without impacting processing.
Outline

- Earth Archive Goals
  - ECVs and CDRs

- Surface Reflectance
  - LEDAPS
  - Spectral Indices

- ECV Products
  - Burned Area
  - Cloud and Cloud Shadow Masks
  - Surface Water Extent
  - Snow Covered Area (SCA)

- Other Upcoming Products
  - Global products
    - Tiled Products?
  - LPVS
Earth Archive Goals

- An archive of calibrated radiances across missions enables development of Climate Data Records (CDRs) and Essential Climate Variables (ECVs)

- **CDRs**
  - Surface Reflectance
  - Surface Temperature
  - Aerosol Optical Thickness

- **ECVs**
  - Burned Area Extent
  - Surface Water Extent
  - Snow Covered Area
  - Global 30m Land Cover
Example of LEDAPS atmospheric correction. Left, top-of-atmosphere reflectance composite (bands 3, 2, 1) for Landsat-7 ETM+ image of the Black Hills and Badlands in South Dakota (September 20, 2002). Right, surface reflectance composite.
LEDAPS Characterization Study

- Landsat Ecosystem Disturbance Adaptive Processing System
- Baseline quality check on provisional USGS EROS Surface Reflectance (SR) products

LEDAPS vs.

- AERONET (Aerosol Optical Thickness)
- Field Spectrometer (Surface Reflectance)
- MODIS (ECVs)

Confirmed expected strengths and weaknesses in LEDAPS

- SR is good in vegetated areas
- Increased temporal variability in shorter wavelengths (likely due to aerosol influence)
- AOT overestimation in certain areas

Landsat Spectral Indices

- Spectral Indices derived from LEDAPS Surface Reflectance CDR
  - NDVI Normalized Difference Vegetation Index
  - EVI Enhanced Vegetation Index
  - SAVI Soil Adjusted Vegetation Index
  - MSAVI Modified Soil Adjusted Vegetation Index
  - NDMI Normalized Difference Moisture Index
  - NBR Normalized Burn Ratio
  - NBR2 Normalized Burn Ratio 2

http://code.google.com/p/spectral-indices/
https://espa.cr.usgs.gov
LEDAPS NDVI vs. MODIS NDVI

- Surface Reflectance-derived NDVI
- Three years
- EOS Validation Sites
  - Mead
  - Bondville
  - Ft. Peck
- General linearity between data sets

![Graph showing the relationship between LEDAPS and MODIS NDVI](image)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>NIR</th>
<th>Red</th>
<th>NDVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²</td>
<td>0.7165</td>
<td>0.5458</td>
<td>0.8238</td>
</tr>
<tr>
<td>RMSD</td>
<td>0.0572</td>
<td>0.0542</td>
<td>0.0866</td>
</tr>
<tr>
<td>Slope (m)</td>
<td>0.9412</td>
<td>0.6368</td>
<td>0.8467</td>
</tr>
<tr>
<td>Intercept (b)</td>
<td>0.0125</td>
<td>0.0433</td>
<td>0.0475</td>
</tr>
<tr>
<td>Points (n)</td>
<td>64</td>
<td>64</td>
<td>64</td>
</tr>
</tbody>
</table>
Burned Area Essential Climate Variable
Algorithm Status

Landsat Path/Row Scenes selected for algorithm training and Stage 2 validation

Algorithm Processing Zones
- Alaska
- Hawaii
- Mountain West
- Arid West
- Western Great Plains
- Eastern Great Plains
- East
Burned Area Essential Climate Variable
Characterization of product uncertainties/validation

p028r033 MTBS & BAECV

Total Burned Area per year
Burned Area within grid (950, 45, 750)

From 1985 to 2010
Total of 523 MTBS polygons
Total MTBS: 6300.95 Km2
Total BAECV: 8848.82 Km2
Agreed: 4970.05 Km2

Commission and Omission error and Accuracy Rate
Burned Area within grid (950, 45, 750)

For all years
Accuracy: 80.72%
Commission Error: 21.12%
Omission Error: 43.84%

Mon, Sep. 16, 2013 (4:18:32 PM) MDT
Cloud and Cloud Shadow Detection

- **ACCA**
  - Spectral threshold algorithm used on Landsat since 1999.

- **See5 CCA**
  - Designed for use on LDCM, limited applicability elsewhere.
  - Does not require thermal band.

- **LEDAPS SR CCA**
  - ACCA based algorithm that uses LEDAPS SR and surface temperature.

- **CFmask** –
  - Port of Function of Mask (Fmask) from Boston University into C.
  - Performs cloud shadow masking.
Cloud and Cloud Shadow Detection
## Cloud Cover Validation Results

<table>
<thead>
<tr>
<th>Cloud Algorithm</th>
<th>Total Pixels Correct (ETM+ Global Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCA</td>
<td>79.90%</td>
</tr>
<tr>
<td>LEDAPS SR CCA</td>
<td>85.70%</td>
</tr>
<tr>
<td>See5 CCA</td>
<td>88.50%</td>
</tr>
<tr>
<td>CFmask</td>
<td>90.97%</td>
</tr>
</tbody>
</table>

**Fmask:**

**ACCA (mask version) and See5 CCA:**
Landsat Surface Water Extent ECV

Uncertainty evaluation chip locations for SWE ECV pilot research.
Landsat Surface Water Extent ECV

Legend
- 0: None
- 1-79158: Not water
- 79159-1010: Waterbody
- 1011-40356: Reservoir
- 40357-213663: Dam
- 213664-112955: ScaryChop
- 112956-122857: ShittyPark
- 122858-2497: Ripley's BelieveIt

The SWE Juvv1 output (0 - not water, 10 - water) and the NHD data (6-3) were added together to create an output comparing the two inputs (Values 0, 1, 2, 3, 10, 11, 12, 13).

Chart showing comparison values (Juvv1 and NHD data).
Snow Covered Area

• Spectral unmixing approach is used to estimate fraction of Landsat pixel covered by snow, vegetation, and soil.

• Algorithm originally developed for hyperspectral instruments, then adapted for use with multispectral MODIS data (MODSCAG).

• Most recently adapted for use with Landsat TM and ETM+ instruments (TMSCAG).

Snow Covered Area

- Landsat Scene-based fSCA Validation
  - High resolution imagery
  - In situ sensor arrays
    - 900+ sensors, 45 sites, 5 states
    - 13 sites above alpine treeline (Colorado), 5 sites in the Arctic
    - 3 sites with 49 sensors per 60 m² to test optimum sensor density

- Product Implementation plan
  - Phase 1: Landsat scene-based viewable fraction of Snow Covered Area (fSCA) product
  - Phase 2: Landsat-MODIS daily 30 m viewable fSCA
  - Phase 3: Landsat scene-based and Landsat-MODIS daily 30 m canopy-adjusted fSCA
SCA – *In Situ* Sensor Arrays

- Buried data loggers recording hourly temperature indicate presence/absence of snow cover
- Arrays allow for monitoring daily *in situ* fSCA for pixel footprints
- Adjacent sensor arrays allow for comparison between fSCA in forests/meadows or different forest densities
Upcoming work

- Global Datasets
  - Historical
  - Current
  - Future
  - Annual Products
  - Mid-Decadal

- Tiled Products
  - Web Enabled Landsat Data (WELD)
Landsat Global 30m Land Cover
Historical, Current, and Future

Annual products –
Land cover continuous variables

• Quantitative annual continuous measures per pixel percent tree, shrub, herbaceous, water, snow/ice, and barren cover.
• Change products

Mid-decadal products – Land cover types

• Land cover categories consistent with FAO Land Cover Classification System (LCCS)
• Maps and statistical estimates of major land cover types
• Complementary with other global land cover products (e.g., MODIS land cover, Globecover)
Global “Cloud-Free” Landsat
Web Enabled Landsat Data (WELD)

- **ETM+ Composites**
  - L-1T mosaics
  - CONUS and Alaska
  - TOA Reflectance
  - Weekly, Monthly, Seasonal, Annual
  - 2003-2012
  - Albers Equal Area

- **Distributed as**
  - Tiles (5,000 x 5,000 pixels)
  - User-defined Area of Interest

- **Illustrates a general goal:** Retire the WRS path/row system
Land Product Validation System (LPVS)

- Allows users to order and analyze co-registered data for multiple satellites, for characterization and cross-validation of land products.
- Currently in beta, next phase in September 2014
- User feedback is encouraged

http://landsat.usgs.gov/LPVS.php
Land Product Validation System (LPVS)

- Planned Data sets include CDRs and ECVs from:
  - Landsat (L4 and L5 TM, L7 ETM+, L8 OLI/TIRS)
  - MODIS
  - GOES-R Advanced Baseline Imager
  - ESA Sentinel series
  - JPSS
    - Suomi NPP
    - JPSS1
  - ...and others

http://lpvsexplorer.cr.usgs.gov/
End of presentation
(Supporting Slides after this)
LEDAPS AOT vs. AERONET

- Nationally distributed sites
- Relationships evaluated by difference classes to remove bias

<table>
<thead>
<tr>
<th>Color</th>
<th>Site Class</th>
<th>Median AOT Difference</th>
<th>Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>&lt; -0.10</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-0.10 - 0.10</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.11 - 0.20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.21 - 0.30</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>&gt; 0.30</td>
<td>7</td>
</tr>
</tbody>
</table>
LEDAPS AOT vs. AERONET

- 95 sites
- 3,471 Observations
- Correlation: 0.63 – 0.87
- Bias: -0.42 – 0.25
- RMSD: 0.29 – 0.65
Temporal Stability Analysis

- South Dakota State University
- CEOS Pseudo Invariant Calibration Sites
  - Algodones Dunes
  - Libya 4
- More variability in shorter wavelengths with most aerosol influence
Burned Area Essential Climate Variable

Product specification

Individual layers:

• Burn probability (0-100) and burn classification (0 or 1) layers with QA masks

Annual summary layers:

• Burned probability and classification layers summarized for a calendar year (Jan-Dec)
• Maximum probability, first date of observed burn, number of observed burns, number of non-cloudy observations
Burned Area Essential Climate Variable

Product specification

Single-Date Image (e.g. derived from 1 landsat scene)
2 bands (8 bit unsigned integers)
   Band 1, categorical:
      -8 = ‘fill’ mask
      -7 = DDV mask (no longer used)
      -6 = cloud mask
      -5 = adjacent cloud mask
      -4 = cloud shadow mask
      -3 = snow/ice mask
      -2 = water mask
      0 = not burned
      1 = burned

Band 2 = burn probability from classification algorithm (0-100) and QA masks (negative numbers)
Forest vs Meadow, Forest Density
fSCA Comparisons

- Data from spring 2013 indicated longer snow persistence in forests than adjacent meadows (4-21 days)
- Additional sites for water year 2014 will compare meadows vs dense forest vs sparse forest
- This data will aid in adjusting fSCA under canopies using data from nearby open areas
SCA – High Res Imagery

- Multispectral and panchromatic imagery classified to produce binary SCA
- Binary SCA then aggregated to 30 m fSCA and compared to Landsat-derived fSCA
- Approach works well in areas without forest cover
- Much more feasible and efficient using Digital Globe NGA archive
Glossary

CDR – Climate Data Record
A time series of measurements with sufficient length, consistency, and continuity to identify climate variability and changes.

SR – Surface Reflectance

ECV – Essential Climate Variable
Products that are based on CDR inputs, are technically and economically feasible for systematic observation, and that create an authoritative basis for regional to continental scale identification of historical change, monitoring current conditions, and helping to predict future scenarios.

AOT – Aerosol Optical Thickness
SCA – Snow Covered Area
fSCA – Forested Snow Covered Area
Glossary (continued)

Spectral Index –

A product derived from CDRs that is useful but does not meet the criteria for ECVs.

NDVI – Normalized Difference Vegetation Index
EVI – Enhanced Vegetation Index
SAVI – Soil Adjusted Vegetation Index
MSAVI – Modified Soil Adjusted Vegetation Index
NDMI – Normalized Difference Moisture Index
NBR – Normalized Burn Ratio
NBR2 – Normalized Burn Ratio, second formulation
Glossary (continued)

LPVS –
Landsat Product Validation System

LEDAPS –
Landsat Ecosystem Disturbance Adaptive Processing System

ACCA –
Automated Cloud Cover Assessment algorithm

CFMask –
C code version of Function of Mask, a cloud and cloud shadow assessment algorithm.