Analysis Ready Satellite Data Access

Jonathan Morton1, Brian Davis2, Duncan McGregor2, Steve Foga2, Brian Sauer2, John Dwyer2

1Element #4, contractor to USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls
2USGS, Reston, Virginia

Abstract
The Landsat satellite missions have systematically acquired multiplex imagery over Earth’s surface for over 45 years, amassing a temporally dense archive of data that are used in numerous scientific studies involving the monitoring, assessment, and projection of land change. The U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, while continuing to operate Landsat data collection, archive, and distribution, has initiated development of an advanced capability that will efficiently deliver user-specified information derivatives that will transform the availability of lower-level data into analysis ready products for use in mapping and modeling applications. Using advanced processing frameworks and Application Programming Interfaces (API), Landsat scenes are deconstructed and stored as pixels in a data cube. From there, science, calibration, georeferenced, and quality-masked areas of interest, co-registered temporal layer stacks, temporal or band composites, and vectors of pixels values for specific point locations drilled down through data layers can be easily extracted. The need to perform time, network, and disk consuming pre-analyses data manipulations is ameliorated by the abstraction of traditional World Reference System-2 (WRS-2) scenes into parcels of information that can be filtered for quality conditions and readily packaged to user specifications for format, map projection, band selection, spatial and temporal extent. This advanced information access methodology is currently evolving through a prototypical phase and is expected to achieve an initial operating capability on the contiguous U.S. by November 2017.

Objectives
A) Explain how Analysis Ready Data (ARD) will improve the end-user interaction with, and analysis of, the data products.
B) Demonstrate how ARD will be used in operational/science environments.

ARD Data Structure

ARD Benefits
- On-line availability
- Spatially bounded
- Incorporates path/row overlap
- QA filtering
- Enables user-driven analysis
- Time series analysis
- Enables change detection and monitoring

ARD Grid Projection Parameter Definition

<table>
<thead>
<tr>
<th>Projection</th>
<th>Elliptical Arc Equidistant</th>
<th>Datum: North American Datum 1983 (NAD83)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conformal U.S.</td>
<td>Alaska</td>
<td>Hawaii</td>
</tr>
<tr>
<td>First standard parallel</td>
<td>29.5°</td>
<td>65.0°</td>
</tr>
<tr>
<td>Second standard parallel</td>
<td>45.0°</td>
<td>65.0°</td>
</tr>
<tr>
<td>Latitude of projection</td>
<td>23.0°</td>
<td>50.0°</td>
</tr>
<tr>
<td>False Easting</td>
<td>154.0°</td>
<td>154.0°</td>
</tr>
<tr>
<td>False Northing</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Status Quo

Landsat satellites constantly acquire image intervals among their orbit, and the intervals are split into scenes organized by path and row (above) to make the data reasonably sized for end-user consumption. Currently, scenes are ordered one at a time or in bulk, and are downloaded, stored, extracted, combined and analyzed by the end user. Since the conversion of the Landsat data policy to no-cost access in 2008, demand has transitioned away from being contained to a single scene to time series analysis (Wulder et al., 2012). With over 30 years of directly comparable (geometrically and radiometrically aligned) Landsat data, improved analysis of climate and land use interactions can be derived using this data record.

However, a stack of rectified-corrected scenes from associated Quality Assurance (QA) data for a single path/row footprint could be over 700 acquisitions, which equates to ~300 GB (compressed). End users are also restricted to the boundaries of a single scene unless mosaicking is performed, further burdening the individual user.

ARD Grid Structure

ARF Data Structure

Source products for ARD

Source sensors for ARD and respective per-scene data volumes (approx.)

ARD Grid Projection Parameter Definition

Volume for ARD stacks in select regions (approx.)

Landsat scenes into parcels of information that can be filtered for quality conditions and readily packaged to user specifications for format, map projection, band selection, spatial and temporal extent. This advanced information access methodology is currently evolving through a prototypical phase and is expected to achieve an initial operating capability on the contiguous U.S. by November 2017.

Discussion and Conclusions
- ARD will make Landsat data easier to access and analyze than scene products.
- The scene-based Landsat data products will persist for the foreseeable future.
- ARD will enable capabilities like LCMAP and CCDC to efficiently process data stacks.
- Landsat MSS data (1972-1982) will eventually be included.
- New data acquisitions will continuously be added to the ARD.

References

Acknowledgements
Work performed under U.S. Geological Survey contract G15PC00012.

More Information

Abbreviations
- ARF: Application Programming Interfaces
- ARD: Analysis Ready Data
- AVHRR: Advanced Very High Resolution Radiometer
- CCDC: Continuous Change Detection and Classification
- WRS: World Reference System

U.S. Department of the Interior
U.S. Geological Survey
September 2016