Analysis Ready Satellite Data Access

Jonathan Morton[1], Brian Davis[2], Duncan McGreggor[1], Steve Foga[2], Brian Sauer[3], John Dwyer[3]

[1]Element 84, contractor to USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198, USA; [2]Stinger Ghaffarian Technologies, contractor to USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198, USA.; [3]USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198, USA.

Abstract

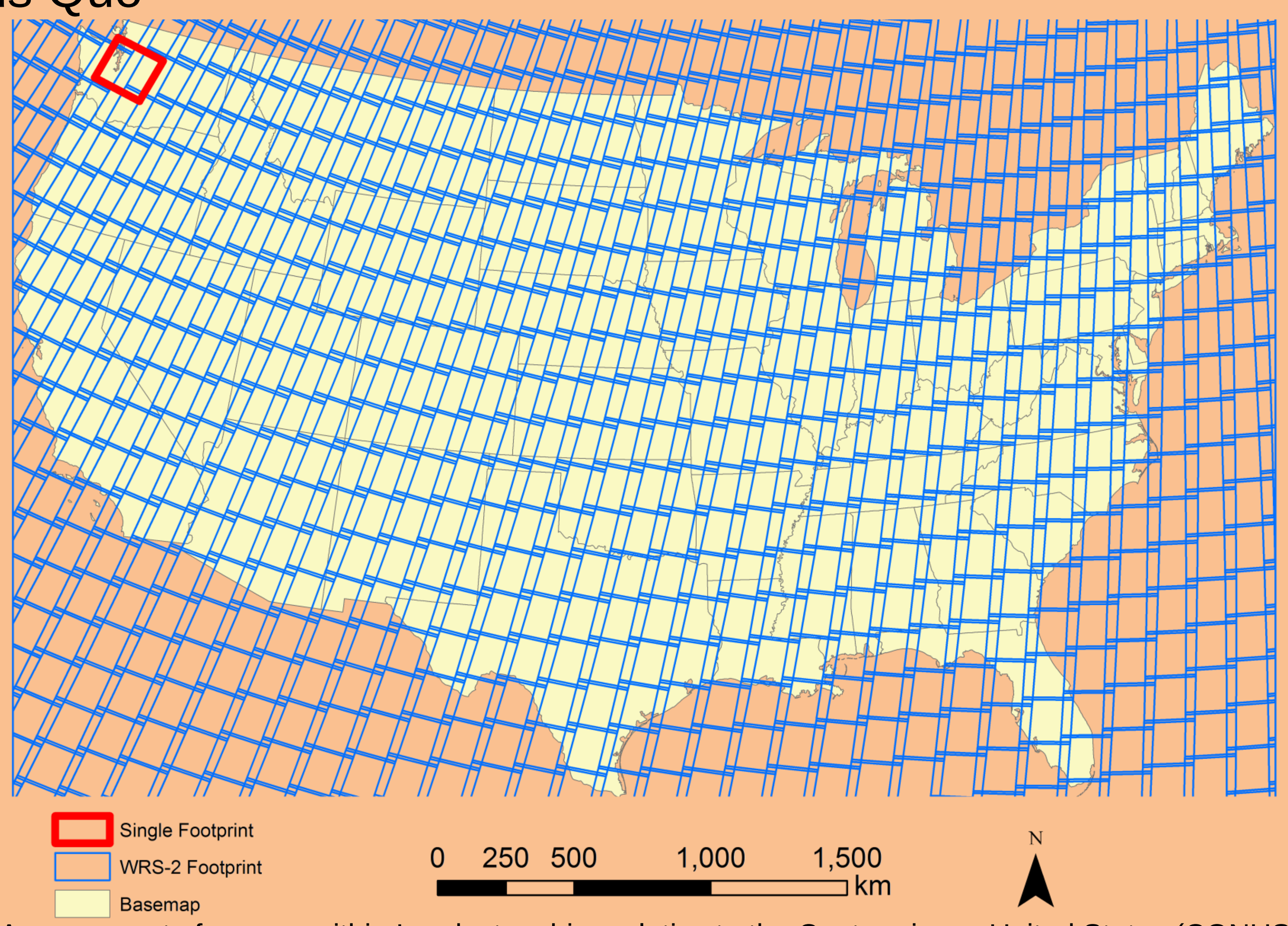
The Landsat satellite missions have systematically acquired multispectral imagery over Earth's surface for over 40 years, amassing a temporally dense archive of data that are used in numerous scientific studies involving the monitoring, assessment, and projection of land change. The U.S. Geological Survey (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 Applications Programming Interfaces (API), Landsat scenes are deconstructed and stored as pixels in a data cube. From there, seamless, calibrated, georeferenced, and quality-masked areas of interest, co-registered temporal layer stacks, temporal or band composites, and vectors of pixel values for specific point locations drilled down through data layers can be easily extracted. The need to perform time, network, and disk consuming pre-analysis data manipulations is ameliorated by the abstraction of traditional Worldwide 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 over the conterminous 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.

Status Quo



Arrangement of scenes within Landsat archive relative to the Conterminous United States (CONUS).

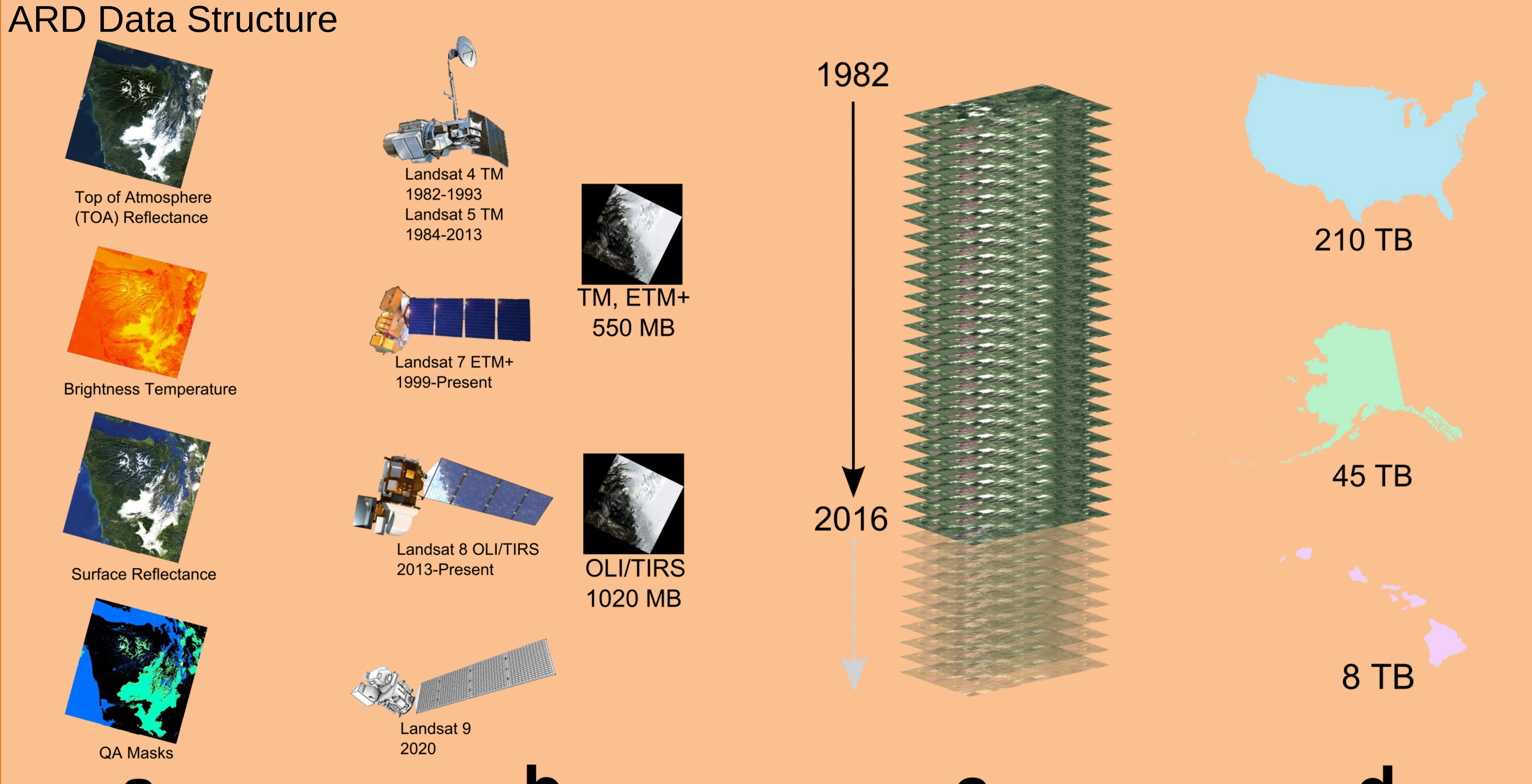
Landsat satellites constantly acquire image intervals along their orbit, and the intervals are split into *scenes* organized by *path* and *row* (above) to make the data reasonably sized for end-user consumption. Currently, scenes are ordered one at a time or in bulk, and are downloaded, stored, extracted, combined and analyzed by the end user.

Since the conversion of the Landsat data policy to no-cost access in 2008, demand has transitioned analysis 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 reflectance-corrected scenes and 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 Data Structure



a Source products for ARD.

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

c Products (**a**) from all sensors (**b**) are gridded and stacked to form ARD.

d Volume for ARD stacks in select regions (approx.)

ARD Benefits

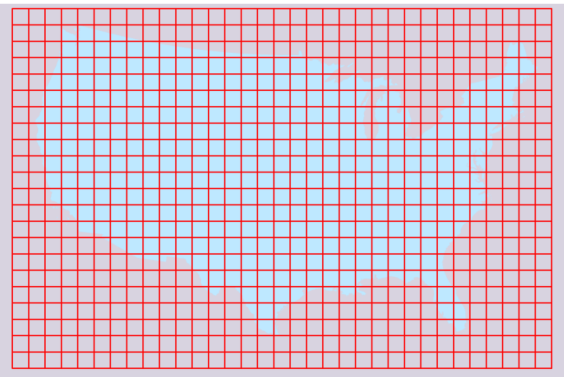
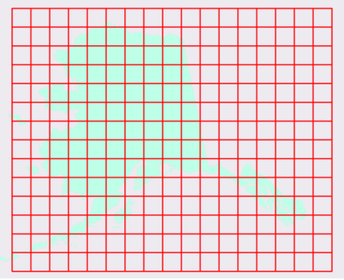
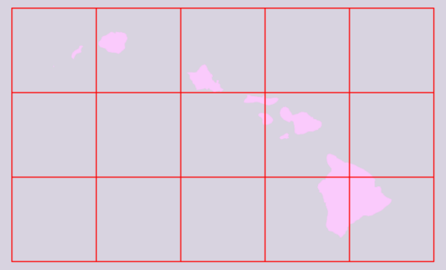
- ✓ On-line availability
- ✓ Spatially boundless
- ✓ Incorporates path/row overlap
- ✓ QA filtering
- ✓ Enables user-driven analysis
- ✓ Time series analysis
- ✓ Enables change detection and monitoring

ARD Grid Projection Parameter Definition

Projection: Albers Equal Area Conic			
Datum: North American Datum 1983 (NAD83)			
	Conterminous U.S.	Alaska	Hawaii
First standard parallel	29.5˚	55.0˚	8.0˚
Second standard parallel	45.5˚	65.0˚	18.0˚
Longitude of central meridian	-96.0˚	-154.0˚	-157.0˚
Latitude of projection origin	23.0˚	50.0˚	3.0˚
False Easting	0.0	0.0	0.0
False Northing	0.0	0.0	0.0

Internal ARD Tiling Scheme

Systematic approach to ingesting, converting, archiving and retrieving data in non a scene-based framework. Ultimately, the tiling scheme will be transparent to end users, and data could be queried at any scale.

Region	ULX	ULY	LRX	LRY	H	V	Graphical Representation
CONUS	-2565585	3314805	2234415	164805	32	31	
AK	-851715	2474325	1698285	374325	16	13	
HI	-444345	2168895	305655	1718895	4	2	

Spatial extents for each grid in meters, and the tile count in horizontal (H) and vertical (V) dimensions. Each tile is 5000 by 5000 30-meter pixels.

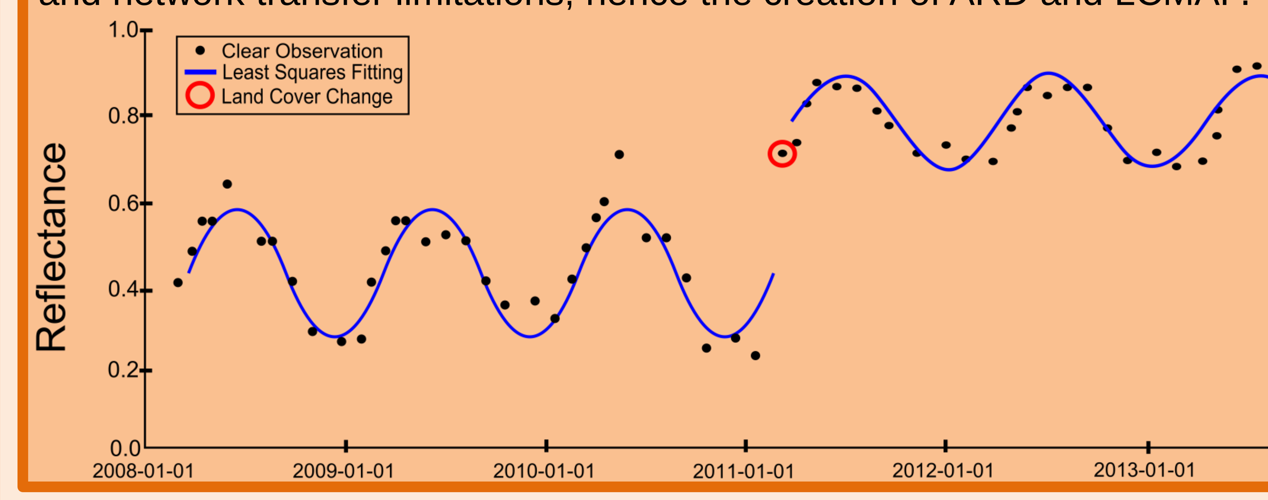
LCMAP

The Land Change Monitoring Assessment and Projection (LCMAP) information system will allow interactive access to the ARD for scientific analysis. Additionally, LCMAP will enable the execution of science algorithms using ARD. LCMAP will be broken into the following components:

- **Data storage** – on-line, scalable storage for ARD and other system components.
- **Data access and exploration** – interactive ARD access to query, filter, spatially and spectrally modify and/or analyze data desired by the end user via an API.
- **Execute science algorithms** – built-in models and algorithms can be run on a continuous time series of data. Many of these results will be pre-computed and available for user analysis.
- **Subscription and notification services** – notification of system updates, added products, product modifications, model completion or when algorithm threshold(s) are met as data are added.
- **Client libraries, documentation, tutorials** – instructions on how to utilize the above functions. Client libraries will initially support language bindings for Python, Clojure, JavaScript, and system shells (e.g. Bash with cURL). R, Julia, and Ruby bindings are planned for a future release.

CCDC

The Continuous Change Detection and Classification (CCDC) algorithm is a major science component in LCMAP. CCDC trends change in pixel attributes over time to determine whether or not land cover has been significantly altered. This algorithm will be run as new data are acquired to allow near-real time assessment of land change. This algorithm requires a dense time series of pixel values which are geometrically co-registered and radiometrically calibrated, and retrievable without cumbersome file-system and network transfer limitations, hence the creation of ARD and LCMAP.



Simulated observations and trending results from the CCDC algorithm. Based upon work by *Zhu et al., 2012.*

Discussion and Conclusions

- ARD will make Landsat data easier to access and analyze than scene products.
- The scene-based Landsat data products will exist 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.

Abbreviations

API	Application Programming Interface
ARD	Analysis Ready Data
AK	Alaska
CCDC	Continuous Change Detection and Classification
CONUS	Conterminous United States
ETM+	Enhanced Thematic Mapper Plus
H	Horizontal tile
HI	Hawaii
LCMAP	Land Change Monitoring Assessment and Projection
LRX	Lower-Right X coordinate
LRY	Lower-Right Y coordinate
MB	Megabyte
OLI	Operational Land Imager
QA	Quality Assurance
TB	Terabyte
TIRS	Thermal Infrared Sensor
TM	Thematic Mapper
ULX	Upper-Left X coordinate
ULY	Upper-Left Y coordinate
V	Vertical tile
WRS	Worldwide Reference System

More Information

- LCMAP Client Documentation: http://usgs-eros.github.io/lcmap-client-docs/current/
- Landsat Collections Definitions: http://landsat.usgs.gov/landsatcollections.php

References

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., & Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. Remote Sensing of Environment, 122, 2-10.

Zhu, Z., Woodcock, C. E., & Olofsson, P. (2012). Continuous monitoring of forest disturbance using all available Landsat imagery. Remote Sensing of Environment, 122, 75-91.

Acknowledgements

Worked performed under U.S. Geological Survey contract G15PC00012.