g

Satellite Data Interoperability Workshop ARD[1] and STAC[2] tracks

Technical Program

August 13-15, 2018
Rambo Auditorium
Earthquake Science Center, USGS
345 Middlefield Rd, Menlo Park, CA 94025

Program Chairs

Ignacio Zuleta
Chris Holmes

Convened by Radiant.Earth

Convening Sponsorship by Maxar and Planet Labs

Sponsors

PCI Geomatics
The Climate Corporation
CosmiQ Works
Astraea

Supporters

Azavea
Vulcan
cosine measurement systems

[1] ARD = Analytics-ready-data
[2] STAC = SpatioTemporal Asset Catalogs

TABLE OF CONTENTS

INTRODUCTION

The Satellite Data Interoperability Workshop aims to advance interoperability and collaboration in the earth observation and remote sensing industry. It builds on the success of the SpatioTemporal Asset Catalog (STAC) sprints, expanding the scope of collaboration beyond just API's and data formats, to include interoperability at the data level. There is a growing trend of Analysis Ready Data, and so with this event we convene a group of experts to start to set down standards for cross-provider ARD.

GOALS

The goal of this workshop is to:

- Advanced the SpatioTemporal Asset Catalog specification for greater stability and more use cases
- Coordinate and survey trends on
  - Data formatting for Analysis Ready Data
  - Contribution and hosting of benchmark ARD datasets
  - Definition of reference implementations of every step of harmonization for ARDs to be implemented by vendors and data providers
- Move beyond discussion to start to establish standards of interoperability for Analysis Ready Data
- Survey new technology, payloads and data sources that support the creation of ARD

Meals and evenings

- Boxed lunch provided all three days at **noon**
- Dinner at **6pm**:
  - **Monday: Nola's,** 535 Ramona St, Palo Alto, CA 94301
  - **Tuesday: Left Bank,** 635 Santa Cruz Ave, Menlo Park, CA 94025
- Coffee and **light breakfast** at **8:30am**
- ~~Optional **Bay Area Space Happy Hour at Planet HQ** on Wednesday evening.~~ **Full, registration closed.**

Monday joint activities

Monday August 13

lunch		
1:00-1:05	**Welcome, Introduction by the co-chairs**	Ignacio Zuleta, *Planet* Chris Holmes, *Radiant.Earth*
1:05-1:20	**Radiant.Earth and Satellite Data Interoperability + Sponsor Welcome**	Hamed Alemohammad, *Radiant.Earth*
1:20-1:40	**ARD Opening Remarks**	Ignacio Zuleta, *Planet*
1:40-2:00	**The State of STAC**	Chris Holmes, *Radiant.Earth*
meet and greet		
3:00-3:25	**Large volume production system to prepare ARD data and integration with the Open Data Cube – Initial lessons learned from deforestation monitoring case study in Canada**	Guillaume Morin, *PCI Geomatics*
3:25-3:50	**Why is ARD + API based access essential for EO?**	Leo Lymburner, Geoscience Australia
3:50-4:15	**ARD for Agricultural Remote Sensing: Challenges and Use Cases**	Keely Roth, *The Climate Corporation*
4:15-4:40	**ARD at Digital Globe**	Chris Helm, *Digital Globe*
4:40-5:05	**HyperScout In-Orbit Demo and future plans**	Marco Esposito, *COSINE*
5:05-5:30	**Open Source Applied Research: SpaceNet and Machine Learning for Remote Sensing Data Sets**	David Lindenbaum, In-Q-Tel

STAC Track Schedule

Tuesday August 14

Time	Title	Description
9:00 - 11:30	**Implementation Talks**	Prepared presentations by those with STAC implementations
	STAC API's	Harris, sat-api by DevSeed, NASA CMR Proxy by Element 84, Hexagon
	Static STAC	CBERS, Spacenet, Google Earth Engine, DigitalGlobe, Planet, Azavea / Radiant.Earth
	Clients	STAC Browser
10:20 - 10:40	**Break**	
11:30 - 2:30	**Small Group Sessions**	Groups are formed to enable smaller discussions. Those with experience implementing the spec will be grouped to resolve specification issues, with github tickets assigned. Those who are newer to the spec can sprint on their own implementation or join an existing tool effort. Check Radiant.Earth Community Sprint Repo on Github for details.
12:00 - 1:00	**Lunch**	Ideally small groups continue through lunch, or can break and network with other STAC and ARD people
2:30 - 3:00	**Full Group Check-in**	Check on status, report on progress, resolve any burning questions with the full group
3:00 - 6:00	**Small Group Sessions (continued)**	Keep working in small groups, updating stac-spec repo
3:30 - 5:00	**Open STAC Session** **OPEN TO ALL!**	**OPEN TO ALL ARD ATTENDEES** Deep Dive on STAC overview, Q&A on how to use STAC, and show and tell of leading STAC implementations. A small group of STAC sprinters will present in an alternate room that anyone from the ARD tracks is welcome to attend and learn more. Presented by Chris Holmes, Michael Smith, Frederico Liporace, Seth Fitzsimmons and ?

Wednesday August 15

Time	Title	Description
9:00 - 11:00	**Full Group Spec Discussion**	Review Pull Requests and their related discussions from the previous day, resolve differences and assign leads to move each aspect forward.
11:00 - 2:30	**Small Group Sessions**	Break in to small groups as makes sense from full group discussions, to move forward on writing up and making Pull Requests on the issues discussed.
12:30 - 1:30	**Lunch**	Ideally small groups continue through lunch, or can break and network with other STAC and ARD people
2:30 - 4:00	**End With Delivery**	Core group shifts to direct specification work – reviewing and merging pull requests, updating the spec copy, preparing for release, etc. Small group sessions on implementations continue, to be able to demo.
4:00 - 5:00	**Demo Session**	**OPEN TO ALL ARD ATTENDEES** Everyone shows what they worked on during the sprint. Close with commitments on what people will continue to work on after.
6:00	**Happy Hour at Planet**	**IN SAN FRANCISCO.** Optional happy hour, is ~1 hour away from Menlo Park. Be sure to register at https://www.eventbrite.com/e/space-in-the-bay-at-planet-tickets-48450732445

ARD Track Schedule

Tuesday August 14

9:00-9:20	**Combining HR and VHR imagery for EO-based land monitoring solutions**	Michael Riffler et al, GeoVille GmbH
9:20-9:40	**VITO's Framework for Analysis Ready Data**	Marcel Buchhorn et al, VITO
9:40-10:00	**Ingesting Astro Digital Landmapper-BC data into the analytics-ready dataset**	Alex Kudriashova, Astro Digital, Inc.
10:00-10:20	**High Quality Analysis Ready Data Starts with Mission Design**	William Parkinson, UrtheCast, Inc.

coffee break

10:40-11:00	**On-the-fly ARD in Practice using Sentinel Hub**	Miha Kadunc, SINERGISE
11:00-11:20	**Planet Data in Context**	Robert Simmon, Planet
11:20-11:40	**Google Earth Engine and ARD**	Tyler Eriksson, Google
11:40-12:00	**Starlight's InsightGIS: Tools for GEOINT Platforms with Analysts as the First Focus**	David Bernat, Starlight

lunch break

1:00-2:15	**Harmonized Landsat/Sentinel-2 Products for Land Monitoring**	Junchang Ju, NASA
	TBD	TBD
	Perspectives from EOFSAC: A NASA Program for Food Security and Agriculture	Michael Humber, University of Maryland
2:15-3:30	**Analysis Ready Data as addressed by the Committee on Earth Observation Satellites (CEOS ARD 4 Land)**	Steven Hosford, ESA
	ARD and Exploitation Platforms	George Percivall, OGC
	Benchmarking datasets and public data access	Ignacio Zuleta (panel host)
3:30-4:45	**Satellogic: processing and ARD challenges**	Slava Kerner, Satellogic
	Importance of analysis-ready data for satellite remote sensing applications: a practical discussion	Sangzi Liang, Corteva Agriscience™, Agriculture Division of DowDuPont
	Sentinel-2 ARD - toward multi-mission interoperable/ harmonised products	Ferran Gascon, ESA
4:45-5:30	**panel discussion | breakout sessions**	

Wednesday August 15

9:00-10:15	**Cloud Detection for the Planetscope Constellation**	Keith Beckett et al., Planet Labs, Inc.
	TBD	TBD
	Surface Reflectance Comparison between OLI and Planet imagery	Minsu Kim, USGS
10:15-11:05	**Validating PCI ARD data with Sen2Cor and LaSRC including BRDF correction and Topographic Normalization**	Wolfgang Lueck, PCI Geomatics
	Analogue-based colorization of remote sensing images using textural information	Mathieu Gravey, Universite of Lausanne
11:05-11:55	**Calibrating PlanetScope using external data sources to produce a cloud-/gap-free daily surface reflectance data**	Sibo Wang, University of Illinois at Urbana-Champaign
	Combining Multiple Data Sources in time series analysis for monitoring land change	Curtis Woodcock, Boston University

lunch break

1:00-2:15	**Data Fusion for Monitoring Rainforest Carbon: A case for global co-registration**	Pramukta Kumar, Carnegie Institution for Science
	Interoperability between Planetscope, RapidEye and MODIS TOAR products	Arin Jumpasut, Planet Labs, Inc.
	Multi-sensor data fusion for enhanced land surface characterization in the era of the CubeSat	Rasmus Houborg, SDSU
2:15-3:30	**Harmonizing Different TOAR and SR Products using Simple Regression Models**	Gina Li, Planet Labs, Inc.
	Operational Biophysical Analysis Ready Data for Landscape Monitoring	Peter Scarth, University of Queensland
	Creating Interoperable Synthetic Aperture Radar Data at the ASF DAAC	Scott Arko, University of Alaska
3:30-4:40	**Global 30-m Landsat-derived Cropland Extent Product for Nominal 2015: produced using Machine Learning Algorithms and Cloud Computing on Google Earth Engine**	Prasad S. Thenkabail, USGS
	PanArctic Shipping: Using satellite Automated Identification System (AIS) data to aid informed decision-making in the Arctic	Greg Fiske, Woods Hole Research Center
4:40-5:00	**next steps discussion**	
6:00	**Happy Hour at Planet**	**IN SAN FRANCISCO.** Optional happy hour, is ~1 hour away from Menlo Park. Be sure to register at https://www.eventbrite.com/e/space-in-the-bay-at-planet-tickets-48450732445

TALK ABSTRACTS

The technical program consists of series of keynote talks, panels and working sessions (standards, benchmark datasets, tutorials). These are the talks submitted as of August 8th.

ARD+STAC INTEROPERABILITY TRACK — LIST OF TALKS		
Title	Contributors	Abstract

ARD at Digital Globe
Chris Helm

Digital Globe
DigitalGlobe's data access project, RDA (Raster Data Access), is a multi-source platform capable of serving DG products as well as other sensors, like Landsat8 and Sentinel2. RDA employs a graph based pattern to provide reproducible images for arbitrary areas of interest. In addition imagery products can be generated at any point along a graph. As standards and best practices around Analysis Ready Data (ARD) take shape, DG believes it's important to explore various patterns of image processing and image access to define goals for what ARD should be. In this talk we'll discuss our approach to ARD as well as some the challenges and discoveries we've made in the development of an ARD based platform.

Why is ARD + API based access essential for EO?
Leo Lymburner, Lan Wei-Wang, Fuqin Li, Joshua Sixsmith, David Gavin, Simon Oliver

Geoscience Australia
Consumers of Earth Observation data derived information products gain insight from those products when they are delivered in a time frame, on a platform and in the context required for the consumer to make a decision. These consumers range in spatial scale from continental to farm scale and are making decisions about assets at these scales on temporal scales that range from daily to decadal. These consumers want that information delivered onto platforms/devices that range from mobile phones through to desktop GIS in the context of other non-EO data products that provide the context for the decision. 'Analysis Ready Data' in combination with API based access is essential so that producers of vast and rapidly growing archives of EO data can meet these delivery requirements. ARD ensures that users can engage with EO measurements that are radiometrically comparable, geometrically aligned and free from non-surface features (cloud etc.) API based access ensures that the user engages with the spatio-temporal volume of data required to support a decision, and makes the integration with non-EO contextual data as seamless as possible. The United States Geological Survey and Geoscience Australia have pioneered the development of ARD products for the Landsat series of sensors. Geoscience Australia is also working with a range of collaborators on API based access mechanisms and analysis platforms. This talk reflects on the lessons learned and explores how ARD + API can enable decision makers to realise the value of both public good and commercial satellite image archives.

Importance of analysis-ready data for satellite remote sensing applications: a practical discussion
Sangzi Liang

Corteva Agriscience™, Agriculture Division of DowDuPont
Accurate measurement of ground properties with satellites at a high temporal resolution throughout the crop cycle can lead to better understanding of crop and soil conditions, which is vital for agricultural research and production alike. In order to create value for field researchers, seed production, and growers, the Corteva remote sensing data science team is responsible for utilizing public and private satellite datasets to propose, validate, and prototype satellite applications and algorithms. In this talk, we will discuss some of our internal approaches to creating analysis-ready data, the use cases that motivated these approaches as well as validation results that highlights the importance of ARD. Topics include cloud and shadow detection, atmospheric correction, and spatial-temporal data fusion. We support the efforts of the ARD and the general remote sensing science community to identify common practices and define specifications for ARD, which helps all subsequent processes and will be crucial for wider adoption of satellite data and analyses.

Global 30-m Landsat-derived Cropland Extent Product for Nominal 2015:
produced using Machine Learning Algorithms and Cloud Computing on Google Earth Engine
Prasad S. Thenkabail, Pardhasaradhi Teluguntla, Jun Xiong, Adam Oliphant, Russell Congalton, Murali Krishna Gumma, Aparna Phalke, Richard Massey, Kamini Yadav, Mutlu Ozdogan, Chandra Giri, Jim Tilton, Temuulen Sankey, and Corryn Smith, Ying Zhong, and Tyler Thatcher

USGS

This study presents the first global cropland extent product at 30-m (www.croplands.org) derived using Landsat-8 16-day time-series data for the nominal year 2015. The binary product (cropland versus non-cropland) was created using six generalized steps. First, a mega-file data-cube (MFDC) of 8-band (blue, green, red, NIR, SWIR1, temp, SWIR2 and NDVI) Landsat 16-day time-series data of 2013-2015 was created. Multiple years (2013-2015) were required to obtain cloud-free images, especially in regions of the world where cloud cover is a perennial issue such as in the tropics and\or during rainy seasons anywhere in the world. Each band of Landsat 16-day time-step were time-composited using maximum value composites over 2 to 4 months of data (longer time-periods during the cloudy rainy season), to obtain cloud-free images. This resulted in 32 to 48 layers of MFDC depending on how the images are composited. Second, the MFDC time-series were smoothed using a peak counting algorithm to minimize any residual cloud issues. Third, the world was divided into 100s of unique agro-climate segments- based on agro-ecological zones (AEZs) or on refined AEZs that consider local conditions and other inputs such as elevation. Fourth, accurate cropland versus non-cropland "knowledge" was generated by compiling over 100,000 reference data samples requested from existing ground-data campaigns and\or derived from inspection of very high spatial resolution imagery (sub-meter to 5-m pixels). Fifth, multiple machine learning algorithms such as the random forest and support vector machines were adopted and run on Google Earth Engine (GEE) cloud computing environment using distinct segments of the MFCD data as input. Sixth, cropland products were iteratively refined until high level of accuracies were attained. Seventh, collaborators used over 20,000 additional validation samples to perform independent accuracy assessments for each of the unique agro-climate segments. Overall accuracies in the 100s of segments were consistently higher than 90%, with producer's accuracies, typically, above 85% and user's accuracies, typically, above 75%. The methods described in this study produced 30-m derived cropland areas for each country of the world as well as for sub-national regions such as provinces, districts, and counties and were compared with the national agricultural statistics, compiled by UN FAO and others. Eighth, global net cropland area (GNCA) was 1.873 billion hectares (~12.6% of the Earth's terrestrial area). Ninth, the four leading net cropland area countries, between 8-10 % of the GNCA, were: India (180 Mha), USA (168 Mha), China (165 Mha), and Russia (156 Mha). These four countries, together had 35% of the GNCA's. Countries with 2-3.5% of GNCA were: Brazil (64 Mha), Ukraine (44 Mha), Canada (43 Mha), Argentina (38 Mha), and Indonesia (38 Mha). All other countries of the world had <2% of GNCA. Tenth, Maldova, San Marino, and Hungary had greater than 80% of the country in croplands. Denmark, Ukraine, Ireland, and Bangladesh had between 70-80% of the country in Croplands. Netherlands, UK, Spain, Lithuania, Poland, Gaza Strip, Czech Republic, Italy, and India had 60-70% of the Country in croplands. Methods and approaches discussed in Teluguntla et al. (2018), and Xiong et al. (2017). You can download data from: https://lpdaac.usgs.gov/about/news_archive/release_gfsad_30_meter_cropland_extent_products

References:

Teluguntla, P., Thenkabail, P.S., Oliphant, A., Xiong, J., Gumma, M.K. 2018. A 30-m Landsat-derived Cropland Extent Product of Australia and China using Random Forest Machine Learning Algorithm on Google Earth Engine Cloud Computing Platform. International Journal of Photogrammetry and Remote sensing (ISPRS) Journal of Photogrammetry and Remote Sensing. In press.

Xiong, J., Thenkabail, P.S., Tilton, J.C., Gumma, M.K., Teluguntla, P., Oliphant, A., Congalton, R.G., Yadav, K., and Gorelick, N. 2017. Nominal 30-m Cropland Extent Map of Continental Africa by Integrating Pixel-Based and Object-Based Algorithms Using Sentinel-2 and Landsat-8 Data on Google Earth Engine. Remote Sens. 2017, 9(10), 1065; doi:10.3390/rs9101065 and also downloadable@: http://www.mdpi.com/2072-4292/9/10/1065. IP-088538

HyperScout In-Orbit Demo and future plans

Marco Esposito

COSINE

HyperScout is the first smart hyperspectral imager for nanosatellites. It has been launched on the 2nd of February 2018 at 8:51 CET, from the Jiuquan Satellite Launch Center in China. The applications for which HyperScout has been conceived for are crop water management, fire hazard monitoring, flood detection, change detection of land use and land coverage and vegetation monitoring. The aim of the demonstration mission is to assess the quality of the data that will be acquired and the consequent suitability for the intended applications. A second flight of HyperScout, including additional spectral channels, is planned for the Q2/2019. We will be reporting about the outcome of the operations in orbit, and about the preliminary results obtained from the data evaluation performed during the demonstration project.

Calibrating PlanetScope using external data sources to produce a cloud-/gap-free daily surface reflectance data

Sibo Wang

University of Illinois at Urbana-Champaign

A high-resolution, high-frequency surface reflectance dataset has great potentials for various land-based applications (e.g. agriculture, forestry). Novel CubeSat missions, such as PlanetScope, solves the long-standing dilemma between spatial resolution and temporal revisit frequency by deploying a fleet of miniaturized platforms. However, the use of many small platforms poses additional challenges for the acquisition of a reliable surface reflectance dataset, including the inconsistency between swaths, the low quality of cloud mask due to limited number of spectral bands, and the difference in spectral profile between PlanetScope and more widely-used missions such as MODIS and Landsat. This study attempts to correct such shortcomings by calibrating PlanetScope data with external data sources, including MODIS, Landsat, and a fused product based on the newly developed STAIR fusion algorithm. We use cumulative distribution function (CDF) matching and transfer learning techniques to produce a cloud-/gap-free daily surface reflectance data cube for cropland whose spectral profile is the same as that of MODIS NBAR.

Cloud Detection for the Planetscope Constellation

Keith Beckett, I. Zuleta, K. Kapadia, A. Collison, K. Lenhard, J. Martinez-Manso, M. Schwarz, Y. Yi, A. Kanwar, M. Weisman, F. Warmerdam

Planet Labs, Inc.

High accuracy detection and mapping of clouds in optical imagery is critical to Earth Observation data users for two reasons: 1) to support effective discovery of and selection from the available imagery to serve a given application's needs, and 2) to support the effective use of those parts of the selected imagery that meet the a given application's needs. Cloud maps are also a key ingredient in ARD (Analysis Ready Data) stacks of optical imagery. Planet has recently developed a new approach leveraging a combination of RTM (Radiative Transfer Model), and advanced Machine Learning and Computer Vision techniques to accurately map out clouds, shadows, haze and snow in 4-band (blue, green, red and NIR) optical imagery from the Dove and SkySat constellations. The focus of our approach was to significantly improve the data discovery experience and, unlike past approaches, to express the usability of the data. We will outline the key aspects of the cloud detection methodology, summarize the performance achieved, compare it to the previous approach and share some of the lessons learned along the way!

Operational Biophysical Analysis Ready Data for Landscape Monitoring

Peter Scarth

University of Queensland

Government agencies in northern and eastern Australia, the University of Queensland and the University of NSW collaborate through the Joint Remote Sensing Research Program (JRSRP) to operationally produce analysis ready remote sensing information about Australian environments. These products include terrain corrected surface reflectance, fractional cover (derived from a spectral unmixing model) and persistent green time series are used by both government agencies in complex policy settings as well as by landholders monitoring paddock conditions, organisations supporting land management initiatives in the rangelands and Great Barrier Reef catchments, and researchers developing tools to understand land condition, land degradation and human health across all of Australia. To service users across these multiple disciplines, we deliver the full times series of national scale data through time series enabled web mapping services and customised web processing services that enable the full time series over any spatial extent to be interrogated in seconds via a RESTful interface. These services interface with browser applications that provide product visualization for any date in the time series, tools to draw or import polygon boundaries, plot time series ground cover comparisons, look at the effect of historical rainfall and tools to run the revised universal soil loss equation in web time. We also make virtual datasets available to access our COGs as both spatial and temporal mosaics, and these are enabling new applications by developers who link these with other cloud hosted data to produce products such as land condition and seasonal forecasting tools. This talk covers the full stack remote sensing process to build, validate and operationally deliver nationally consistent ARD biophysical products from Landsat and Sentinel 2. While the field data component and computational requirements are large, we show that the largest barriers are in working with users to understand the data, and working with developers to maximise value from the information and delivery systems. NOTE: This talk covers JRSRP data available from and downstream services:
http://data.auscover.org.au/xwiki/bin/view/Product+pages/WebHome

Combining HR and VHR imagery for EO-based land monitoring solutions

Michael Riffler, Andreas Walli, Jürgen Weichselbaum, Christian Schleicher, Christian Hoffmann

GeoVille Information Systems
and Data Processing GmbH

As one of the leading satellite based land monitoring service providers, we are interested in combining various data streams including high resolution (HR) and very high resolution (VHR) optical imagery in a fully automatized way, with the aim to provide more accurate and customized land monitoring applications. This includes, for instance, dedicated services for the agricultural sector in assessing the growing conditions of particular crops, the detection of changes for various land cover and land use types (e.g., urban, forest, etc.) as well as environmental applications to monitor coastline or wetland dynamics at various levels of detail and scale. In one of our services, we address the requirements of the European AgroFood value chain industry for example on logistics management support. The methodologies applied to extract the relevant information require frequent (daily) data coverage. Usually large efforts are necessary to prepare and harmonize HR and VHR imagery most often containing geometric errors and differences in the spectral reflectance values hampering the automated analysis of the data. This also includes the mechanisms of data access, which are ideally based on API-services in order to integrate them into fully automated B2B solutions. In the presentation we will focus on the requirements for Analysis Ready Datasets we have in order to combine the most important HR and VHR data streams to be used in the information solutions we build.

VITO's Framework for Analysis Ready Data

Buchhorn, Marcel; Dries, Jeroen; Smets, Bruno; Van De Kerchove, Ruben; Benhadj, Iskander

VITO - Flemish Institute for Technological Research NV

Various new EO systems, the opening of EO archives to the public, and the technological advances to handle huge data amounts makes it possible to harmonize data of various EO platforms in the spectral, spatial and temporal domain. Users get more and more challenged with data search, pre-processing, interoperability, storage and access instead of running analysis. One solution is to provide Analysis Ready Data (ARD) to the users where the data provider handles all the data preparation work on their own operational platforms. The remote sensing unit of VITO (Flemish Institute for Technological Research) implemented first versions of operational ARD's for several EO datasets (e.g. SPOT VGT, PROBA-V, Sentinel-2) with a focus on time series analysis which are accessible/customizable via various portals (www.vito-eodata.be, http://terrascope.be/en). We also offer a Virtual Research Environment (VRE) via our PROBA-V Mission Exploitation Platform (MEP) which gives users direct access to the data plus the possibility to directly run analysis in a Hadoop/Spark private cloud. This platform allows global on-the-fly data analytics using a continuous full processed archive (20 years for SPOT VGT and PROBA-V) at various resolutions where the user can focus on the application at hand without the hassle of data preparation.
In our presentation we will give an overview of our framework for ARD with a focus on time series analysis. We will present in detail our approach to generate harmonized datasets for PROBA-V and Sentinel-2. In order to get interoperability between PROBA-V and Sentinel-2, we went back to the PROBA-V Level 1C data and implemented a fully new designed operational processing chain for the pre-processing (geometric correction including pixel alignment to the Sentinel-2 grid, atmospheric correction, cloud and cloud shadow detection) and an advanced processing chain for the post-processing (image compositing to generate time series stacks, image clipping to the Sentinel-2 tiling grid, temporal cloud/outlier screening, gap-filling in the time series of a pixel via a Kalman-filter approach which uses the PROBA-V 300m corresponding side cameras information, compositing to 5-daily observations, generation of status masks – usable data masks). Moreover, several Level-3 products (e.g. NDVI, EVI) are generated from the time-series surface reflectance data stacks. Over 50 million single acquisitions were processed to generate a global PROBA-V ARD at 100m resolution spanning over 5 years (since 2014 until today) and fully aligned to the Sentinel-2 tiling grid and naming. Finally, metadata are injected and the products are implemented into a SpatioTemporal Asset Catalog (STAC) called VITO's Earth Observation Collection Catalogue. The catalog is based on an ElasticSearch database to allow easy searching of the json documents. VITO ingests various products into a Geotrellis based backend, allowing on the fly extraction of time-series cubes for specified areas or whole tiles. In the second part of the presentation, we will show some ARD-derived products (e.g. continental land cover map, agriculture applications) and the improvements in data analysis through usage of ARD at VITO. Next to our experiences we will present challenges we encountered on the way e.g.: the need for maintainable/flexible/usable metadata formats and the catalogue specification for them, since especially our specific products, e.g. the ten daily composites, have certain properties that are not always easy to represent in current formats (for instance: periodicity, nominal date). And the necessity of defining standards for scalable Global Reference Grid systems from coarse to high resolution will also be stated.

Starlight's InsightGIS: Tools for GEOINT Platforms with Analysts as the First Focus

David Bernat

Starlight

Analysts' central responsibility to provide knowledgeable interpretation of analysis-ready data (ARD) will remain the backbone of geospatial data analysis far into the future of AI and automation. Starlight specializes in tools which put analysts and remote sensing workflows into a platform model. These tools transform the analyst experience and re-imagines AI automation as a simpler byproduct of work done on our platform. Starlight will walk through case studies from natural language processing to show the surprising consequences of our platform for automation, provenance, and granular trust. We will connect this to contemporary GEOINT by showing our plan to embed the analyst workflow used recently to discover an overlooked nuclear enrichment site in North Korea at the James Martin Center for Nonproliferation Studies and Planet.

Multi-sensor data fusion for enhanced land surface characterization in the era of the CubeSat

Rasmus Houborg

South Dakota State University

Constellations of small CubeSats are emerging as a novel observational resource with the potential to overcome the spatiotemporal constraints associated with conventional single-sensor satellite missions. With a constellation of more than 175 active CubeSats deployed in low-earth orbits, Planet has realized daily RGB and near-infrared (NIR) imaging of the entire Earth land area at very high spatial resolution (~3 m), which was unimaginable just a few years ago. The spatial and temporal enhancements enabled by CubeSat systems can provide critical insights into rapidly evolving vegetation dynamics with a factor of 10 increase in spatial resolution compared to Landsat. However, the large number of constellation satellites, and relatively cheap sensor components and design, introduce cross-sensor inconsistencies and cross-calibration challenges. While superior in terms of spatiotemporal resolution, the radiometric quality and spectral resolution is clearly not equivalent to that of rigorously calibrated space-agency satellites such as Landsat and Sentinel. A CubeSat Enabled Spatio-Temporal Enhancement Method (CESTEM) has been developed to realize the full potential of CubeSat systems by exploring synergies with more conventional satellite systems. CESTEM serves as a multi-purpose data fusion scheme for radiometric normalization, smooth temporally-consistent phenology reconstruction, and spatiotemporal enhancement of biophysical properties. CESTEM can use high quality Landsat 8 or Sentinel-2 surface reflectance and vegetation biophysical retrievals as references to transform relatively noisy CubeSat time series observations into Landsat or Sentinel consistent estimates. CESTEM offers a unique data-driven avenue towards

cross-platform interoperability enabling spatiotemporal enhancement of spectral metrics and biophysical properties retrievable from conventional large satellite systems. The application of CESTEM over a variety of agricultural landscapes and environments highlights the resolution advantage and impressive capability to provide repeatable time-critical insights into within-field vegetation dynamics, the rate of vegetation green-up, and the timing and progress of key phenological transitions, that are largely uncaptured by Landsat or Sentinel-2 imagery. Opportunities for near real-time processing over broad areas will be discussed.
Ingesting Astro Digital Landmapper-BC data the analytics-ready dataset Alex Kudriashova, Chris Biddy *Astro Digital, Inc.* To ingest imagery of new multispectral sensor into an analytics-ready time series it should be cross-calibrated. Nowadays leading satellite data community took effort to make a consistent dataset of cross-calibrating publically available Landsat 8 and Sentinel-2. This dataset has become an industry standard and commercial providers are currently tuning their data to be aligned with this dataset to improve the fullness of Earth Observation dataset. After a successful launch of the first Landmapper-BC in 2018 Astro Digital was actively working at ingesting the Landmapper-BC data into the analytics-ready dataset of Landsat 8 and Sentinel-2. As a result we were able to demonstrate the consistency of Landmapper-BC, Landsat 8 and Sentinel-2 pixel values for the same location for recent dates. Creating of an analytics-ready product from Landmapper-BC data required to make a proper geometric and radiometric correction, make an orthorectification, apply a surface reflectance model and make minor changes to address minor changes between the sensors. We did it and validated the final product against industry-standard methods including flat field analysis and consistency of the time series for known targets. Overall this work demonstrates possibility of ingesting data from Landmapper-BC constellation to make a Landsat 8 + Sentinel-2 dataset more frequent and cloud-resistant. This finding enables a lot of industrial applications from disaster response to early crop disease identification at medium resolution globally.
High Quality Analysis Ready Data Starts with Mission Design William Parkinson *UrtheCast, Inc.* The core components of Analysis Ready Data (geometry, atmospheric correction, cloud masking, and radiometry) pose complex challenges when working with heterogenous constellations of earth observation data. Unique radiometric, geometric, and orbit characteristics all hamper efforts, increase error, and limit product quality. UrtheCast's UrtheDailyTM maximizes ARD quality by design. We are concurrently building a constellation of identical satellites that will be launched along a single orbit plane with space and ground systems designed to prioritize product consistency and quality. This gives UrtheCast the unique ability to build a holistic processing and calibration strategy that increases the quality of ARD products over those created from heterogeneous constellations.
Perspectives from EOFSAC: A NASA Program for Food Security and Agriculture Michael Humber, Alyssa Whitcraft, Chris Justice, Inbal Becker-Reshef *University of Maryland* The Earth Observations for Food Security and Agriculture Consortium (EOFSAC) is a new, multidisciplinary program commissioned by NASA and led by the University of Maryland to enhance the use of satellite data in decision making related to food security and agriculture domestically and globally. The program brings together over 40 field leaders in agriculture and food security research from universities, governments, NGO's, and the private sector with the goal of advancing the state of the science and the adoption of Earth observations. A holistic view of these issues requires working with remote sensing instruments across the multiple remote sensing resolutions (spatial, temporal, spectral, radiometric). These sources of remotely sensed data support food security and agricultural applications from the global to local scales, with each application inherently providing a unique set of challenges and opportunities. In this context, interoperable analysis ready data is an important asset as it facilitates scaling between applications and leveraging preexisting workflows. We provide our perspectives on the issues facing users of remotely sensed data across resolutions based on our experience working with data ranging from coarse satellite observations (e.g. MODIS) to high resolution satellite observations (e.g. Planet Doves) and UAV imagery. We highlight some of the challenges with scaling across different sensor resolutions, providing end-users with analysis ready data, creating useful tools, and harmonizing data which is acquired from a wide variety of partner organizations and other sources.
Planet Data in Context Robert Simmon *Planet Labs* Planet's constellation is one of the newest Earth observing platforms in a lineage of land remote sensing that goes back to at least 1972, with the launch of Landsat-1. Our data—with a near-daily cadence and resolution from 6.5 and 0.8 meters per pixel—fits between the daily observations of MODIS, VIIRS, & OLCI at up to 250 m/px, 5- and 16-day revisit and 10- and 30 m/px Sentinel-2 and Landsat data. Visible and near infrared bands are suitable for mapping and vegetation monitoring tasks. Daily PlanetScope observations at 3-5 meter resolution enable new types of analysis, along with SkySat's 0.8 m/px data suitable for validation and tip-and-cue applications, and the long-term and high dynamic range archive of RapidEye.

Analysis Ready Data as addressed by the Committee on Earth Observation Satellites (CEOS ARD 4 Land)
Steven Hosford,
Adam Lewis , Jenn Lacey, Susanne Mecklenburg, Ferran Gascon

ESA, Geoscience Australia, USGS
Over the past two years, the space agencies who are members of the Land Surface Imaging "Virtual Constellation" of the Committee on Earth Observation Satellites (CEOS) have been developing a framework for generating Analysis Ready Data for land applications. This framework, which has been called CEOS Analysis Ready Data for Land or CARD4L, is foreseen to enable users to access satellite data products that are 'ready to use' for a wide range of land applications. Moreover, CARD4L aims to enable non-expert users access to products that have been processed 'far enough' to be suitable for immediate analysis for a range of applications, while ensuring they are not too specific to only be used for particular topics or areas. Data products aligned with the CARD4L specifications should be an important enabler of diverse end-user EO processing environments such as datacubes and exploitation platforms. CARD4L's ambition is to enable users to easily locate products that are suitable for ingestion into such EO processing environments, and for them to have confidence that these different CARD4L products will limit as far as possible barriers to interoperability between data from different sensors. The CARD4L framework defines specifications for several CARD4L "product families" including surface reflectance, land surface temperature and radar backscatter. Additional product families are currently being defined for other products derived from SAR imagery. Each product family specification includes requirements for general and pixel level metadata and corrections to ensure that the product is a valid measurement and can be accurately located in space. The CEOS CARD4L framework aims to be generic, with the potential to support analysis ready data products from other data sources. A consistent, standardised approach to describing and assuring analysis ready data products, suitable for government and private sector data, would be highly advantageous to data providers and users alike.

Creating Interoperable Synthetic Aperture Radar Data at the ASF DAAC
Scott Arko

University of Alaska
Synthetic Aperture Radar (SAR) provides a weather-independent and cloud-free dataset that can be valuable in a variety of applications. However, SAR data have historically been provided as non-projected images in formats that are difficult to use and unsupported by many modern tools. In this way they have not been interoperable with existing EO datasets and typical user workflows. The Alaska Satellite Facility (ASF) Distributed Active Archive Center (DAAC) is working to make SAR data more accessible through the development of cloud-based processing systems and adoption of new formats for both data and metadata. This presentation will focus on the processing of Sentinel-1 SAR data using ASF's Hybrid Pluggable Processing Pipeline (HyP3). HyP3 provides a scalable architecture for processing large volumes of remote sensing data in the Amazon Web Services (AWS) cloud. Using this engine, ASF DAAC provides advanced processing and conversion of image products to usable formats, including cloud-optimized geotiffs. Example products include radiometrically terrain corrected (orthorectified) SAR image products and interferometric SAR products. Current work is focused on using the high-cadence data from Sentinel-1 to generate analysis ready time series of both amplitude and interferomtric SAR products. ASF DAAC is investigating the best way to publish these ARD products to the larger remote-sensing community.

Open Source Applied Research: SpaceNet and Machine Learning for Remote Sensing Data Sets
David Lindenbaum

In-Q-Tel
SpaceNet is a partnership initiative between CosmiQ Works, DigitalGlobe, and Radiant Solutions designed to accelerate machine learning against geospatial problems using remote sensing data. Over the last 2 years, SpaceNet has released over 5700 sq km of imagery with 685,000 building footprint labels and 8000 km of road network labels, run three competitions to solve the challenge of extracting building footprint and road networks from satellite imagery, and open sourced 13 separate algorithms. SpaceNet has also undertaken hosting other, related data sets such as IARPA's Functional Map of the World. In this presentation, we will have provide an overview of the SpaceNet initiative, discuss upcoming competitions for 2018 and 2019, how derived algorithms could be used in an operational setting, and how we see ARD assisting future machine learning research and development. https://spacenetchallenge.github.io/.

Surface Reflectance Comparison between OLI and Planet imagery
Minsu Kim

USGS
Surface reflectance is being compared in selected region around the globe. Landsat OLI data that covers wide range of reflectance range over wide geospatial distribution is selected. Corresponding surface reflectance image from Planet will be compared.

ARD and Exploitation Platforms
George Percivall

Open Geospatial Consortium

The Open Geospatial Consortium (OGC) has activities relevant to the workshop scope of "the current state-of-the-art in satellite data interoperability". This presentation will focus on two main topics with the option to discuss other relevant topics that the participants may wish to discuss, e.g., WFS3. The two focus areas of development: 1) Geospatial Datacubes and 2) Earth Observation Exploitation Platforms. 1) A Geospatial Datacube provides access to and analytics on analysis ready data (ARD) organized with coordinate axes of space and time with cells in the cube containing data of geospatial features, e.g., imagery. OGC members implementing geospatial datacubes are documenting common practices to spur development and leading to the possibility to federated geospatial datacubes. 2) OGC is forming a Earth Observation Exploitation Platform Domain Working Group with the goal of defining a standards-based framework for cloud-based access to and analysis of EO data. An ad-hoc meeting was held in March 2018 to scope the working group with the results issued in a request for comment: http://www.opengeospatial.org/pressroom/pressreleases/2792

Satellogic: processing and ARD challenges

Slava Kerner

Satellogic

With 3 satellites of Newsat family in production, we will review pre-processing pipeline. In this context, ARD challenges common to the industry will be discussed. Need for balance between absolute and relative geo-accuracy is important to ARD products. We will consider some lessons we learned about geometric stability for narrow-baseline imagery. Focusing on the spectral challenge, we will describe vicarious calibration of Newsat, as well as continuous cross-calibration performed with respect to Landsat-8 and Sentinel-2. Finally, we will discuss the need for spectral normalization across the different satellites. With Landsat-8 and Sentinel-2 serving as the standard dataset for cross-calibration, there is a clear need for larger variation of data sources. We call for more satellite manufacturers to make TOA measurements over RadCalNet sites publicly available, as a way to promote better ARD-validation; we suggest to brainstorm the framework for that.

On-the-fly ARD in Practice using Sentinel Hub

Miha Kadunc, Anže Zupanc, Matej Batič, Rok Močnik, Grega Milčinski

SINERGISE LABORATORY FOR GEOGRAPHICAL INFORMATION SYSTEMS, LTD.

Sentinel Hub services provide on-demand tailored access to satellite data. They feature a JavaScript processing engine, which allows users to perform pixel-wise operations on the server before data retrieval. In 2017 we added support for multitemporal processing, a form on on-the-fly analysis-ready data (ARD), which enables measurements from several acquisitions to be analyzed in a single function.

We have since used this functionality in several experiments and production projects, facing many challenges and learning valuable lessons in producing and using on-the-fly ARD. I will present our experiences in using these capabilities to develop the time-lapse and time series functions of the EO Browser web application. I will describe how we produced cloud-free mosaics from Sentinel-2 data for New Zealand and from Sentinel-3 data for Europe, and what we learned when using the platform for the Sentinel-2 Global Mosaic project. I will report on our experiences using machine learning to develop "s2cloudless" - a open source cloud screening algorithm for Sentinel-2 - and how this led us to design "eo-learn" - a python library for easy manipulation of in-memory patches of analysis-ready data. Within all of these efforts, we aimed to balance the cost, performance and quality of the output, while trying to keep the solutions applicable globally. We believe our experiences will be valuable in future efforts to define interoperable ARD services.

Harmonized Landsat/Sentinel-2 Products for Land Monitoring

Junchang Ju (NASA GSFC), Sergii Skakun (U. Maryland), Jean-Claude Roger (U. Maryland), Martin Claverie (U. Louvain)

NASA
University of Maryland
Université catholique de Louvain

Many land applications require more frequent observations than can be obtained from a single “Landsat class” sensor. Agricultural monitoring, inland water quality assessment, stand-scale phenology, and numerous other applications all require near-daily imagery at better than 1ha resolution. Thus the land science community has begun expressing a desire for a "30-meter MODIS" global monitoring capability. One cost-effective way to achieve this goal is via merging data from multiple, international observatories into a single virtual constellation. The Harmonized Landsat/Sentinel-2 (HLS) project has been working to generate a seamless surface reflectance ARD product by combining observations from USGS/NASA Landsat-8 and ESA Sentinel-2. Harmonization in this context requires a series of radiometric and geometric transforms to create a single surface reflectance time series agnostic to sensor origin. Radiometric corrections include a common atmospheric correction using the Landsat-8 LaSRC/6S approach, a simple BRDF adjustment to constant solar and nadir view angle, and spectral bandpass adjustments to fit the Landsat-8 OLI reference. Data are then resampled to a consistent 30m UTM grid, using the Sentinel-2 global tile system. Cloud and shadow masking are also implemented. Quality assurance (QA) involves comparison of the output 30m HLS products with near-simultaneous MODIS nadir-adjusted observations. Prototoype HLS products have been processed for ~7% of the global land area using the NASA Earth Exchange (NEX) compute environment at NASA Ames, and can be downloaded from the HLS web site (https://hls.gsfc.nasa.gov). A wall-to-wall North America data set is being prepared for mid-2018. This talk will review the objectives and status of the HLS project, and illustrate applications of high-density optical time series data for agriculture and ecology. We also discuss lessons learned from HLS in the general context of implementing virtual constellations.

Data Fusion for Monitoring Rainforest Carbon: A case for global co-registration
Pramukta Kumar, Joe Mascaro, Greg Asner

Carnegie Institution for Science
Measuring and monitoring forest carbon content is an essential component of the UN Reducing Emission from Deforestation and forest Degradation (REDD+) process. The Carnegie Airborne Observatory (CAO) uses an advanced airborne LiDAR instrument to make precise estimates of rainforest carbon over a wider area than can be accomplished efficiently by researchers on the ground. While an essential scientific instrument, the airborne LiDAR is cost and time prohibitive to use as an operational monitoring system. Developing models that utilize commercial electro-optical (EO) satellite-based sensors to recover information measured using an instrument such as the CAO is one promising way to scale-up measurements to a national and regional level. Essential to this imagery use case is reliable co-registration of commercial EO and airborne LiDAR data. In the context of rainforest monitoring, this talk will discuss the consistent data positioning across data sources as an important component of Analysis Ready Data (ARD), how well Planet and CAO data currently match, and attempt to make the case that global co-registration of imagery across data sources is both useful and attainable with current technology.

ARD for Agricultural Remote Sensing: Challenges and Use Cases
Keely Roth

The Climate Corporation
Analysis Ready Data (ARD) are a critical input to many agricultural applications within remote sensing from both a research and commercial production perspective. Rapid exploration and development of generalizable and portable algorithms often requires researchers' ability to focus on science and innovation over data ingest and pre-processing typical of L0 to L2 data transformations. Having the ability to ensure timely access and delivery of consistent remote sensing raw and derived products required by various Agricultural models and processing chains depends heavily on mature imagery platforms that incorporate STAC and ARD. The Climate Corporation leverages remote sensing data in both research and production capacities from a variety of sensors and platforms that include commercial and publicly available datasets. In this talk, we'll share some of our specific Agricultural use cases for ARD, focusing on the challenges that ARD can address including data quality control and tracking, data accessibility, custom stack creation and data cubing, atmospheric correction, cross-platform calibration, and common spatial gridding. Implications for building interoperable data standards as a broader technical requirement for other RS science capabilities within Agriculture will also be highlighted.

Interoperability between Planetscope, RapidEye and MODIS TOAR products
Arin Jumpasut et al.

Planet Labs, Inc.
Analytic ready data sets provide a convenient way of combining geospatial data from multiple sources without any pre-processing of the data. This is important as it can make analysis done on the data more consistent as well as providing a convenient starting point. One aspect of this is the radiometric interoperability of the data from different satellite constellations. This presentation will discuss work done to characterise the radiometric differences between Dove data, Rapideye data and MODIS data. Here we show that the radiometry is stable not only within constellations when considering different multi-satellite constellations (like the PlanetScope constellation) but also through the growing season and within the geospatial location of the area of interest (the US Midwest region). This means that a single correction is applicable for an entire constellation within the area of interest and time period. We will present per-band linear correction factors that can be used to create a more interoperable data product. In the future, this will be improved by considering spectral unmixing, expanding the geographic area considered, different seasons and considering the effect of atmospheric correction.

Harmonizing Different TOAR and SR Products using Simple Regression Models
Gina Li et al.

Planet Labs, Inc.
Analysis Ready Data (ARD) provide geospatial end-users with preprocessed and ready to use time-series satellite imagery. Individual constellations provide unique sets of pros and cons to interoperability -- while one satellite will revisit at a higher rate, another will be better radiometrically calibrated, providing a challenge of having a consistently well-calibrated product throughout a given time series. ARD stacks that are composed of multiple satellite constellations (PlanetScope, SkySat, RapidEye, Sentinel-2, Landsat and MODIS) are currently capable of being generated internally at Planet and can provide insight to radiometrically calibrate small agricultural sites. In this study, a multi-constellation ARD is created to radiometrically correct PlanetScope radiometric values to match those of Sentinel-2 in individual agricultural fields within Imperial Valley, CA. Per-band random forest regression models are created to match PlanetScope RSRs to Sentinel-2 RSRs. In cross-validation, it is shown that per-band corrections yield similar NDVI values for PlanetScope-corrected top of atmosphere and surface reflectance values as compared to Sentinel-2. These results demonstrate a use-case scenario for per-band radiometric correction for individual agricultural fields, which can provide agricultural growers with a daily, well-normalized product for monitoring purposes. It is anticipated that similarly well-calibrated results can be drawn for other small-scale, homogenous land covers.

Validating PCI ARD data with Sen2Cor and LaSRC including BRDF correction and Topographic Normalization

Wolfgang Lueck, Arnold Hougham, Guillaume Morin, Kevin Jones

PCI Geomatics

To facilitate hyper-temporal image analysis using imagery acquired by both Landsat 8 and Sentinel 2 A/B, imagery needs to be processed to Analysis Ready Data (ARD). We have implemented automated workflows that process both Landsat 8 and Sentinel 2 A/B in a consistent manner to an analysis ready data level. A super-registration of imagery to 1/10th of the ground sampling distance (GSD) ensures both radiometric and geometric consistency within the data stack. With the strong correlation between spectral bands, high frequency information captured in the Landsat 8 pan or Sentinel 2 10m VIS-NIR bands is transferred to the lower spatial resolution coastal, vegetation and SWIR bands, reconstructing all spectral bands to a GSD of 15m in the case of Landsat 8 and 10m in the case of Sentinel 2. This excludes the use of the Cirrus and water vapor bands. A spectral pre-classification guides the atmospheric correction process based on the DLR Atcor code and is further used for bidirectional reflectance distribution function (BRDF) corrections and topographic normalizations, producing Nadir BRDF Adjusted Reflectance (NBAR) products for both Landsat 8 OLI and Sentinel 2 MSI.

In this paper we compare and evaluate algorithms implemented in Sen2Cor, LaSRC and HLS ARD products generated by the USGS over the USA, against PCI's algorithms implemented within our products to produce ARD products. Simultaneous acquisitions of Landsat 8 and Sentinel 2 imagery over Canada, USA and South Africa as well as a dense time series are used to analyze between scene variability and fit to spectral libraries. The reduction of variability in NBAR between simultaneous acquisitions and multitemporal comparisons over pseudo invariant features has proven to be the sound scientific method for validating our algorithms. Data processed with our multi-sensor workflows are suitable for the ingestion and analysis in the Open Data Cube and can lead to an improved signal / noise ratio when compared to products generated by HLS LaSRC and Sen2Cor respectively.

Large volume production system to prepare ARD data and integration with the Open Data Cube – Initial lessons learned from deforestation monitoring case study in Canada

Guillaume Morin, Masroor Hussain, Kevin Jones

PCI Geomatics

PCI's Large volume production system to prepare ARD data and integration with the Open Data Cube – Initial lessons learned from deforestation monitoring case study in Canada. PCI Geomatics has established a close collaboration with the Pacific Forestry Service (PFC),located in Victoria Canada, to prototype a large volume production system to automate: the preparation of multi-temporal ARD data stacks, its integration with the open data cube, and multi-temporal analysis. The PFC wishes to establish an automated operational system to determine the locations and the extents of deforestation across Canada by using high resolution image data (from Landsat and Sentinel initially). PCI Geomatics has leveraged its GXL large scale volume production system in conjunction with specific workflows to automate the production of ARD data. Methods and protocols have been put in place to process imagery into format suitable for ingesting into the CEOS Open Data Cube. ARD product for Landsat-8 and Sentinel-2 over a two-year period (roughly 350 images) were created. PCI MRA fusion algorithm, which achieve higher accuracy than industry standard, was used to pan-sharpen Sentinel-2 imagery to attain 10m resolution for the VIS-NIR bands. Automated procedures were used to generate required metadata for Open Data Cube ingestion. Both the original vegetation anomaly detection algorithms, available in the CEOS Open Data Cube, and the PCI's modified version were used to analyse Landsat and Sentinel-2 data. The presentation will focus on the lessons learned from this prototype implementation over Fort McMurray, Alberta. We will also discuss the additional research at PCI to extend existing Open Data Cube. At the core is multi-resolution multi-temporal analysis by using geographic object-based image analysis (GEOBIA) instead of per-pixel analysis. We are also exploring the potential benefits of using geographic median composite. Also we are investigating cross sensor calibration and fusion to allow the integration of Planet Scope, Skysat, Sentinel 2 and Landsat data in the open data cube.

Sentinel-2 ARD - toward multi-mission interoperable/ harmonised products

Ferran Gascon, Steven Hosford

ESA
CNES

There is a growing need from several applications (e.g. smart farming) of having Analysis Ready Data (ARD) providing surface reflectance measurements with frequent temporal revisit. In parallel, several optical imaging sensors are flying, and even more foreseen to be launched in the coming years. All these missions together have a strong potential if they would be combined in a single and coherent data stream providing consistent surface reflectance measurements with a higher revisit frequency. In this direction and complementing the currently available Sentinel-2 ARD product (Level-2A), ESA plans to develop an harmonised ARD product combining data from several sensors and taking as reference mission Sentinel-2. The data from all missions will be converted to a Sentinel-2 like stream, i.e. data having Sentinel-2 spatial resolution and spectral responses. As a first step, ESA is developing a prototype processor (sen2like) for combining Sentinel-2 and Landsat-8 in an harmonised data stream. This processor will be expanded in the future to integrate data from other missions such as Sentinel-3 or PROBA-V.

The presentation will include a description of the objectives of the project, the current status (including the methods used) and the next steps. Additionally, the currently available Sentinel-2 ARD products (Level-2A) will be introduced.

Google Earth Engine and ARD
Tyler Eriksson,
The Google Earth Engine Team

Google
Google Earth Engine is a cloud-based geospatial processing platform that brings Google's massive computational capabilities to bear on a variety of high-impact societal issues including deforestation, drought, disaster, disease, food security, water management, climate monitoring and environmental protection. Earth Engine consists of a multi-petabyte analysis-ready data catalog co-located with a high-performance, intrinsically parallel computation service. It is accessed and controlled through an Internet-accessible application programming interface (API) and an associated web-based interactive development environment (IDE) that enables rapid prototyping and visualization of results. The data catalog houses a large repository of publicly available geospatial datasets, including observations from a variety of satellite and aerial imaging systems in both optical and non-optical wave-lengths, environmental variables, weather and climate forecasts and hindcasts, land cover, topographic and socio-economic datasets. All of this data is preprocessed to a ready-to-use but information-preserving form that allows efficient access and removes many barriers associated with data management. [1] This talk will give an overview of the Earth Engine platform, focusing on the data catalog and data storage aspects of the system.
1. Gorelick, Noel, et al. "Google Earth Engine: Planetary-scale geospatial analysis for everyone." Remote Sensing of Environment 202 (2017): 18-27. [https://doi.org/10.1016/j.rse.2017.06.031]
2. https://earthengine.google.com/datasets
3. https://github.com/google/earthengine-api

Analogue-based colorization of remote sensing images using textural information
Mathieu Gravey

Universite of Lausanne
Environmental studies typically use remote sensing imagery coming from a range of different sensors. In particular, studies on long-term change detection face the problem of fusing datasets having different spectral and spatial resolution. Here, we are only considering the fusion of products having a different number of spectral bands. The easiest solution is to reduce the information of the most recent product, thereby degrading the information to a common representation. To avoid such loss of information, we propose a new method based on multiband spatial pattern matching. Our method uses analogous scenes taken from sensors where all desired bands are informed. These analogs have conceptually the same role as the training images used in multiple-point geostatistics simulation. The spectral characteristics of the training image are then transferred to a target image, where new synthetic spectral bands are generated. A spatial pattern matching procedure is used to control this transfer, resulting in preservation of spatial and spectral coherence in the results. We illustrate the method by increasing the spectral resolution of archive Corona satellite images to the same level of spectral resolution and coverage as Landsat-8 modern imagery. The same methodology could be used to enhance the spectral coverage of modern sensors, such as those used by Planet Labs. Furthermore, we envision that these approaches can be used not only for multi-spectral imagery but also for hyperspectral data.

PanArctic Shipping: Using satellite Automated Identification System (AIS) data to aid informed decision-making in the Arctic
Greg Fiske

Woods Hole Research Center
In concert with the International Maritime Organization guidelines for ships operating in Arctic Ice-Covered waters in 2002, the Arctic Council initiated the Arctic Marine Shipping Assessment (AMSA), which issued its final report in 2009. The history of AMSA involves questions arising with urgency into evidence and options that contribute to informed decisions regarding Arctic marine safety and environmental protection, underlying a process of science diplomacy. Primary sources of data for AMSA involved ship tracking with the Automatic Identification System (AIS) from ground stations, shore-based radar systems, and Arctic nation details of fishing vessels and other smaller ships. Monitoring Arctic ship traffic fundamentally changed the year of the AMSA report, when satellite AIS records began providing continuous, synoptic, pan-Arctic coverage of individual ships with data pulsed over seconds to minutes. Our analysis investigates the oldest and longest continuous satellite AIS record for the Arctic (from 1 September 2009 through 31 December 2016). The data were produced via an array of micro satellites and provided by SpaceQuest Inc. Initial results of our analysis indicate that the mean center of shipping activity has moved 300 kilometers north and east—closer to the North Pole—over the 7-year span. Additionally, compared to International Maritime Organization (IMO) registered ships with Class A transponders, smaller vessels with Class B transponders (without IMO registry) are increasing faster in the Arctic Ocean. Noting that shipping and sea ice are recognized as primary drivers of change in the Arctic, sea-ice coverage data from the National Snow and Ice Data Center also was assessed alongside the ship traffic patterns, revealing ship-ice interactions are increasing and moving northward. Satellite AIS analyses provide the core observational framework for the next generation AMSA to inform operational and infrastructure decision-making processes for safe, secure and reliable maritime traffic in the Arctic Ocean.

Combining Multiple Data Sources in time series analysis for monitoring land change

Curtis Woodcock

Boston University
By combining data from multiple sources it is possible to improve the temporal frequency of observations, which improves monitoring capability. Our efforts to date have focused on time series analysis of Landsat data, but we are now integrating both radar and Sentinel-2 optical data. Our hope is to be able to add Planet imagery to this effort. One of the benefits of our approach is that we build and "carry" a model for the expected surface reflectance that could be used to help calibrate Planet data. One of my hopes for the meeting will be to learn more about Planet data and its processing, with the hope of improving our ability to combine it with other sources for land change monitoring.