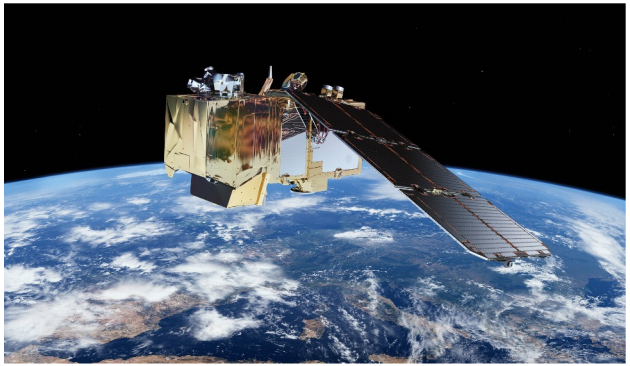
Cloud-Based Data Production at Scale: NASA's Harmonized Landsat Sentinel-2

May 30, 2023

Brian Freitag
NASA Marshall Space Flight Center (MSFC)/IMPACT Project

What is HLS?

- What is “harmonized”?
  - Using data from two similar instruments and constructing an algorithm so products from each instrument can be used interchangeably
- HLS reprojected onto an analysis-ready grid
  - Analysis-ready means data pixels cover the exact same location in both HLS datasets
  - Reduces analysis overhead for integrating data from both sensors



Sentinel-2 photo credit: ESA

Landsat 8 photo credit: NASA

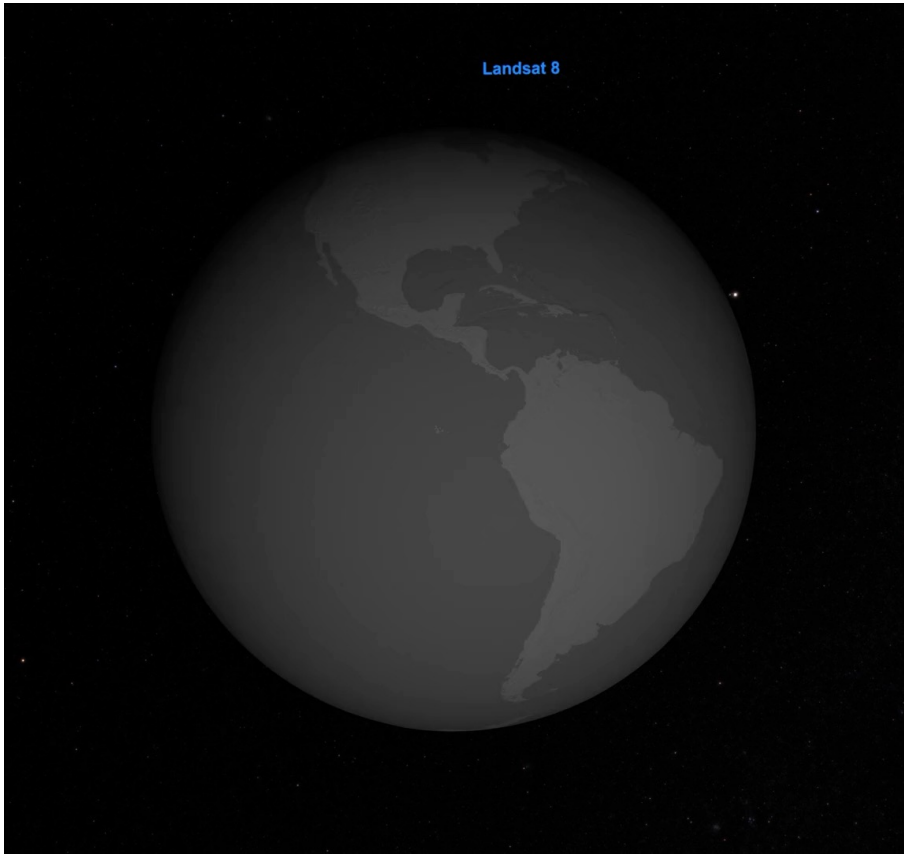
HLSS30 Band Information

Band Name	S30 Band	Wavelength (µm)	Units
Coastal Aerosol	B01	0.43 - 0.45	reflectance
Blue	B02	0.45 - 0.51	reflectance
Green	B03	0.53 - 0.59	reflectance
Red	B04	0.64 - 0.67	reflectance
Red-Edge 1	B05	0.69 - 0.71	reflectance
Red-Edge 2	B06	0.73 - 0.75	reflectance
Red-Edge 3	B07	0.77 - 0.79	reflectance
NIR Broad	B08	0.78 - 0.88	reflectance
NIR Narrow	B8A	0.85 - 0.88	reflectance
Water Vapor	B09	0.93 - 0.95	reflectance
Cirrus	B10	1.36 - 1.38	reflectance
SWIR 1	B11	1.57 - 1.65	reflectance
SWIR 2	B12	2.11 - 2.29	reflectance

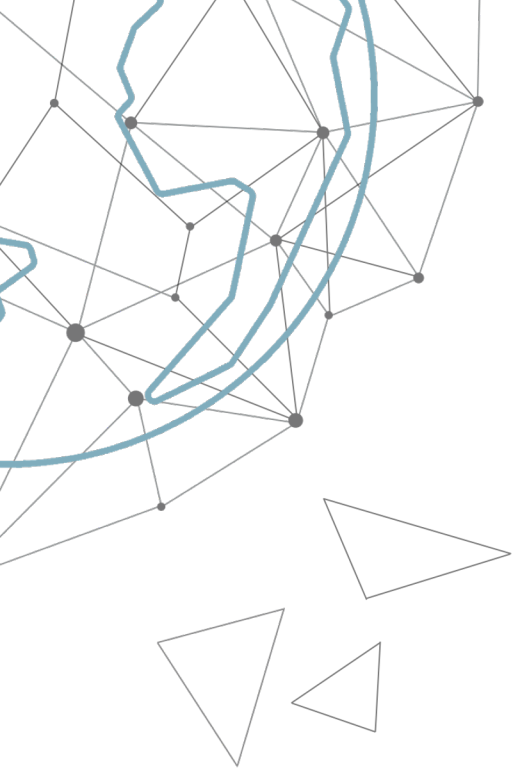
HLSL30 Band Information

Band Name	L30 Band	Wavelength (µm)	Units
Coastal Aerosol	B01	0.43 - 0.45	reflectance
Blue	B02	0.45 - 0.51	reflectance
Green	B03	0.53 - 0.59	reflectance
Red	B04	0.64 - 0.67	reflectance
NIR Narrow	B05	0.85 - 0.88	reflectance
SWIR 1	B06	1.57 - 1.65	reflectance
SWIR 2	B07	2.11 - 2.29	reflectance
Cirrus	B09	1.36 - 1.38	reflectance
Thermal IR 1	B10	10.60 - 11.19	degrees (C)
Thermal IR 2	B11	11.50 - 12.51	degrees (C)

Benefits of S2/L8/L9 Harmonization

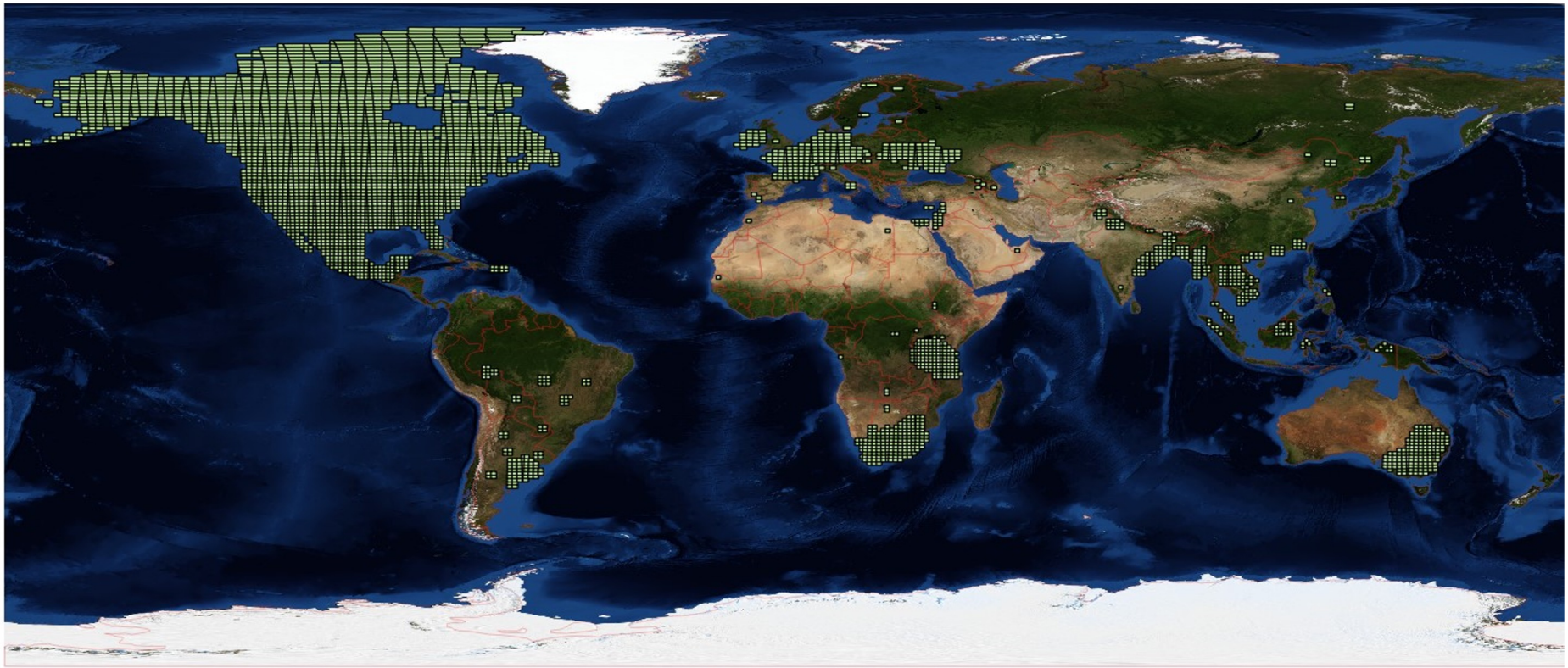


Credit: NASA SVS

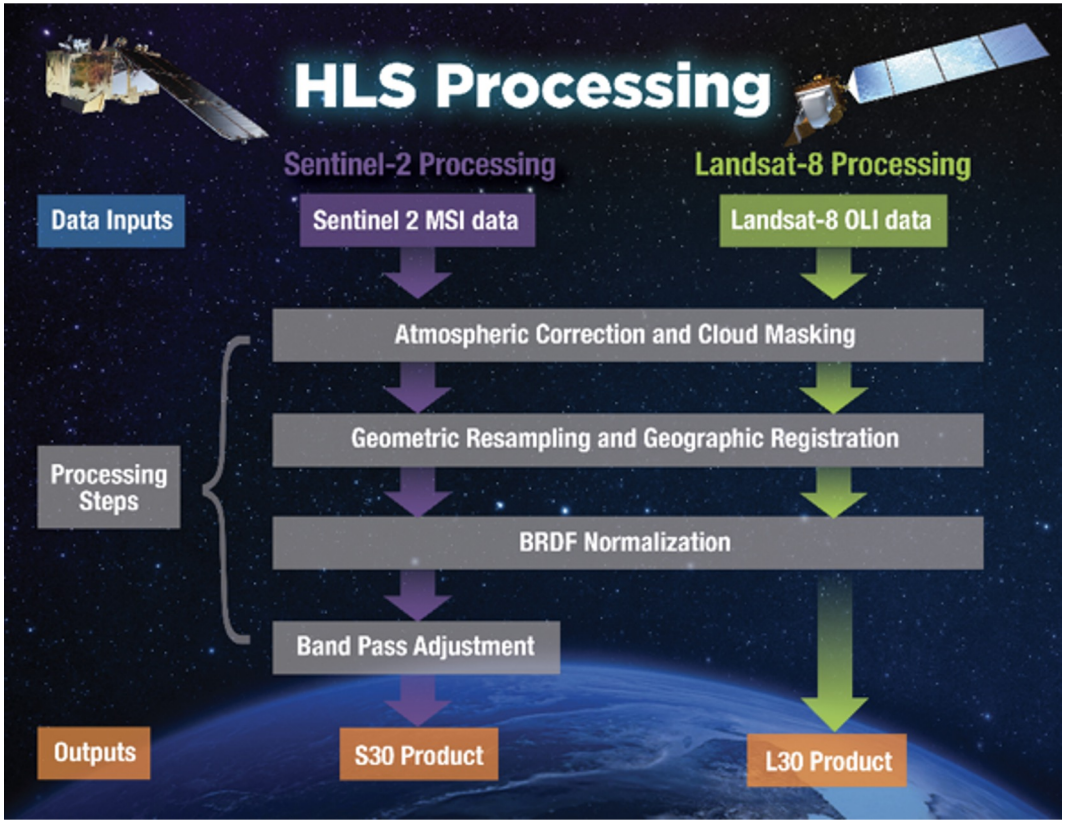


Scaling to Global HLS Data Production

HLS Coverage Prior to Global Implementation



HLS Algorithm Workflow



Band Pass Adjustment

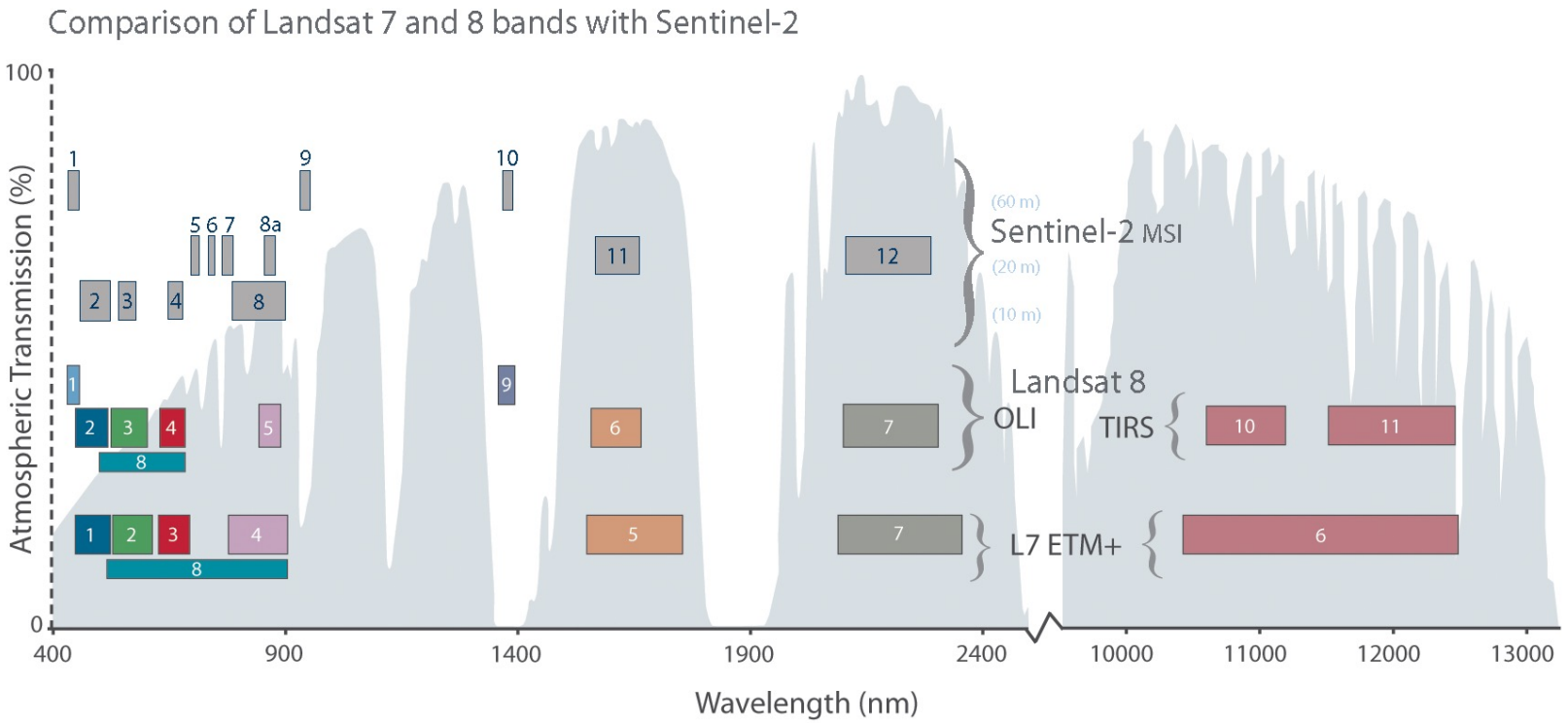
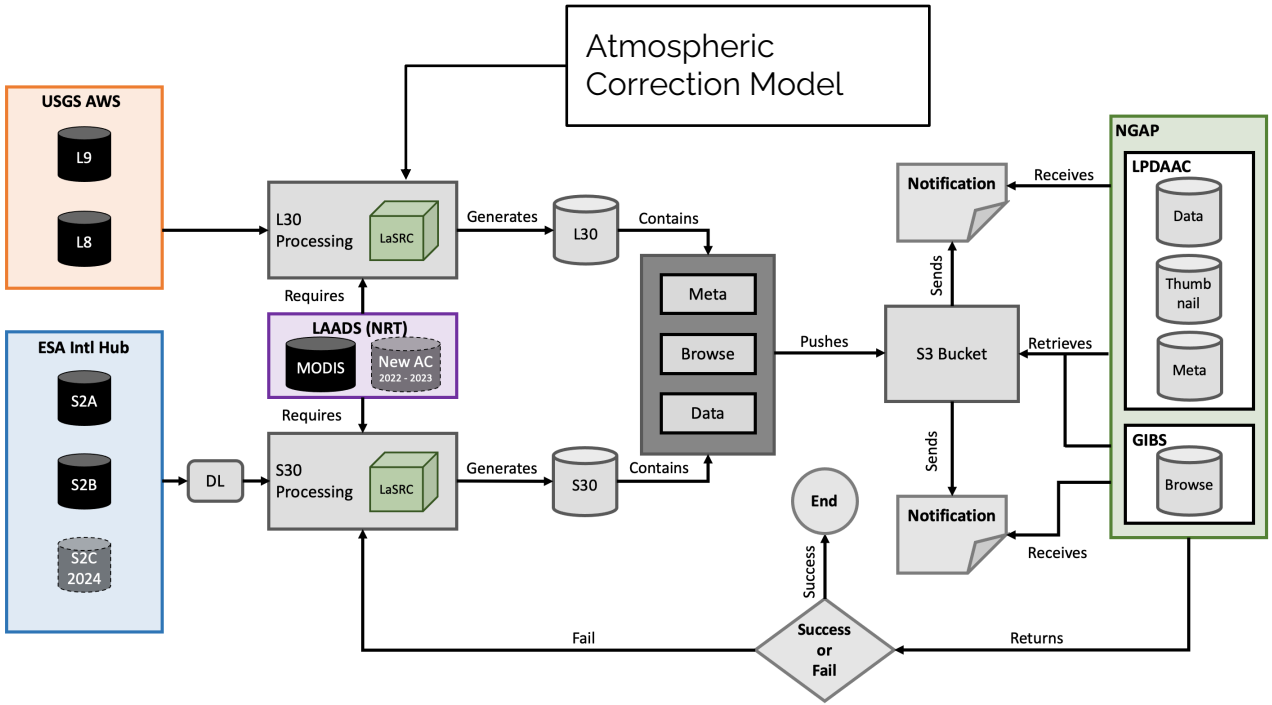


photo credit: USGS

“Global” HLS Coverage Area

Data Processing Workflow

1. Data downloaded or notifications received from external agencies.
2. Processing pipeline triggered with each new granule.
3. Access MODIS data from LAADS for atmospheric correction (2-3 day latency)
4. Generate data, metadata, and browse imagery.
5. Notify LPDAAC and GIBS of data availability with SQS message containing manifest of new files.

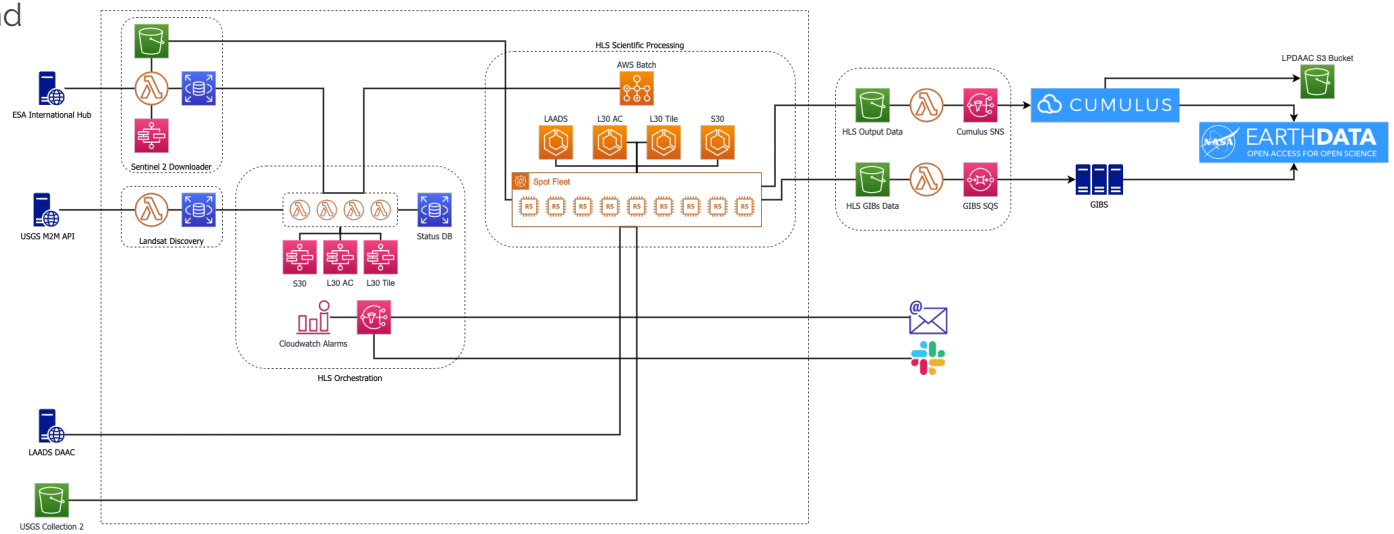


System Architecture

NASA IMPACT AWS: Stage data for LPDAAC and GIBS retrieval, notify LPDAAC and GIBS for data availability, 90-day temporary archive

DAAC AWS: Ingest, archive, and publication of HLS data products, full archive backup

NASA GIBS AWS: Ingest browse imagery and make available for NASA Worldview client

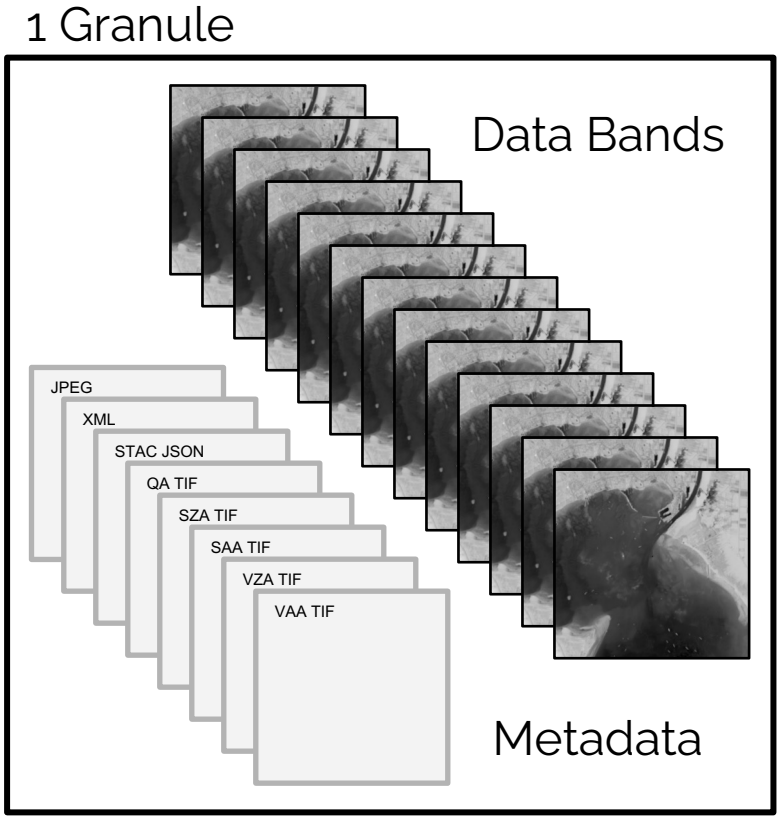


Infrastructure Demand for Processing Modes

- Forward
  - New data collected by USGS and ESA input into HLS production system
  - Typically runs 1 day at a time
- Historical
  - Archived USGS and ESA data input into HLS production system
  - Scales to n-days depending on timeline for completion

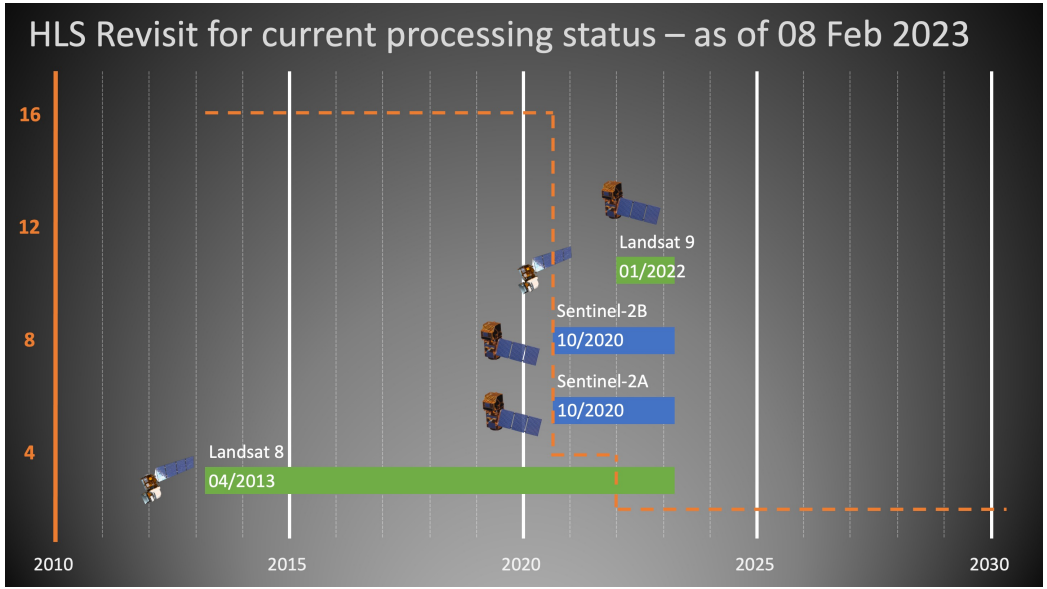
~15000 per day

150000+ per day



Current Status of Data Production

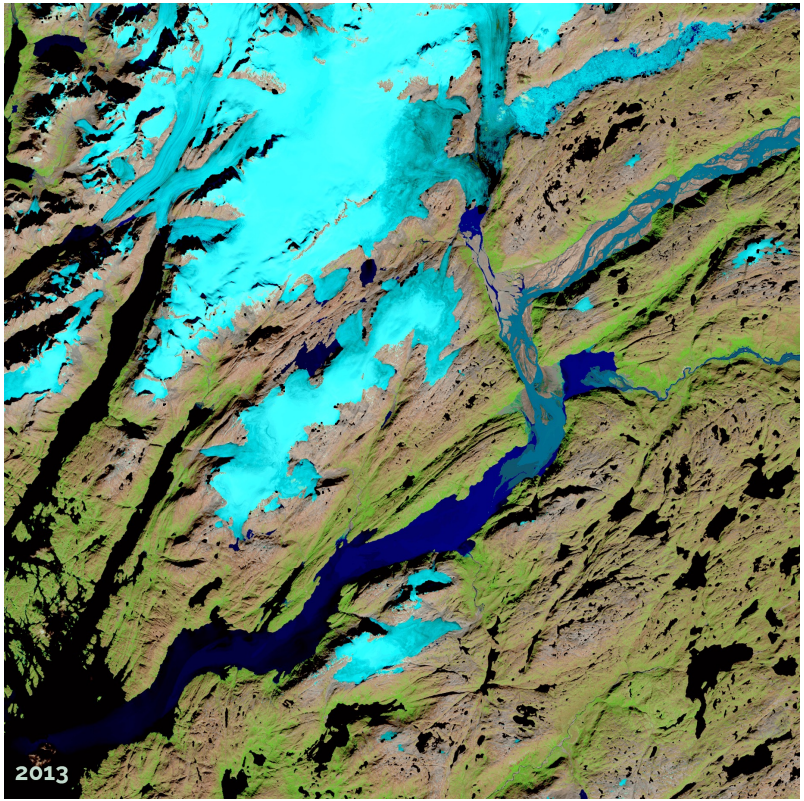
- Currently 4 satellites in virtual constellation
  - Landsat 8 – data available from 04/11/2013
  - Landsat 9 – data available from 01/01/2022
  - Sentinel-2A – data available from 09/01/2017
  - Sentinel-2B – data available from 09/01/2017



Dataset Characteristics

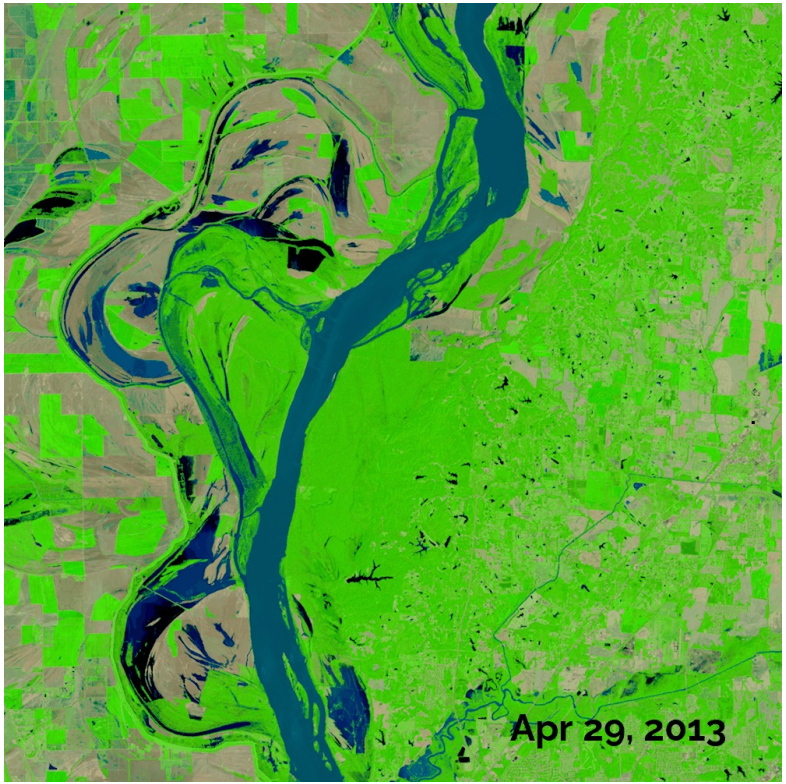
Resolution	30 meters
Latency	2-3 days
Format	Cloud-optimized GeoTIFF
Distribution	Band-separated files
Provider	LP DAAC

- Each granule includes:
  - 13 (10) band-separated data files
  - 4 angle bands
  - 1 QA band
  - 1 browse image
  - 1 xml metadata
  - 1 stac json

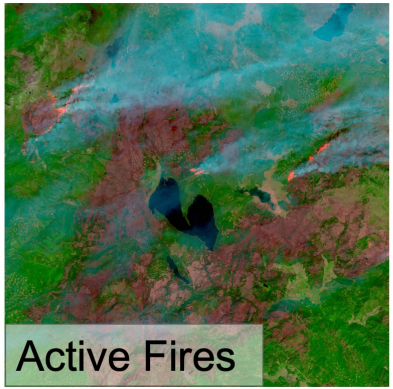


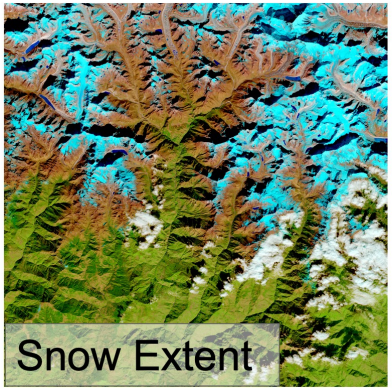
Dataset Statistics

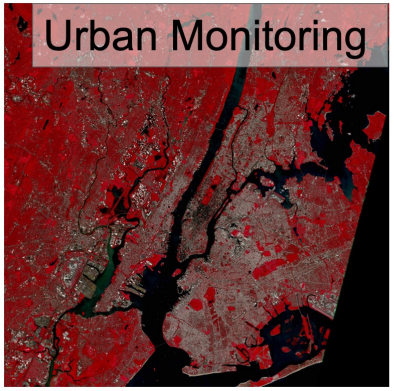
	HLSL30	HLSS30
Current data start date	11-Apr-2013	1-Sep-2017
Number of Granules	10.05M	12.31M
Average daily files	2719	5884
Average daily volume (TB)	0.88	1.91



HLS Science Applications



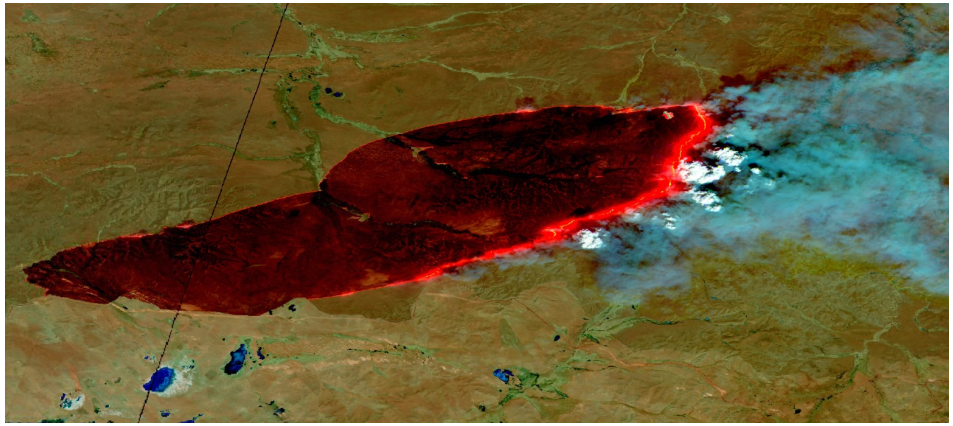




HLS SWIR False Color Composite (FCC) Imagery for FIRMS

- Available from March 15, 2022 – present on both US/Canada and Global FIRMS
- Dynamically generated imagery using COGs hosted in LP DAAC S3 bucket
    - Slower response times than pre-generated imagery as in GIBS

Satellite	L30	S30
FCC Band 1	7	12
FCC Band 2	5	8A
FCC Band 3	4	4



Thank you.