The Joint ESA-NASA Multi-Mission Algorithm and Analysis Platform: Next-Generation Collaboration Tool for Scientific Algorithms and Datasets

Laura Jewell[1], George Chang[1], Hook Hua[1], Manil Maskey[2], Rahul Ramachandran[3], Kaylin Bugbee[2], Laura Duncanson[4], Marco Lavalle[1], Aimee Barciauskas[5], Chris Lynnes[6], Clement Albinet[7], Amanda Whitehurst[8], and Bjoern Frommknecht[9]

[1]Jet Propulsion Laboratory
[2]University of Alabama in Huntsville
[3]NASA Marshall Space Flight Center
[4]University of Maryland
[5]Development Seed
[6]NASA Goddard Space Flight Center
[7]ESA
[8]ASRC Federal Holding Company
[9]European Space Research Institute

November 24, 2022

Abstract

The ESA-NASA Multi-mission Algorithm and Analysis Platform (MAAP) is a platform designed to meet the need of the international scientific community to collaborate on the generation and analysis of increasingly massive datasets from space-based, airborne, and field observations for aboveground terrestrial carbon dynamics. The MAAP is an open-source, cloud-based platform that distinguishes itself from other science platforms by being agnostic to any science disciplines and allows scientists to write, develop, and execute their own algorithms in shared workspaces that are tailored to their specific area of research. We present an overview of the capabilities of the Pilot implementation of MAAP. Users can explore and visualize ESA and NASA data that is ingested in the MAAP data catalog, develop and test algorithms to generate new datasets and act upon existing ones, launch matured algorithms as large-scale processing jobs, and analyze the results. This allows scientists to follow the entire scientific process within the platform, from the conception of a hypothesis to the creation of scientific results in a version-controlled and reproducible environment. Users can document their process, collaborate with other users, and share algorithms and datasets. We will conclude with an outlook on the features that the next phase of development will bring to the platform.

The Joint ESA-NASA Multi-Mission Algorithm and Analysis Platform (MAAP):
Next-Generation Collaboration Tool for Scientific Algorithms and Datasets

AGU Fall Meeting
December 9, 2019
IN11D-0692

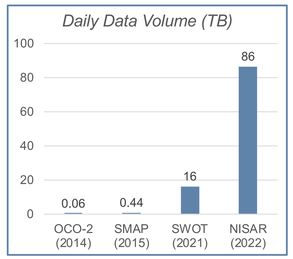
Laura Alisic Jewell[1] (Laura.A.Jewell@jpl.nasa.gov), George Chang[1], Hook Hua[1], Manil Maskey[2], Rahul Ramachandran[2], Kaylin Bugbee[3], Marco Lavalle[1], Laura Duncanson[4], Aimee Barciauskas[5], Chris Lynnes[6], Amanda Whitehurst[7], Björn Frommknecht[8], Clément Albinet[8]

[1]Jet Propulsion Laboratory, California Institute of Technology; [2]NASA Marshall Space Flight Center; [3]University of Alabama Huntsville; [4]University of Maryland, College Park; [5]DevelopmentSEED; [6]NASA Goddard Space Flight Center; [7]ASRC Federal Holding Company; [8]European Space Agency

Summary

- Multi-Mission Algorithm and Analysis Platform (MAAP): science-focused, cloud-based environment to discover, process, analyze, and share NASA and ESA satellite, airborne, and field data
- ESA-NASA collaboration focuses on NISAR, BIOMASS, and GEDI missions, for improving understanding of aboveground terrestrial carbon dynamics
- Cloud-based architecture to co-locate algorithms and analysis with the data, instead of bringing data to users
- Potential blueprint for future ESA/NASA collaboration and governance of scientific data, algorithms, and software
- Project currently in Pilot phase, next development phase underway

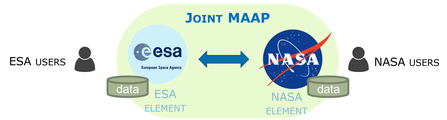
Why Do We Need MAAP?



- New satellite missions being launched produce larger and larger data volumes
- Need scaling up of algorithms and processing for large datasets
- Remote sensing retrieval algorithms are often not sufficiently mature; require multi-mission fusion efforts
- Need co-location of complementary data from multiple missions
- Scientific community needs improved data and algorithm sharing and collaboration across users and organizations

International Collaboration

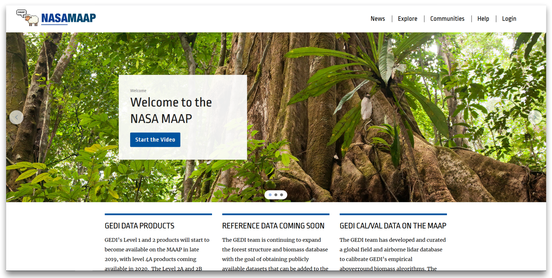
- Key objective is to improve how we share data and algorithms across organizations
- Common entry point, separate ESA and NASA implementations
- ESA-NASA interoperability for user access management, data catalog, API; currently being extended to algorithms and visualization



Key Concepts of MAAP

- Establish analysis collaboration framework to share data, algorithms, compute resources to accelerate development cycle
- Decouple algorithm development, cataloging, large scale processing deployment
- Provide interfaces to develop, version, and share algorithms and data products
- Open-source, cloud-based, with software re-use
- Agnostic to science discipline, allows scientists to write, develop, and execute their own algorithms in shared workspaces tailored to specific area of research

MAAP Capabilities & The Scientific Process



- Scientists can follow the entire scientific process within the platform: From conception of a hypothesis to creating scientific results in version-controlled and reproducible environment
- Users can document their process, collaborate with other users, and share algorithms and datasets

Explore & visualize data in catalog

↓

Collaboratively develop, modify & test algorithms

↓

Perform large-scale processing for matured algorithms

↓

Analyze results

↓

Share data, algorithms, documentation, and results; interact with community

Use Case: Terrestrial Biomass Investigation

EXPLORE ***RUN*** ***DEVELOP***

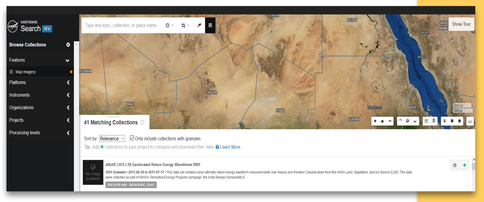
Multi-sensor search and comparison

Image visualization and exploration

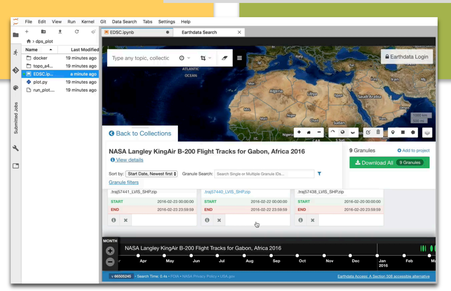
Subsetting to different formats and zonal statistics

Biomass product generation and validation with user-provided *in situ* data

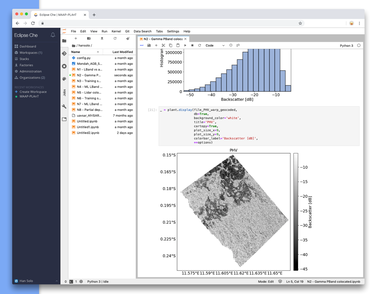
Algorithm code development



Search for and explore data in the MAAP data catalog (based on Common Metadata Repository and EarthData Search Client)



Interact with and map data queries and their results from within Jupyter notebooks, using the MAAP API and Jupyter extensions



Run and share Eclipse Che workspaces with tailored Jupyter environments to develop, test, and mature algorithms, run processing jobs, and visualize results

Datasets in MAAP

Pilot:

- Airborne and field campaigns to support Cal/Val for GEDI, NISAR, BIOMASS missions: LVIS, UAVSAR, AfriSAR, BioSAR, TropiSAR, INDREX2, PALSAR, ALOS
- Ancillary data used in biomass algorithms: SRTM, Sentinel-2, GPM, GlobCover, Landsat 7

Next phase:

- GEDI, NISAR, BIOMASS products
- Additional supporting data

Future Work

- Develop algorithms supporting Analysis-Ready Datasets (ARDs) for fast analysis
- Enable on-demand analysis on ARDs using multiple frameworks
- Provide access to data products from GEDI, BIOMASS, NISAR as they become available
- Establish cloud resource cost model and governance framework for algorithm and data sharing

References

Related AGU 2019 Fall Meeting presentation:
[IN53A-02] *Bugbee et al*, "Creating the Future: The Joint ESA-NASA Multi-Mission Algorithm and Analysis Platform's Data Ecosystem"

Albinet et al, "A Joint ESA-NASA Multi-mission Algorithm and Analysis Platform (MAAP) for Biomass, NISAR, and GEDI", Surveys in Geophysics, 2019; https://doi.org/10.1007/s10712-019-09541-z

Acknowledgements

We are grateful for the development work by the ESA and NASA developer teams, and for the support provided by Kevin Murphy and Henri Laur. This work was in part conducted at the Jet Propulsion Laboratory (JPL), California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA).