

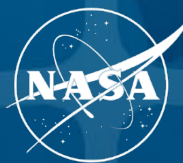


A cloud-based, science enabling system for visualizing and analyzing data at a DAAC

Geoffrey T. Stano, Navaneeth Selvaraj, and Will Ellett

GHRC DAAC Earthdata Webinar

October 16, 2024



- **Presenters**
- **GHRC Overview**
- **Problems the Field Campaign Explorer (FCX) Addresses**
- **FCX Overview**
- **FCX Demonstration**
- **FCX Playground**
- **Jupyter Notebooks**
- **Future Work**

Today's Presenters

- **Dr. Geoffrey Stano (left)**

- GHRC DAAC Scientist – coordinating combination of atmospheric and computer science
- Research to applications focus
- 20 years experience in lightning applications

- **Mr. Navaneeth Selvaraj (right)**

- GHRC lead front-end developer – leading team of developers on projects, such as the Field Campaign Explorer
- Emphasis on open source activities
- 8 years experience in software development



Dr. Geoffrey Stano



Mr. Navaneeth Selvaraj

NASA's Earth Science DAACs and GHRC



- **Distributed Active Archive Centers (DAAC)**

- NASA's Earth Science Data and Information System (ESDIS)

- **Role**

- Process, archive, document, and freely distribute Earth Science data
- Enable the use of these data by users in their research

- **GHRC**

- Global Hydrometeorology Resource Center
- 1 of 12 NASA DAACs
- Collaboration between NASA Marshall Space Flight Center and the University of Alabama in Huntsville

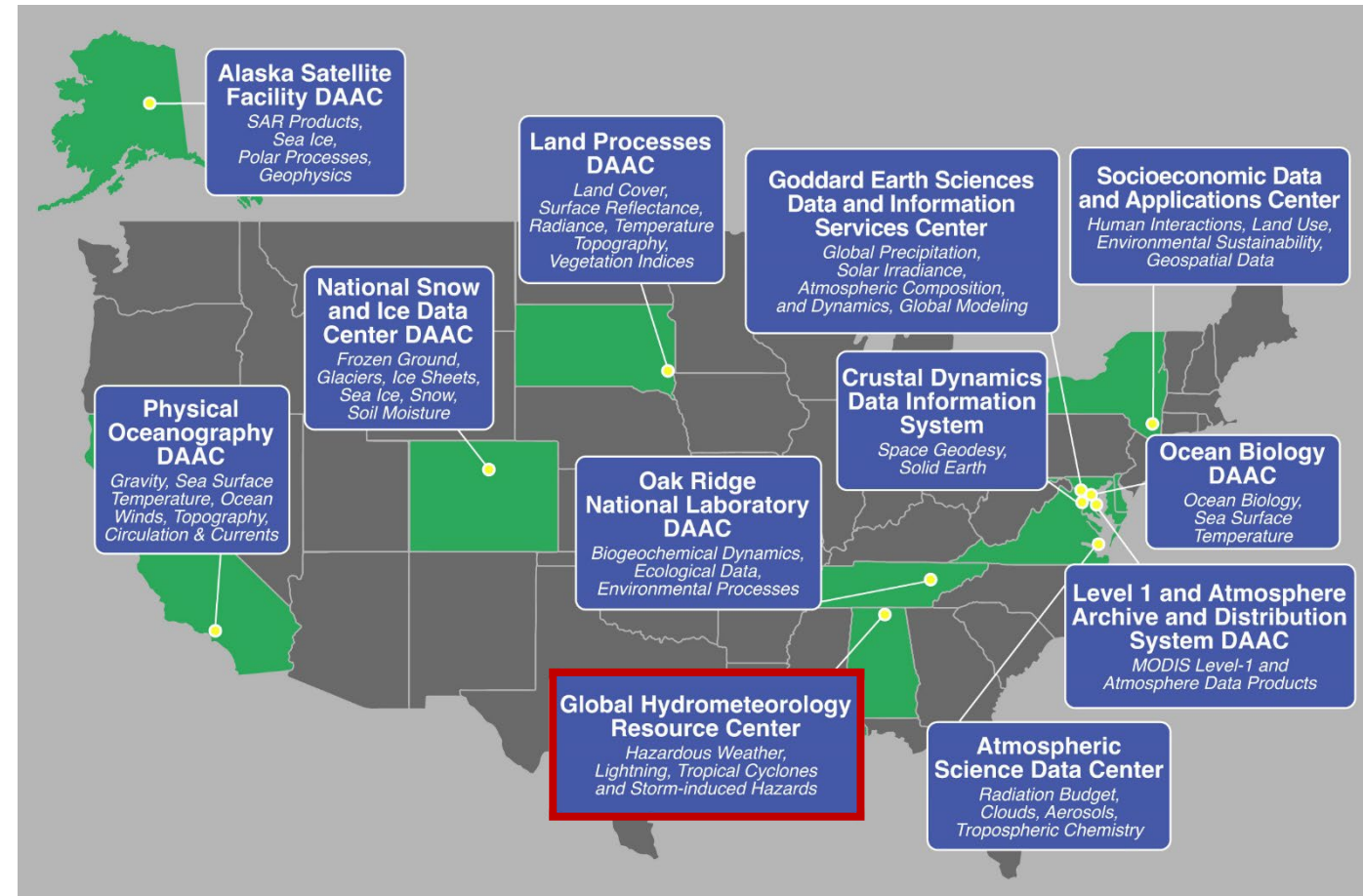


Image courtesy NASA ESDIS

GHRC Mission and Holdings



To provide a comprehensive archive of datasets for the analysis of dynamical and physical processes of storm hazards, lightning, precipitation, tropical systems, and field campaigns. Emphasize cloud-based tools, science expertise, and open science enabling users to more fully use GHRC's unique holdings.

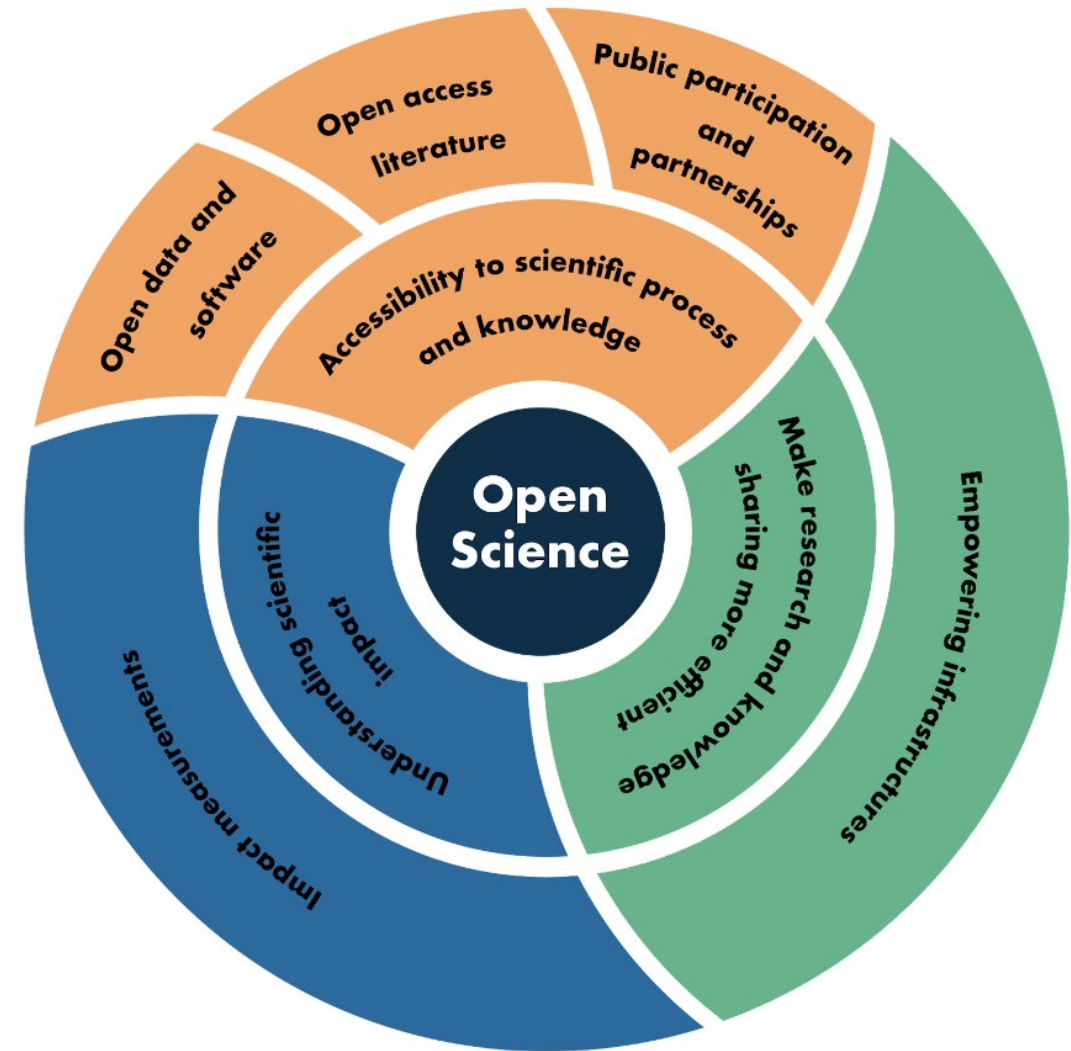


• Data Holdings (600+ datasets with 150+ TB of data)

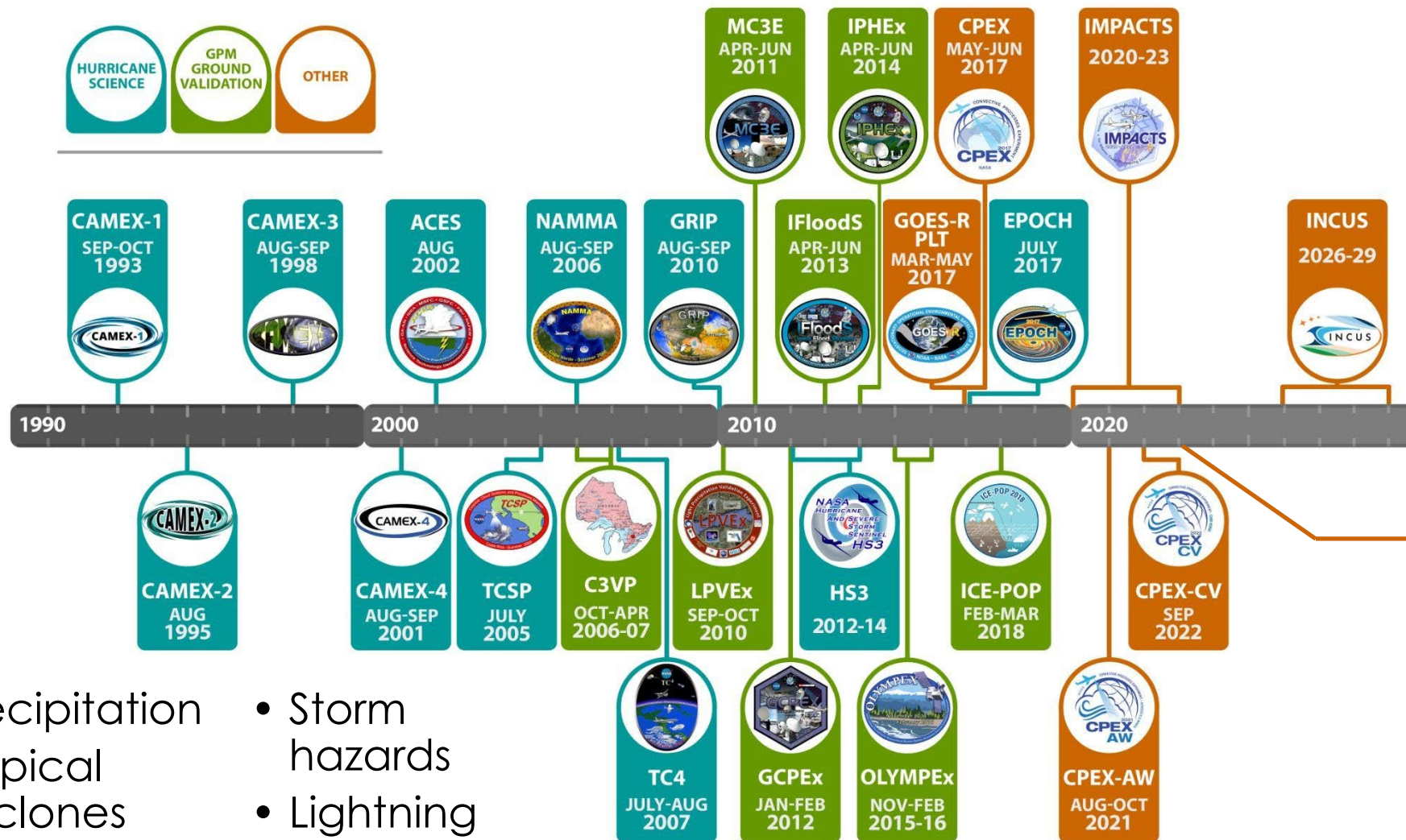
- Precipitation
- Hurricanes
- Storm hazards
- Lightning
- Field campaigns

What Are We Trying To Address?

- Advocate for open science
- Data stewardship and expertise
- Community outreach and participation
- Technology development to support accessibility, sharing, and communicating
- Expertise in airborne / field campaign data, information, and knowledge sharing
- Field Campaign Explorer (FCX) is one approach GHRC uses to enable open science



GHRC's Field Campaign and Mission Holdings



- Precipitation
- Tropical cyclones
- Storm hazards
- Lightning

Just approved for archival!



The Field Campaign Explorer (FCX)



- 3D data exploration tool for visualization and analysis of diverse, coincident datasets
- Originally built for field campaigns, it is not limited to this now
- Open source – No downloads
- **25** different instruments visualized
- Available in NASA GitHub
 - <https://github.com/nasa/GHRC-FieldCampaign-eXplorer-UI>
 - <https://github.com/nasa/GHRC-FieldCampaign-eXplorer-core>
- Published to Python Package Index (PyPI)
 - <https://pypi.org/user/ghrc/>

See Separate Media File Recording

Recording of FCX display abilities

Challenges FCX Addresses

- **Data Complexity**

- Challenging to handle massive, multi-dimensional Earth science datasets

- **Fragmented Data Systems**

- Lack of seamless integration across diverse NASA data

- **Visualization Gaps**

- No standardized tool for multi-source 3D rendering

- **Analytical Limitations**

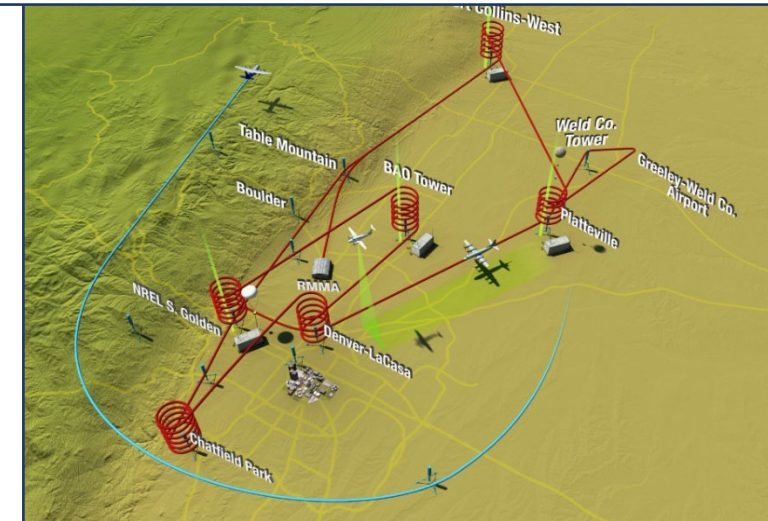
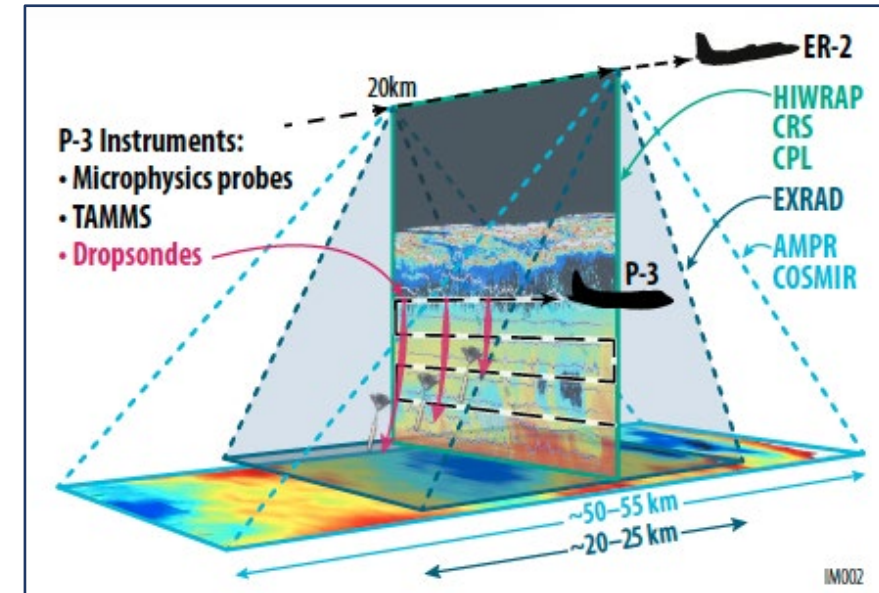
- Insufficient integration of advanced analytics

- **Accessibility and Open Science**

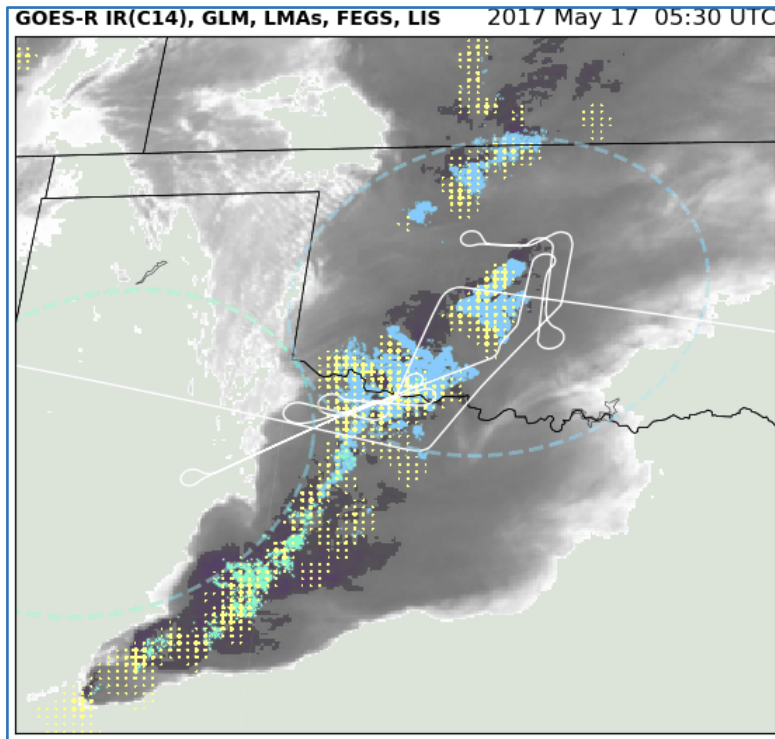
- Need for a cloud-based, open-source platform to support collaboration

- **Performance Issues**

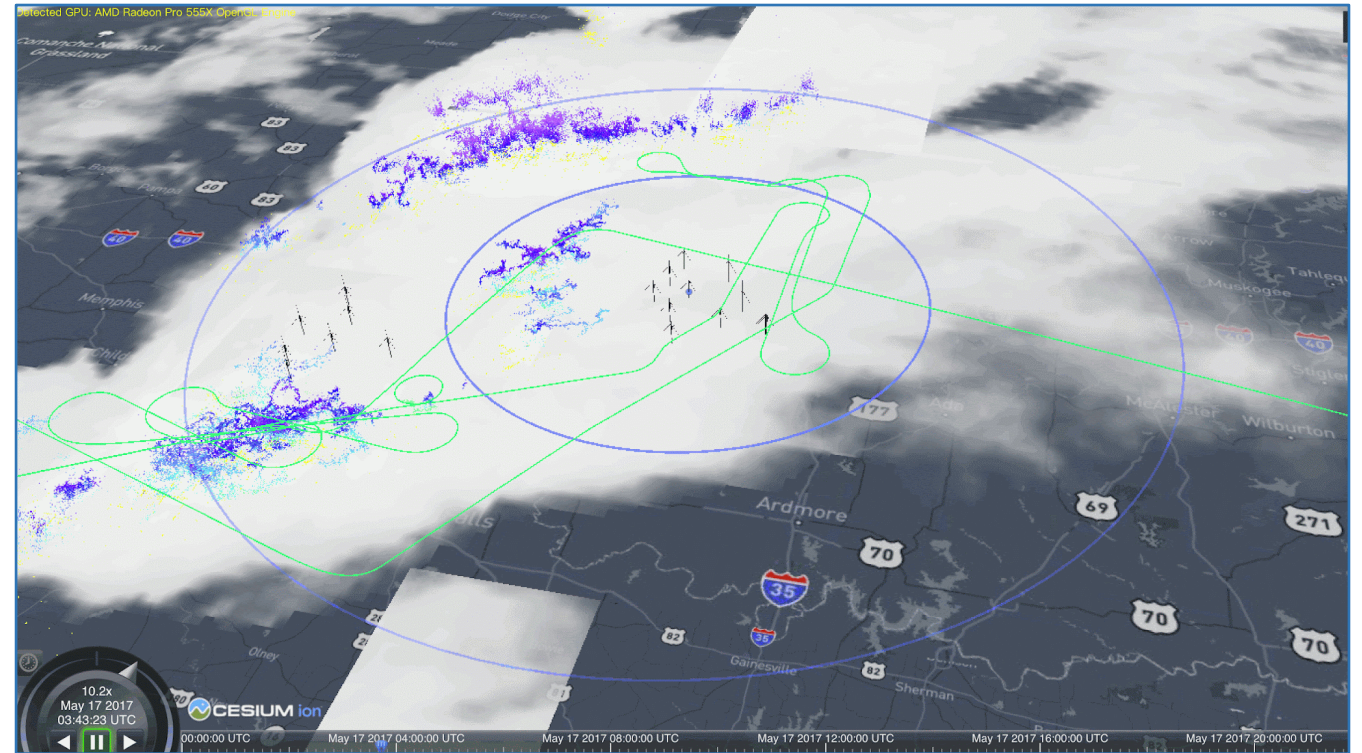
- Current tools lack GPU acceleration for high-resolution, real-time visualizations



2D versus 3D Display



2D, non-interactive animation of lightning observed from various coincident measurements.



The same scene but with a 3D, interactive animation/exploration from FCX (running in the cloud)



- **Modularized extension of FCX**
- **Allows users to interact with individual visualization components as standalone elements**
- **Simplifies learning process**
 - Users can understand and develop FCX components effectively
- **Amazon Web Server environment**
 - Data streaming via Earthaccess
 - Dataset transformation with Pangeo Forge
 - NetCDF metadata viewer

- **Earthaccess**

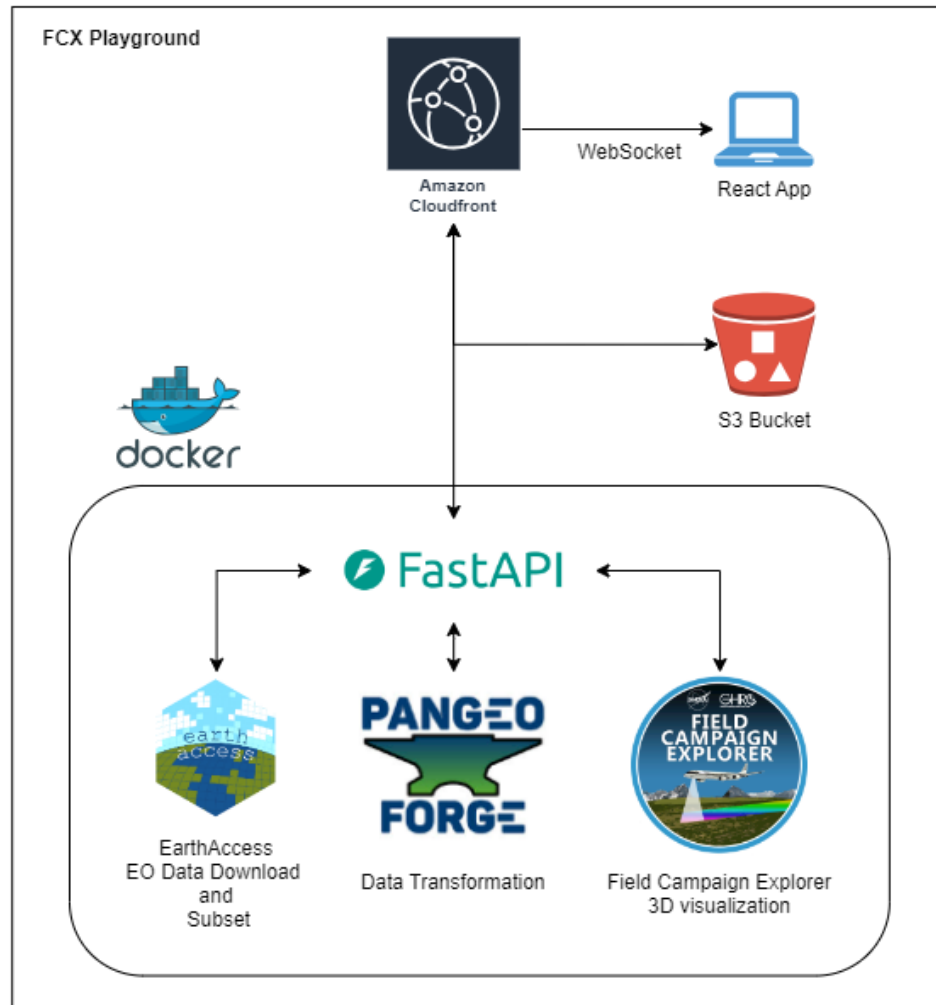
- Python library to search for and download/stream NASA Earth science data
- <https://github.com/nsidc/earthaccess>

- **Pangeo Forge**

- Open-source platform for data Extraction, Transformation, and Loading (ETL)
- Extract data from traditional data repositories and deposit in cloud object storage in analysis-ready, cloud-optimized (ARCO) formats
- <https://github.com/pangeo-forge>

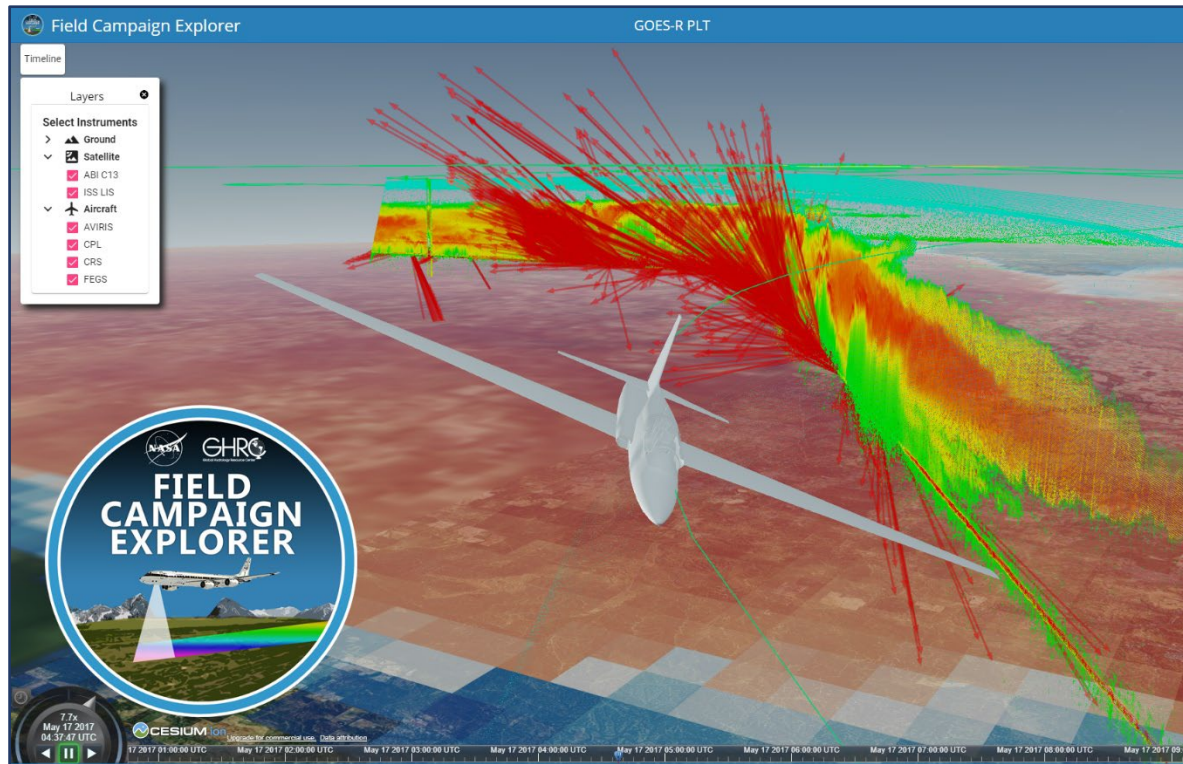
- **Field Campaign Explorer (FCX)**

- 3D data exploration tool to visualize and analyze diverse, coincident datasets
- Can support more than just field campaigns



- **Containerized services using Docker for consistent and scalable deployments**
- **Fast API orchestrates:**
 - Earth observation data access
 - Data transformation
 - 3D visualizations
- **S3 Bucket manages processed datasets with CloudFront Content Delivery Network (CDN)**
- **React App provides real-time data access and visualization through WebSocket**

FCX Playground User Benefits



Early example of FCX with electric field mill (arrows), Advanced Baseline Imager 10.35 micron imagery (below), cloud radar (curtain), and ER-2 flight track (green line).

- Better understand FCX
- Integrates multiple NASA Earth science tools into a single, cohesive platform for seamless data processing
- Simplifies deployment with an end-to-end data processing experience
- Provides comprehensive user interface for 3D visualization and geospatial analysis
- Reduces costs and development time by combining functionalities into one application

Earth Access – Input Form

FCX Playground

Visualizations

- czml
- 3dTile
- wmts
- pointPrimitive
- subsettingTool
- histogramTool
- earthaccess
- pangeoforge

INPUT FORM STATUS FILE PATH METADATA

Short Name
GPM_3IMERGDF

Start date
11 / 19 / 2022

End date
11 / 30 / 2022

Bounding Box (Xmin)
-180

Bounding Box (Ymin)
-90

Bounding Box (Xmax)
180

Bounding Box (Ymax)
90

Concept id

UPLOAD GEOJSON FILE SUBMIT

- Enter Collection Short name
- Temporal and Spatial Filter
- Concept ID (Optional)
- Upload Shapefiles for advanced spatial filter
- Submit

Earth Access - Status



FCX Playground

Visualizations

- czml
- 3dTile
- wmts
- pointPrimitive
- subsettingTool
- histogramTool
- earthaccess
- pangeoforge

INPUT FORM STATUS FILE PATH METADATA

GPM_3IMERGDF - 8/5/2024, 11:20 AM ^

Querying data

In progress - downloading files 1/3

Uploading 3B-DAY.MS.MRG.3IMERG.20221119-S000000-E235959.V07B.nc4 to S3

In progress - downloading files 2/3

Uploading 3B-DAY.MS.MRG.3IMERG.20221120-S000000-E235959.V07B.nc4 to S3

In progress - downloading files 3/3

Uploading 3B-DAY.MS.MRG.3IMERG.20221121-S000000-E235959.V07B.nc4 to S3

Job completed

GPM_3IMERGDF - 8/5/2024, 11:17 AM ∨

MCD43A4 - 7/29/2024, 3:21 PM ∨

CLEAR

- Track progress of the data download

Metadata Viewer for NC Files

FCX Playground

Visualizations

- czml
- 3dTile
- wmts
- pointPrimitive
- subsettingTool
- histogramTool
- earthaccess
- pangeoforge

INPUT FORM STATUS FILE PATH METADATA

GPM_3IMERGD - 8/5/2024, 12:26 PM ^

No granules found

GPM_3IMERGDF - 8/5/2024, 11:20 AM ^

coords

lon: {"dims":["lon"],"attrs":{"units":"degrees_east","long_name":"Longitude"},"dtype":"float32","shape":[3600]}

lat: {"dims":["lat"],"attrs":{"units":"degrees_north","long_name":"Latitude"},"dtype":"float64","shape":[1800]}

time: {"dims":["time"],"attrs":{"units":"days since 1980-01-06 00:00:00Z","standard_name":"time","long_name":"time","bounds":"time_bnds"},"dtype":"float64","shape":[1]}

attrs

BeginDate: "2022-11-19"

BeginTime: "00:00:00.000Z"

EndDate: "2022-11-19"

EndTime: "23:59:59.999Z"

- View NC files metadata to understand the dataset

Demo - Video



See Separate Media File Recording

Jupyter Hub Notebooks in the Cloud



- **Modular Design**
 - Make for easy experimentation
- **Customizable Code**
 - Users can use or code with their data
- **Data Recipes**
 - Expanding library to simplify analysis workflows
- **Real-time Collaboration**
 - Multiple users working together in the cloud
- **Scalable Resources**
 - Leverage cloud infrastructure for high-performance computing
- **Accessible anywhere with internet connectivity**

See Separate Media File Recording

Example notebook output of the ER-2 flight navigation.

<https://github.com/ghrcdaac/ghrc-playground>

Looking Forward (FY25)



- **Interactive Integrated Development Environment for updating visualizations**
- **Integrate lightning datasets and capabilities into the Playground**
 - GLM / LMA tools (Dr. Eric Bruning – Texas Tech)
 - Existing GHRC Jupyter notebooks
- **Add GPU acceleration and machine learning libraries for visualizations and analysis**
- **Include GraphQL API to efficiently access and query metadata**
- **Add Python code for more plotting libraries**
- **Integrate other ESDIS tools to make the FCX Playground more robust and cohesive**

Looking Forward (FY25 - continued)

- GHRC will be the DAAC to archive data from the ALOFT campaign and the Investigation of Convective Updrafts (INCUS) mission
- GHRC working with both science teams to integrate their data into FCX
- Opportunity to enable FCX growth via user collaborations



<https://ghrc.earthdata.nasa.gov/fcx>

<https://ghrc.earthdata.nasa.gov/fcx-playground>



THANK YOU!

Dr. Geoffrey Stano – geoffrey.stano@uah.edu

Navaneeth Selvaraj – ns0066@uah.edu

