National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology

NASA's Open Source Integrated Analytic Platform for Air Quality and Wildfire

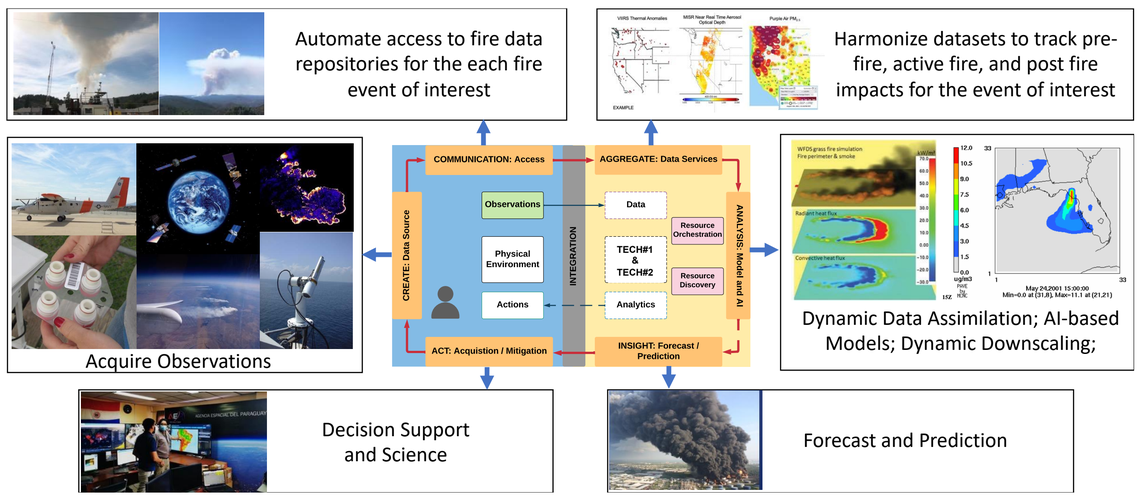
EARTH SCIENCE TECHNOLOGY OFFICE

Sina Hasheminassab[1], Abhishek Chatterjee[1], Nga Chung[1], Gary Doran[1], Thomas Huang[1], Olga Kalashnikova[1], Riley Kuttruff[1], Kyo Lee[1], Grace Llewellyn[1], Kevin Marlis[1], Wai Phyo[1], Joe Roberts[1], Jeanne Holm[2], Mohammad Pourhomayoun[3], Dawn Comer[2], Chaowei Yang[4], Seren Smith[4], Anusha Srirenganathanmalarvizhi[4]

[1] Jet Propulsion Laboratory, California Institute of Technology [2] City of Los Angeles [3] California State University, Los Angeles [4] George Mason University

- **AIST Fire Alarm Platform:**
  - Integrates the AIST Air Quality Analytic Collaborative Framework (AQ ACF) with the AIST Integrated Digital Earth Analysis (IDEAS) digital twins framework.
  - Focuses on the impact of wildland fire on air pollution, providing data, machine learning (ML), and numerically driven assimilation for air quality analysis, visualization, and prediction.
- **Analytic Collaborative Framework (ACF)**:
  - Concept for scalable computational and data infrastructures to harmonize data, tools, and computation resources.
  - Enables scientific investigations by integrating various technologies and optimizing data partitions for parallel analytics.
- **Apache Science Data Analytics Platform (SDAP)**:
  - A professional open-source solution under ACF.
  - Combines big data technologies for Earth science, optimized for cloud or on-premises computing clusters.
  - Offers capabilities like satellite and model data analysis, in situ data integration, fast data subsetting, and OpenSearch.
- **Capabilities of SDAP**:
  - Comprises two main subsystems: Analysis Optimized Storage (manages data partitions for parallel analytics) and Services and Workflows (provides analytic capabilities and workflows).
  - Automates pre-computed statistics (e.g., min/max, standard deviation, regional averages) for predefined time windows.
- **Visualization and Analytics**:
  - Leverages OGC standards-compliant solutions (OnEarth, MapServer, GDAL) similar to NASA's GIBS.
  - Provides robust visualization capabilities for complex data sets.
  - Streamline data onboarding and image generation.
  - Expand data analytics by combining satellite, model, and ground measurements.
  - Explore low-cost cloud-based storage for accessing and analyzing large remote sensing data sets.
- **Public Accessibility**:
  - Offers a one-stop Air Quality Information Platform for researchers and decision-makers.
  - Features web-based tools and downloadable Jupyter notebooks.
  - Supports public and private cloud or on-premises cluster deployments, with RESTful APIs for applications.

Professional Open-Source Earth System Digital Twin Framework



https://ideas-digitaltwin.jpl.nasa.gov/airquality/

Current Dataset

Satellites

Sentinel 5P TROPOMI
- Carbon Monoxide (CO)
- Methane (CH_4)
- Sulphur Dioxide (SO_2)
- Ozone (O_3)
- Nitrogen Dioxide (NO_2)

MODIS Aqua & Terra
- Aerosol Optical Depth (AOD)

Suomi NPP VIIRS
- Deep Blue Aerosol Angstrom Exponent
- Deep Blue Aerosol Optical Thickness

OCO-3
- Snapshot Area Maps XCO_2

EMIT
- Methane Plumes

In-Situ Sensors

PurpleAir ($PM_{2.5}$)
AirNow (CO, O_3, NO, NO_2, PM_{10}, $PM_{2.5}$)

QuantAQ (CO, O_3, CO_2, NO, NO_2, PM_{10}, $PM_{2.5}$)

Models

High Performance GEOS-Chem (GCHP)
- Methane (CH_4)
- Nitrogen Oxide (NO)
- Nitrogen Dioxide (NO_2)
- Sulphur Dioxide (SO_2)
- Ozone (O_3)

MERRA-2
- Planetary Boundary Layer Height (PBLH)
- Surface Air Temperature (TLML)
- Surface Exchange Coefficient for Heat (CDH)

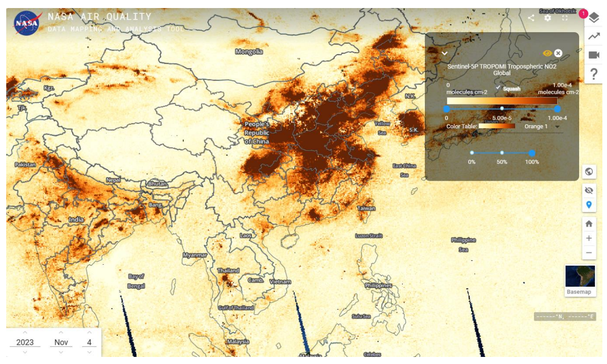
GEOS Composition Forecasts (GEOS-CF)

TROPESS Chemical Reanalysis (O_3)

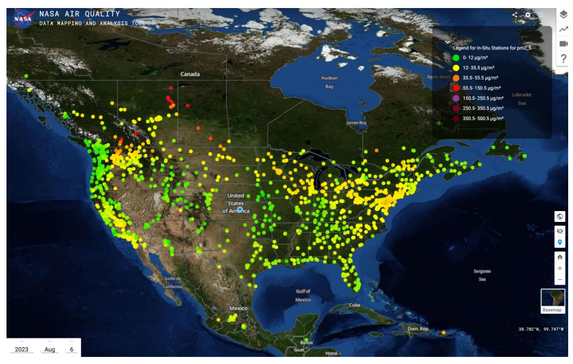
Surface $PM_{2.5}$ from Washington Univ of St. Louis

Integrated Air Quality Information Platform

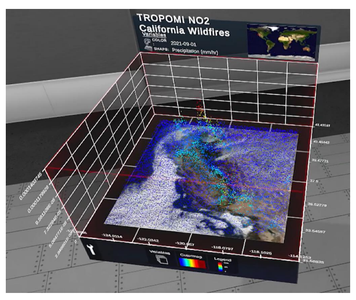
Web-based Interactive Data Analysis Tool

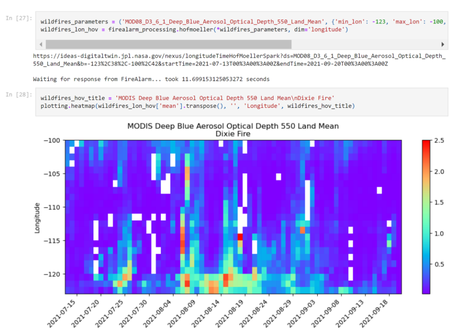


Immersive Science



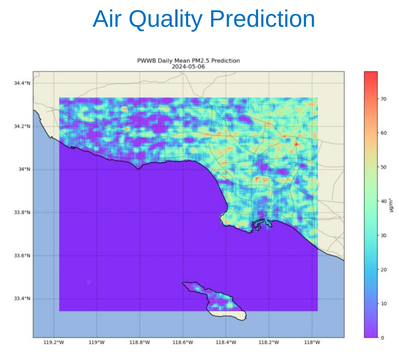
Interactive Jupyter Notebook

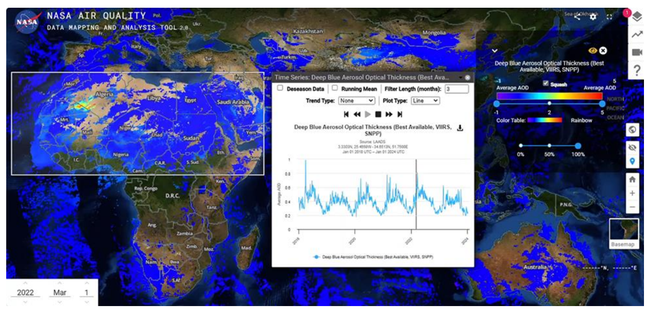


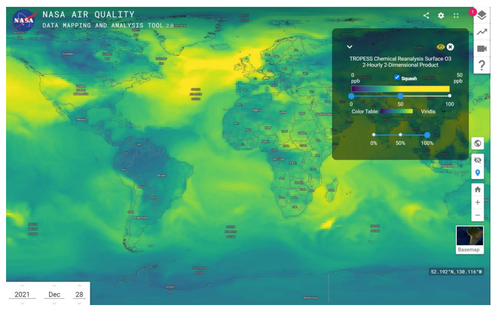


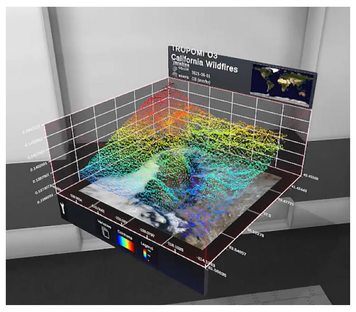
ML/AI Models

Air Quality Prediction









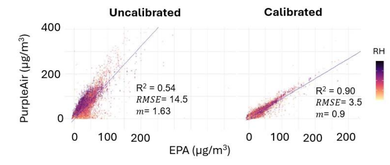
Satellite and Model data inventory

```
In [3]: aq_datasets = firealarm_processing.get_datasets()
In [4]: pd.concat([aq_datasets[['GlobalAttrTitle', 'GlobalAttrSource']].head(), aq_datasets[['GlobalAttrTitle', 'GlobalAttrSource
```

Out[4]:

	GlobalAttrTitle	GlobalAttrSource
0	MERRA-2 Global CDH	GES DISC
1	MERRA-2 Global PBLH	GES DISC
2	MERRA-2 Global TLML	GES DISC
3	MODIS Terra AOD 550 Dark Target Deep Blue Comb...	Giovanni
4	MODIS Terra Aerosol Optical Depth Land Ocean Mean	Giovanni
24	Sentinel-5P TROPOMI Methane CH4 California	GES DISC
25	Sentinel-5P TROPOMI Carbon Monoxide CO California	GES DISC
26	Sentinel-5P TROPOMI Tropospheric NO2 California	GES DISC
27	Sentinel-5P TROPOMI Total Ozone California	GES DISC
28	Sentinel-5P TROPOMI Sulphur Dioxide SO2 Califo...	GES DISC

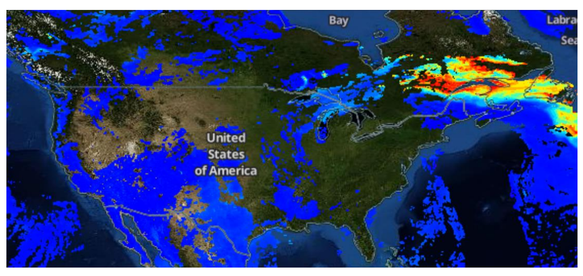
PurpleAir Calibration
(LSTM, RNN, DNN, RainFarm, etc.)



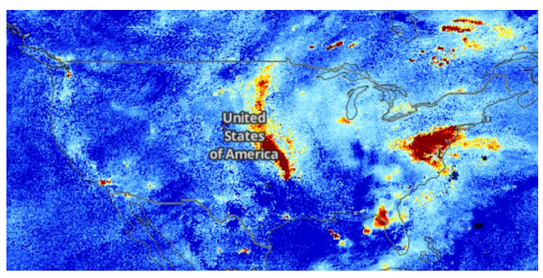
Air Quality & Greenhouse Gases Use Cases

2023 Canadian Wildfire
(June 2023)

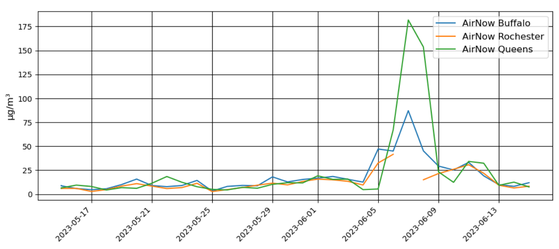
VIIRS Aerosol Optical Thickness



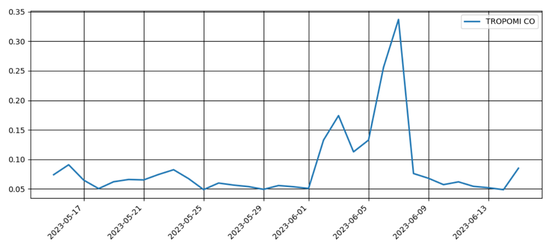
TROPOMI Tropospheric NO_2 Vertical Column



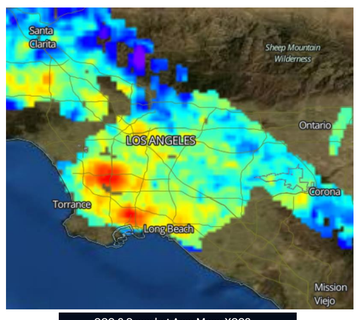
Time-series of in-situ $PM_{2.5}$ from selected AirNow sensors in NE US



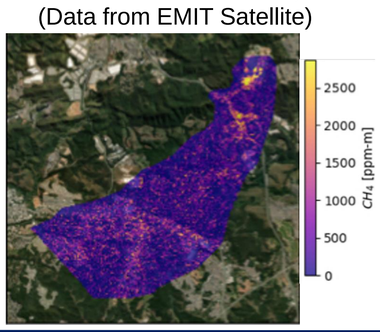
Time-series of TROPOMI CO averaged over NE US



CO_2 over Los Angeles, CA
(Data from OCO-3 Satellite)



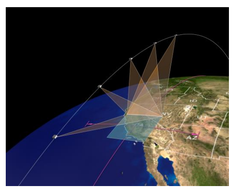
Methane Plume in São Paulo, Brazil
(Data from EMIT Satellite)



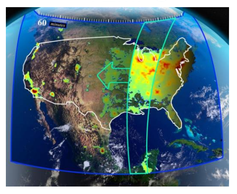
Current Developments & Future Directions

New Satellite Data

MAIA

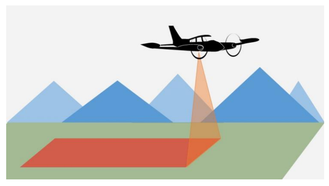


TEMPO



Field Campaigns Data

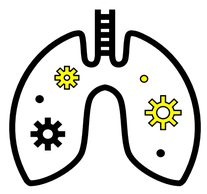
Aircraft measurements; mobile monitoring



Wildfire Prediction, Spread Projection, and Air Quality Impact

Socioeconomic Data for Environmental Justice Analysis

Health Effects Analysis



Scenario-based Analysis