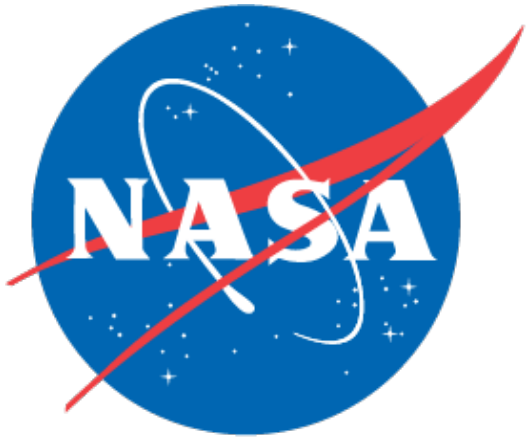


HEALTH & AIR QUALITY

EARTH SCIENCE
APPLIED SCIENCES

Supporting local government public health and air quality decision-making with a sub-city scale air quality forecasting system from data fusion of models, satellite, in-situ measurements, and low-cost sensors

PI: K. Emma Knowland

Science PIs: Carl Malings, Nathan Pavlovic

HAQ Applications Annual Team Meeting, March 29, 2023

Project Summary

EARTH SCIENCE
APPLIED SCIENCES

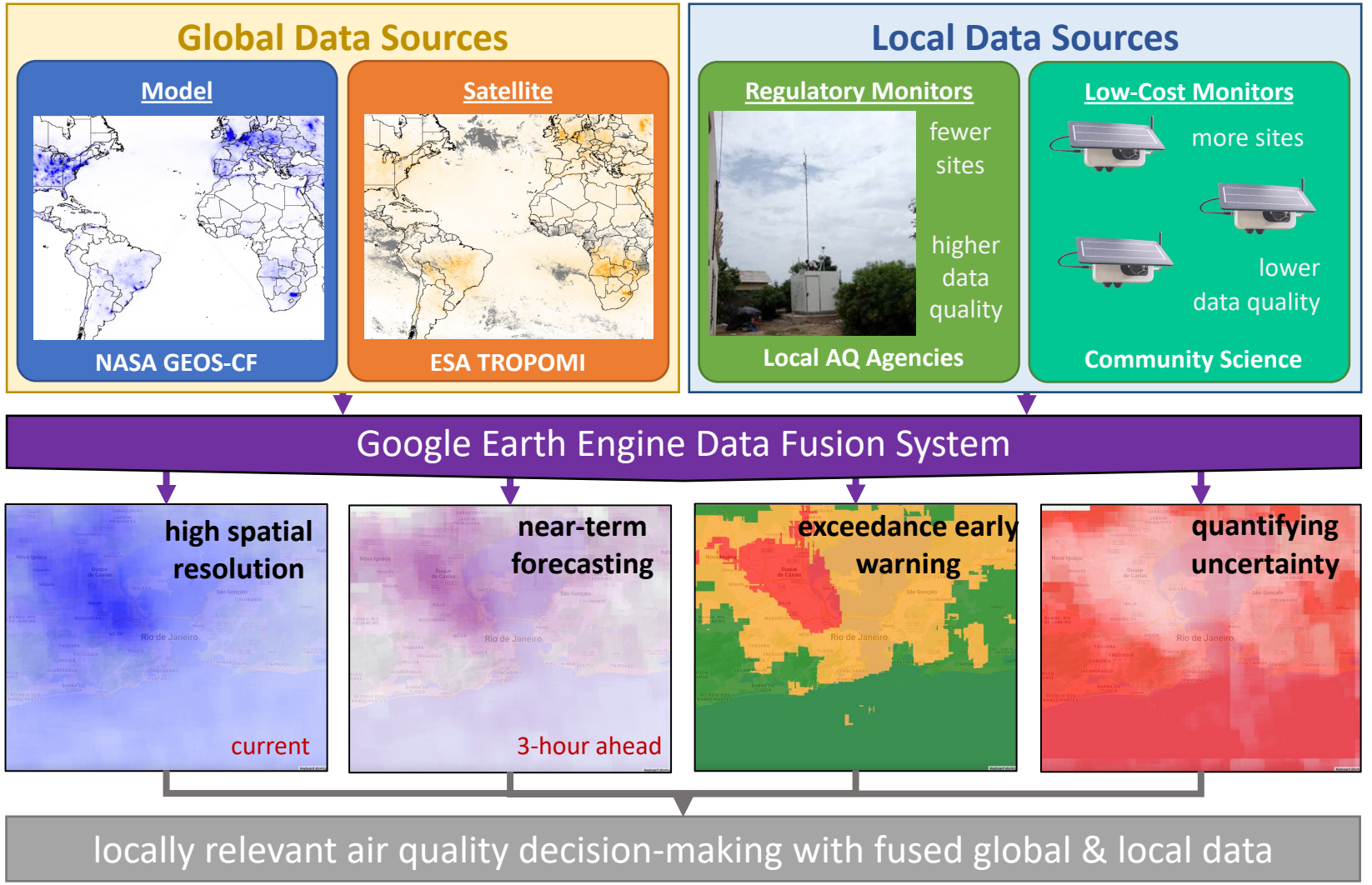
This project's goal is to…

…integrate diverse **global** and **local** air quality data sources…

…using the cloud computing platform of **Google Earth Engine**…

…to provide synthesized **estimates** and **forecasts** of air quality at a **local scale** but with a **global scope**…

…which will be freely accessible by air quality managers worldwide, facilitating their **decision-making** processes.



Project Partners/Collaborators

EARTH SCIENCE
APPLIED SCIENCES

Role	Name	Affiliation	Organization Type
Science PI	Carl Malings	Morgan State University/GESTAR-II, NASA GSFC GMAO	University, Co-operative Agreement
Science PI	Nathan Pavlovic	SonomaTech, Inc	Private Sector
Co-I, Institutional PI	Stephen Cohn	NASA GSFC GMAO	Government
Co-I	Christoph Keller	Morgan State University/GESTAR-II, NASA GSFC GMAO	University, Co-operative Agreement
Collaborator	John White	U.S. Environmental Protection Agency (EPA)	Government
Collaborator	Sean Khan	United Nations Environment Programme (UNEP)	Intergovernmental organization
Collaborator	Randall Martin	Washington University in St. Louis	University
Collaborator	Dan Westervelt	Lamont-Doherty Earth Observatory	University
Collaborator	Sean Wihera	Clarity	Private Sector

Project End-users & Stakeholders

EARTH SCIENCE
APPLIED SCIENCES

Organization Name	Organization Type	Decision Making Activity
City of Rio de Janiero, Brazil (Instituto Pereira Passos, Prefeitura do Rio de Janeiro; Contacts: Felipe Mandarino & Bruno Boscaro)	Government	• Advance warning system for potential air-quality related hospital visits • Detecting “invisible” pollutant sources not included in current inventories
City of Dakar, Senegal (Contacts: Aminata Mbow Diokhané [CGQA], Ngongang Danube [UNEP])	Government	• Analysis and forecasting for transportation policy AQ impacts • Including dust forecasts for early warning
US EPA (Contact: John White)	Government	• Incorporating low-cost sensors into their Air Quality forecasting and information systems (with associated uncertainties)

Engagement plan and recent updates

Project “kick-off” meetings held with end-users to confirm decision-making support needs.
Targeted follow-up meetings to establish data pipelines and better quantify anticipated outcomes.

Schedule & Milestones

GMAO: Knowland, Malings, Keller, Cohn | STI: Pavlovic, AQ scientists, software engineers | US EPA (White) & UNEP (Khan) supporting end-users

* with LDEO (Westervelt) & Clarity (Wihera)
+ with WUSTL (Martin)

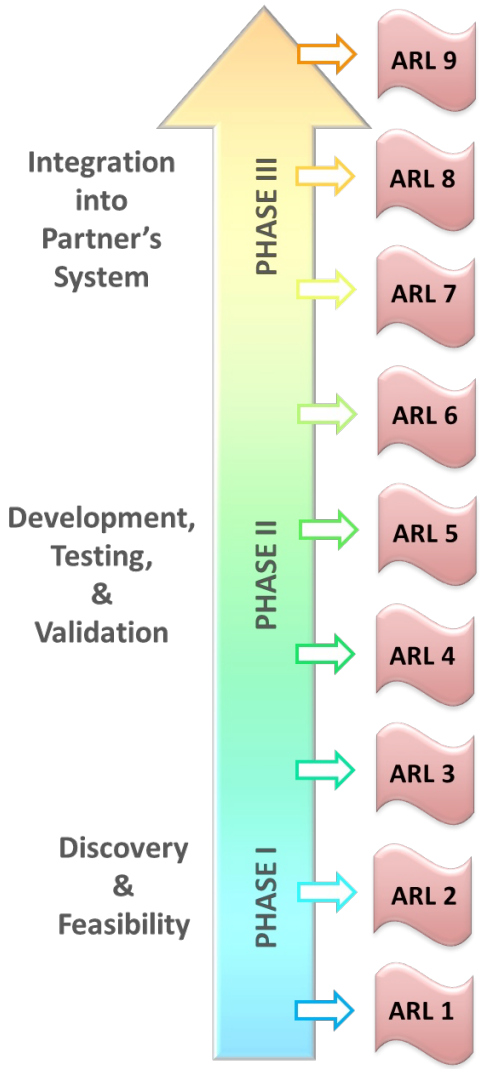
Objective & Tasks		Year 1		Year 2		Year 3	
1	a. Import new datasets to GEE, including GEOS-CF and RGM and LCS data						
	b. Merge existing GMAO and UNEP/STI data fusion methodologies						
	c. Refine data fusion system, including uncertainty quantification capabilities+						
	d. Incorporate LCS in the data fusion system, with uncertainty quantification*						
	e. Implement data fusion system into GEE tool, including interfaces						
	f. Refine GEE tool, especially input/output capability and interfaces						
2	a. Assess the status and identify key end-user needs for the GEE tool						
	b. Train end-users in the functionality and capabilities of the GEE tool*						
	c. Pilot deployment of GEE tool in end-user domains of interest						
3	a. Validate data fusion system in end-user domains of interest+						
	b. Integrate GEE tool into end-user decision-making processes						
	c. Evaluate the project's impact on decision-making outcomes						
	d. Integrate GEE tool with US EPA AirNow-Tech and UNEP GEMS Air systems						
	e. Transfer GEE tool to US EPA, UNEP, end-users for operational use						
Anticipated Application Readiness Level (ARL) metric		2	3	4	5	6,7	8,9

Year 1 through Q2: Most tasks within Objective 1 have started. For Objective 2, email correspondence with end-users to establish next steps

Progress on Project Tasks

EARTH SCIENCE
APPLIED SCIENCES

- Task 1a: Developed list of data sets and sources for GEE ingest. GEOS-CF replays are available and GEOS-CF forecasts datasets are in the process of being imported to GEE.
- Task 1b: A conceptual framework for the combined data fusion methodology has been created.
- Task 1c: A conceptual framework for uncertainty quantification in the combined methodology has been created.
- Task 1d: A prototype of LCS integration was tested, results presented at ASIC 2022 conference.
- Task 1e: Developed draft framework application specification. A prototype of the data fusion method has been implemented in GEE, using synthetic ground monitor data.
- Task 2a: Project "kick-off" meetings with end-users confirmed decision-making support needs; next steps include targeted follow-ups to establish data pipelines and quantify metrics for assessing outcomes.



EARTH SCIENCE APPLIED SCIENCES

ARL Performance

- Start-of-Project ARL = 2 *(July 26, 2022)*
  - Fundamental data fusion methodology published in peer-reviewed journal: Malings et al. 2021: https://doi.org/10.1029/2021EA001743
  - Nathan Pavlovic (Sonoma Technologies, Inc.) had implemented a similar methodology in Google Earth Engine on behalf of end-user UNEP
- Goal ARL = 9
- Current ARL = 3 *(as of March 29, 2023)*
  - Proof of Application Concept
  - Prototype data fusion system used to demonstrate potential application capabilities & use-cases to end users
  - Feedback from end-users about their most important decision-making tasks to be supported by the tool

Current ARL-Supporting Evidence

EARTH SCIENCE APPLIED SCIENCES

ARL 3, Milestone 1: Components of the application tested and validated independently

Task 1a: Import new datasets to GEE, including GEOS-CF and RGM and LCS data

Data Set Type	Data Sources	Status	Notes
Satellite Data Sets	MODIS AOD	**Active GEE Ingest**	Instrument may be retired
	VIIRS AOD	**Not Ingested**	Potential replacement for MODIS AOD
	TROPOMI NO2	**Active GEE Ingest**	
Model Data Sets	CAMS	**Active GEE Ingest (PM2.5, AOD)**	Total column (CO, O3, NO2, AOD components), No surface concentrations beyond PM2.5.
	GEOS-CF Replay	**Active GEE Ingest**	
	GEOS-CF Forecast	**GEE Ingest in Development**	GMAO working with GEE to ingest
RGM Data Sets	Publicly-available data sets	**Reviewed sources. Some data available in Google Cloud; others available through public API.**	Data refresh frequency in Google Cloud likely insufficient for this application.
	End-user data sets	**For discussion with stakeholders.**	
LCS Data Sets	Publicly-available data sets	**Reviewed sources and assessed potential ingest pathway.**	Need to assess which sensors and sources are priority for end-user locations.
	End-user data sets	**For discussion with stakeholders.**	

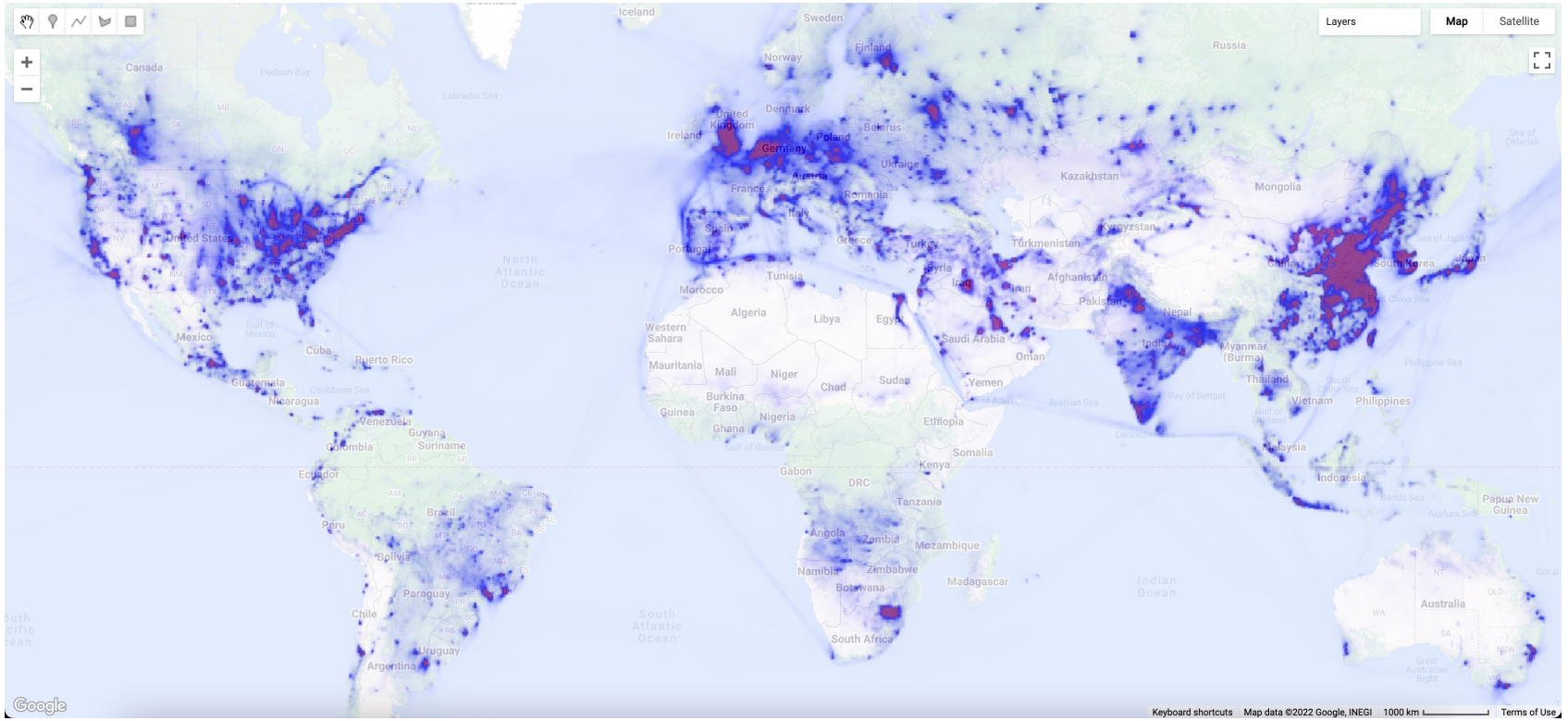
A preliminary list of key data sets, characteristics, and real-time sources has been developed. The list of data sources will be refined to meet the needs of methodologies and reflect available stakeholder datasets and needs.

Current ARL-Supporting Evidence

EARTH SCIENCE
APPLIED SCIENCES

ARL 3, Milestone 1: Components of the application tested and validated independently

Task 1a: Import new datasets to GEE, including GEOS-CF and RGM and LCS data



GEOS-CF Surface NO_2, average week of Sept 8-15, 2022, visualized in Google Earth Engine

GEOS-CF historical estimate files ("replay" files) are currently available on Google Earth Engine as ee.ImageCollection("NASA/GEOS-CF/v1/rpl/tavg1hr").

Next steps:
GEE is working on pulling in GEOS-CF forecast data. GMAO has early access to test data in algorithms.

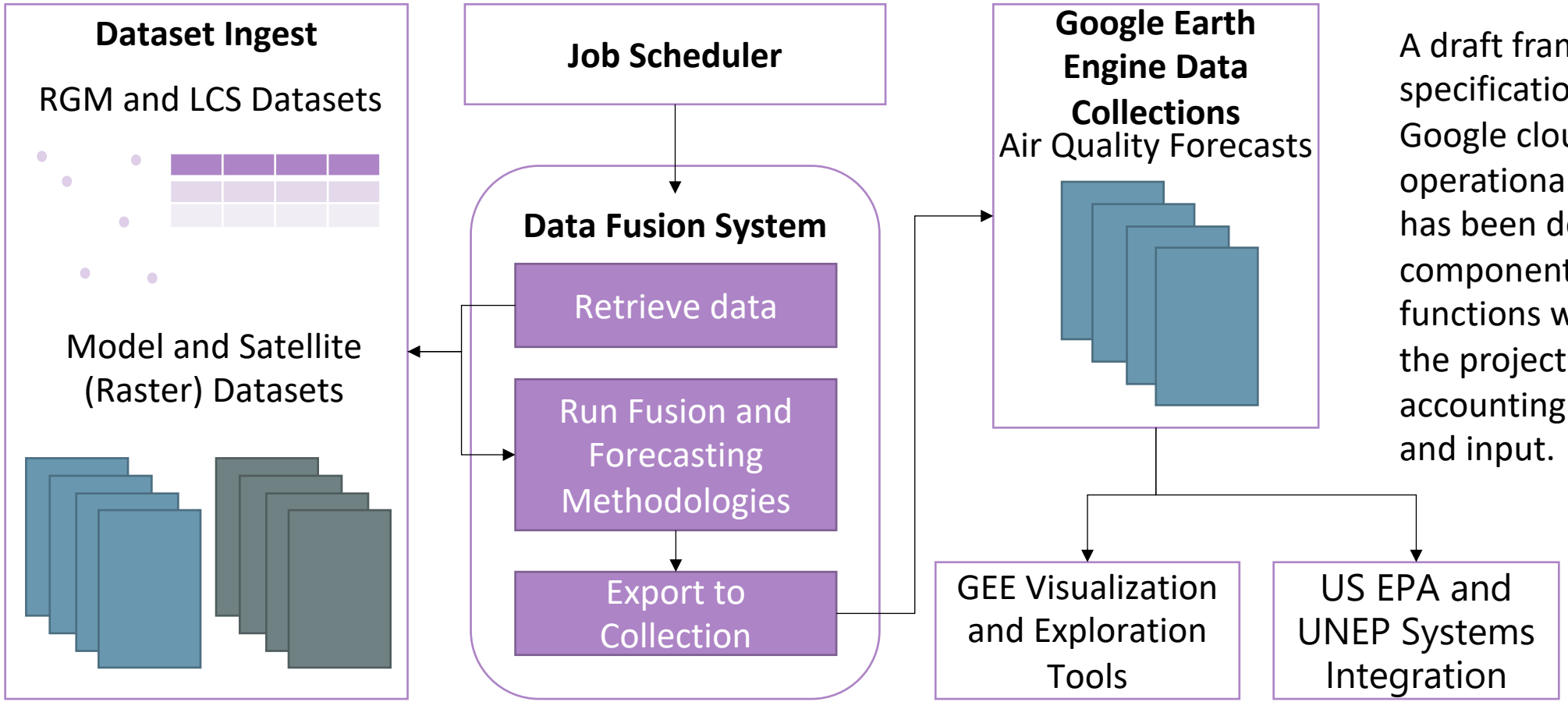
Need ground-based observations in GEE

Current ARL-Supporting Evidence

EARTH SCIENCE APPLIED SCIENCES

ARL 3, Milestone 1: Components of the application tested and validated independently

Task 1e: Implement data fusion system into GEE tool



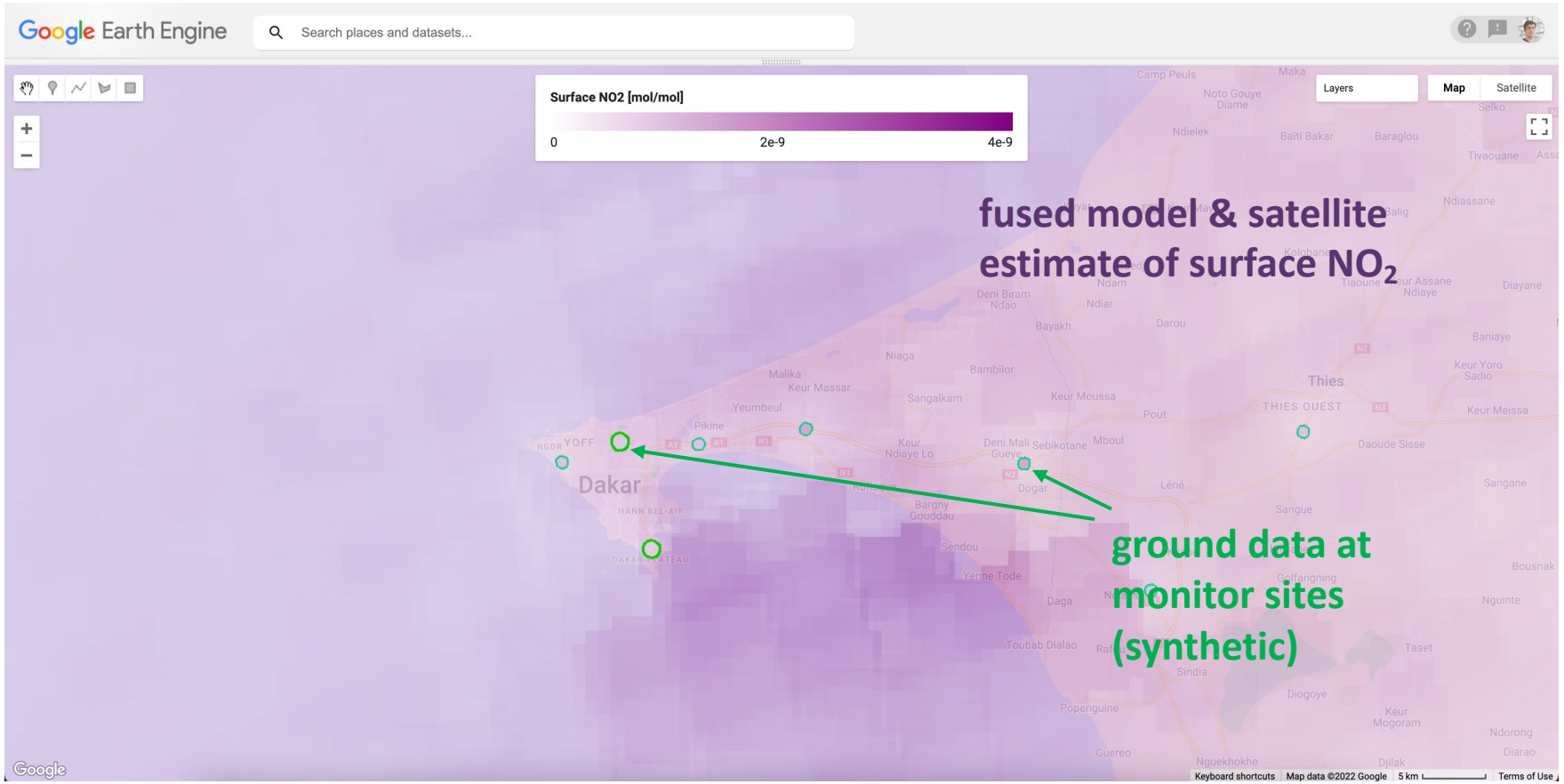
A draft framework application specification, outlining key Google cloud and GEE operational system components, has been developed. System component specifications and functions will be developed as the project progresses, accounting for end user needs and input.

Current ARL-Supporting Evidence

EARTH SCIENCE
APPLIED SCIENCES

ARL 3, Milestone 1: Components of the application tested and validated independently

Task 1e: Implement data fusion system into GEE tool



A prototype of the data fusion tool was implemented in Google Earth Engine. The prototype operated with actual (historical) GEOS-CF model and TROPOMI satellite data, while using synthetic surface monitor data. Comparison of synthetic surface data with the results of data fusion using different inputs showed the expected outcomes. Results presented at AGU 2022 and AMS 2023 conferences and in meetings with end-users.

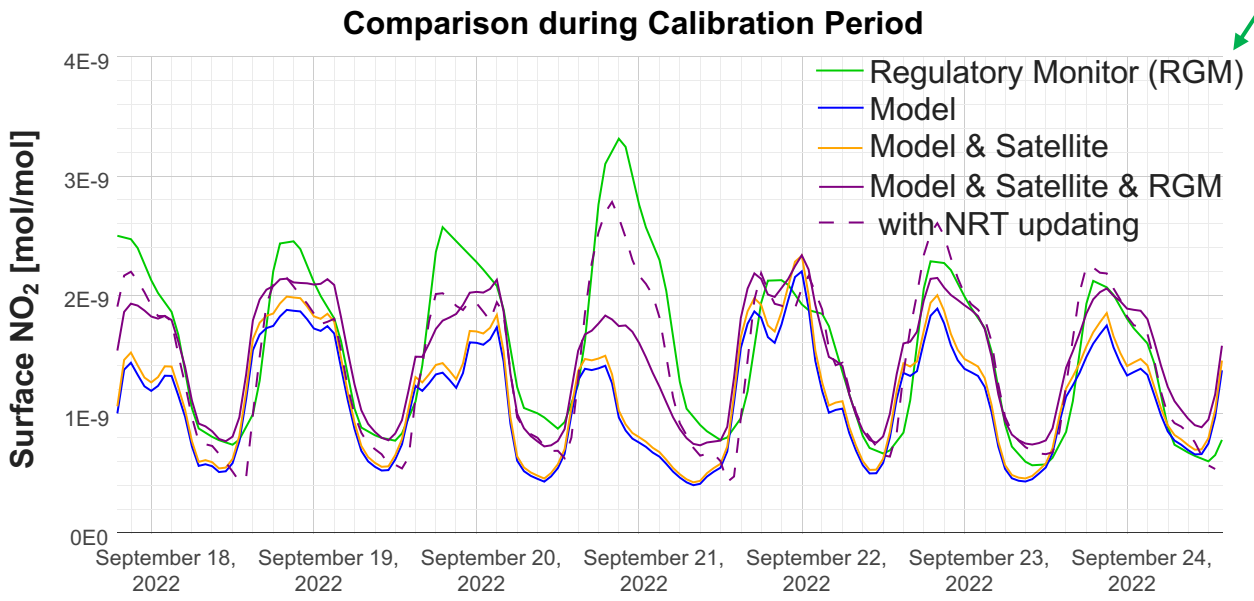
Supporting decision-making with a sub-city scale data fusion system

Current ARL-Supporting Evidence

EARTH SCIENCE
APPLIED SCIENCES

ARL 3, Milestone 1: Components of the application tested and validated independently

Task 1e: Implement data fusion system into GEE tool



Synthetic surface data (target)

Subsequent steps of the data fusion methodology

A prototype of the data fusion tool was implemented in Google Earth Engine. The prototype operated with actual (historical) GEOS-CF model and TROPOMI satellite data, while using synthetic surface monitor data. Comparison of synthetic surface data with the results of data fusion using different inputs showed the expected outcomes. Results presented at AGU 2022 and AMS 2023 conferences and in meetings with end-users.

Current ARL-Supporting Evidence

EARTH SCIENCE
APPLIED SCIENCES

ARL 3, Milestone 2: Detailed characterization of user decision making process completed

Task 2a: Assess the status and identify key end-user needs for the GEE tool

Organization Name	Key Decision-Making Activity	GEE Data Fusion Tool Capability Requests
City of Rio de Janiero, Brazil (Instituto Pereira Passos, Prefeitura do Rio de Janeiro; Contacts: Felipe Mandarino & Bruno Boscaro)	• Advance warning system for potential air-quality related hospital visits • Detecting "invisible" pollutant sources not included in current inventories	• High-resolution NO_2 mapping • AQ category forecasting • Inter-operability with existing GIS tools
City of Dakar, Senegal (Contacts: Aminata Mbow Diokhané [CGQA], Ngongang Danube [UNEP])	• Analysis and forecasting for transportation policy AQ impacts • Including Saharan dust forecasts for early warning	• High-resolution NO_2 and PM mapping • Integration of global models (especially for dust) to inform local forecasts
US EPA (Contact: John White)	• Incorporating low-cost sensors into their Air Quality forecasting and information systems (with associated uncertainties)	• Low-cost AQ sensor data integration • Uncertainty quantification

Challenges and Risks

EARTH SCIENCE
APPLIED SCIENCES

Rank	Type*	Risk	Mitigation Action	Date first noted/Date resolved (if applicable)
1	T	Lack of ground-based monitor data (especially in Dakar)	Use historical datasets (2010-2019) for validation; design the GEE tool to be robust to missing data	2/28/2023
2	T	NASA EOS satellites approaching end of life	Need to test alternative satellite products than MODIS	Start of project
3	T	Google Earth Engine service changes and availability	Work with Google team and collaborators to understand and adapt to service changes	Start of Project
4	ES	Changes in end-user priorities	Can be adjusted for so long as requirements remain within project scope	Start of project
5	T	Changes to collaborator systems and development timelines	Coordinate with collaborating entities to adjust plans within scope	2/28/2023
6	PM	Stakeholder personnel changes	Good communication with contacts to manage handover of project to new employee	1/15/2023
7	T	Air Quality Model updates may impact their availability in GEE	Coordinate with GEE to ensure minimum disruption in data availability	Start of project

* Please designate risk type as: Technical (T), Budget (B), End-User/Stakeholder (ES), or Project Management (PM)

Presentations Related to the Project

EARTH SCIENCE
APPLIED SCIENCES

- K. Emma Knowland, Carl Malings. “Air quality forecasting at sub-city-scale by combining models, satellites, and surface measures”. 6th AfriGEO Symposium. Virtual. October 31, 2022. Invited Presentation.
- Carl Malings. “Air quality forecasting at sub-city-scale by combining models, satellites, and surface measures”. Air Sensors International Conference Fall Webinar Series. Virtual. November 21, 2022. Invited Presentation.
- Carl Malings, K. Emma Knowland, Christoph Keller, Stephen Cohn, Nathan Pavlovic, Sean Khan, John White, Daniel Westervelt, Sean Wihera, Randall Martin, Alan Chan, Callum Wayman. “Towards a Flexible Data Fusion Tool for Air Quality Estimation and Forecasting with a Global Scope in Google Earth Engine”. Invited presentation, Google exhibition booth, American Geophysical Union Annual Meeting. Chicago, IL, USA. December 13, 2022.
- Carl Malings, K. Emma Knowland, Christoph Keller, Stephen Cohn, Nathan Pavlovic, Sean Khan, John White, Daniel Westervelt, Sean Wihera, Randall Martin, Alan Chan, Callum Wayman. “Towards a Flexible Data Fusion Tool for Air Quality Estimation and Forecasting with a Global Scope in Google Earth Engine”. American Geophysical Union Annual Meeting. Chicago, IL, USA. December 16, 2022.
- Carl Malings, K. Emma Knowland, Christoph Keller, Stephen Cohn, Nathan Pavlovic, Sean Khan, John White, Daniel Westervelt, Sean Wihera, Randall Martin, Alan Chan, Callum Wayman. “Towards a flexible data fusion tool incorporating model, satellite, regulatory monitor and low-cost sensor data for air quality estimation and forecasting”. American Meteorological Society 103rd Annual Meeting, Invited presentation. Denver, CO, USA. January 12, 2023.
- Carl Malings, K. Emma Knowland, Christoph Keller, Viral Shah, Callum Wayman, Stephen Cohn, Bryan Duncan, Nathan Pavlovic, Alan Chan, Sean Khan, John White, Daniel Westervelt, Sean Wihera, Randall Martin, Lazrak Noussair. “Creating & Using Open Global and Local Air Quality Data at NASA GMAO”. ESIP Meeting January 2023, Invited presentation. Online only. January 26, 2023.