



Partner



Air Quality, NASA Open Data, Google Earth Engine Data Fusion

Dr. Carl Malings

NASA Goddard Space Flight Center & Morgan State University

GESTAR II Cooperative Agreement

Agenda

- **Motivation**

- What is the global air quality challenge?
- Why are open data needed to meet this challenge?

- **NASA Open Data**

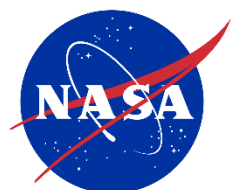
- Satellites – key missions and datasets for air quality
- Models – GEOS-CF air quality forecasts
- Finding NASA Data – Earthdata Search

- **Data Fusion in Google Earth Engine**

- Our NASA-funded project to support air quality managers
- Walkthrough of the data fusion method
- Capabilities and challenges in GEE

- **Further Resources to Learn More**





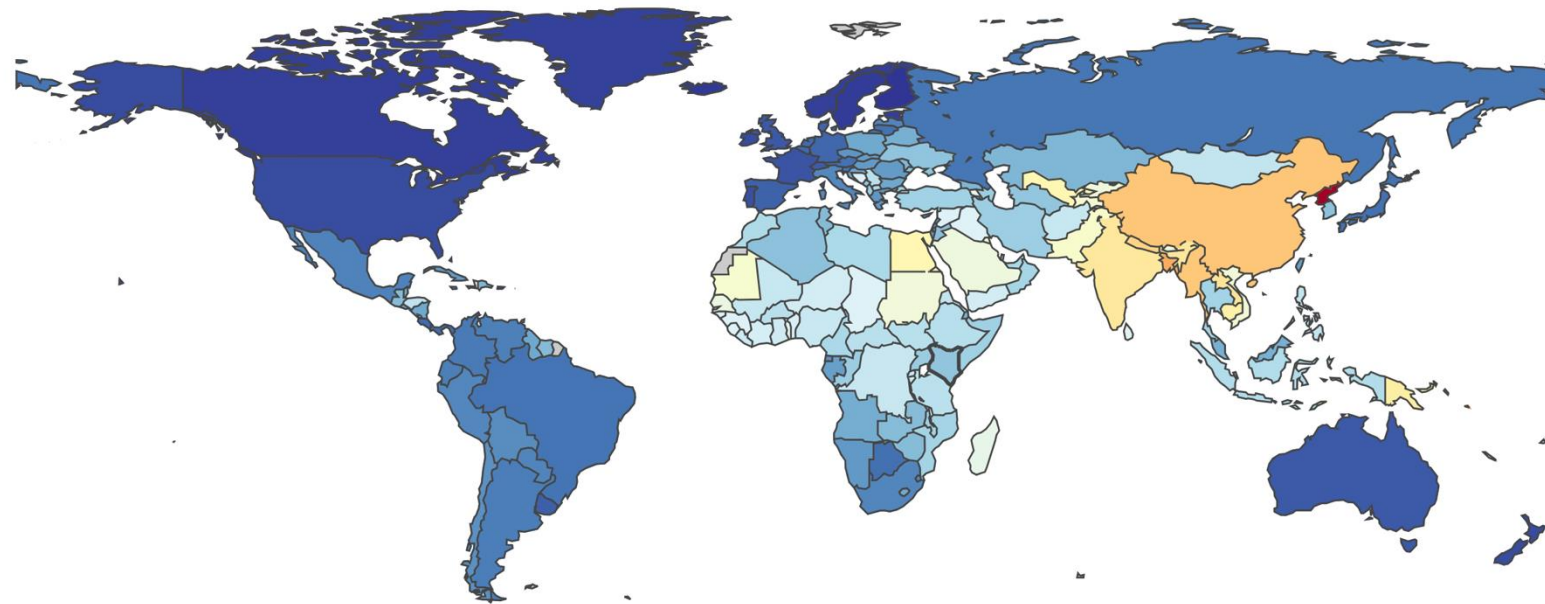
Partner



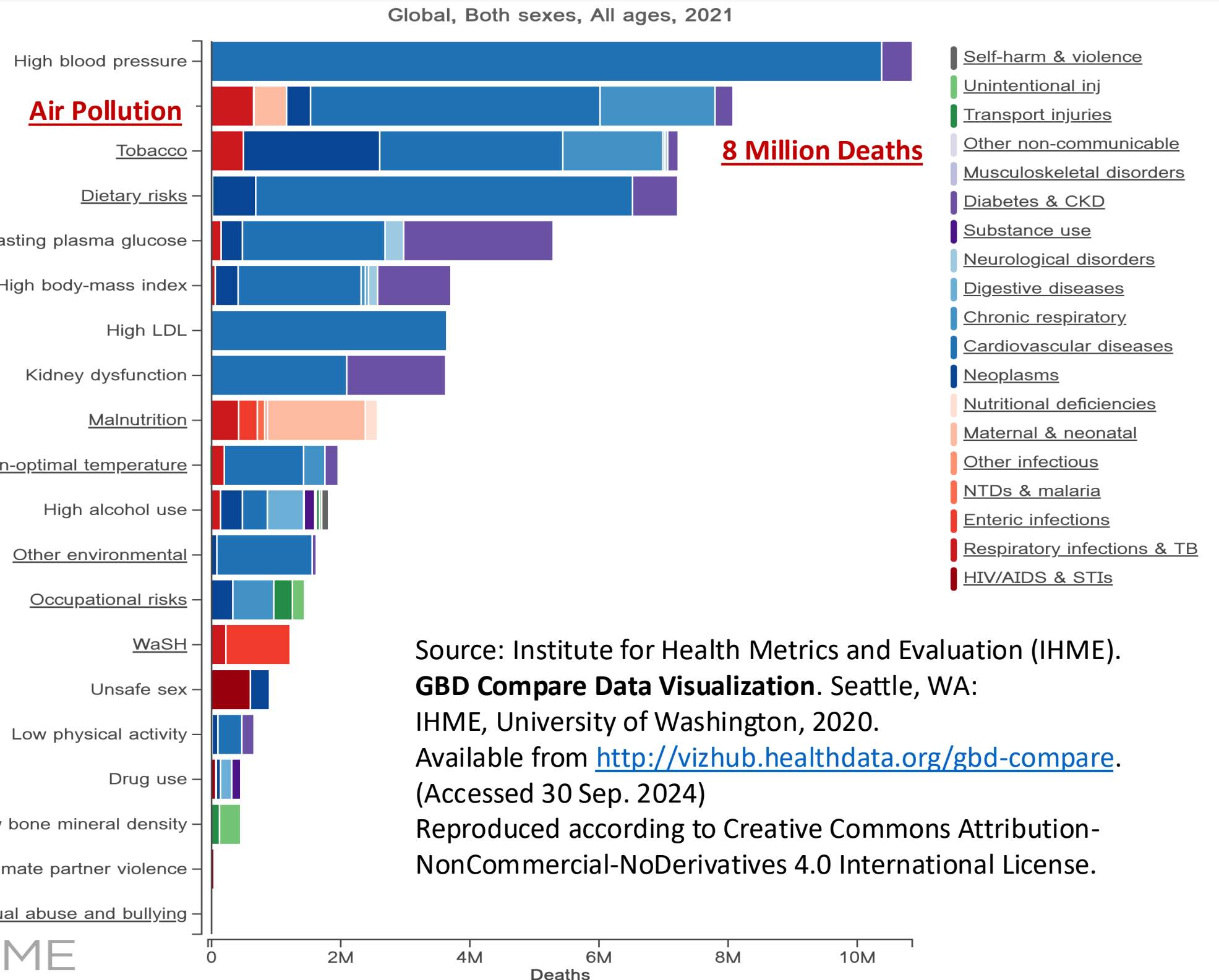
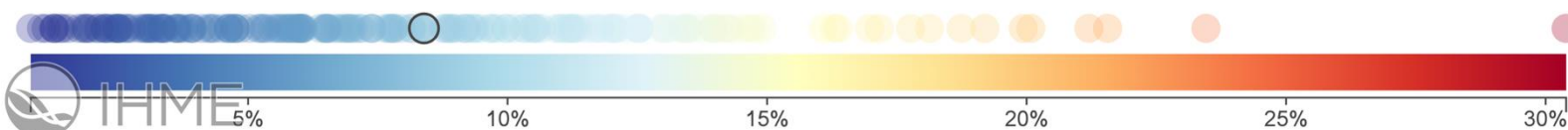
Global Air Quality

The global risks of poor air quality are severe...

Air pollution
Both sexes, All ages, 2021, Percent of total deaths



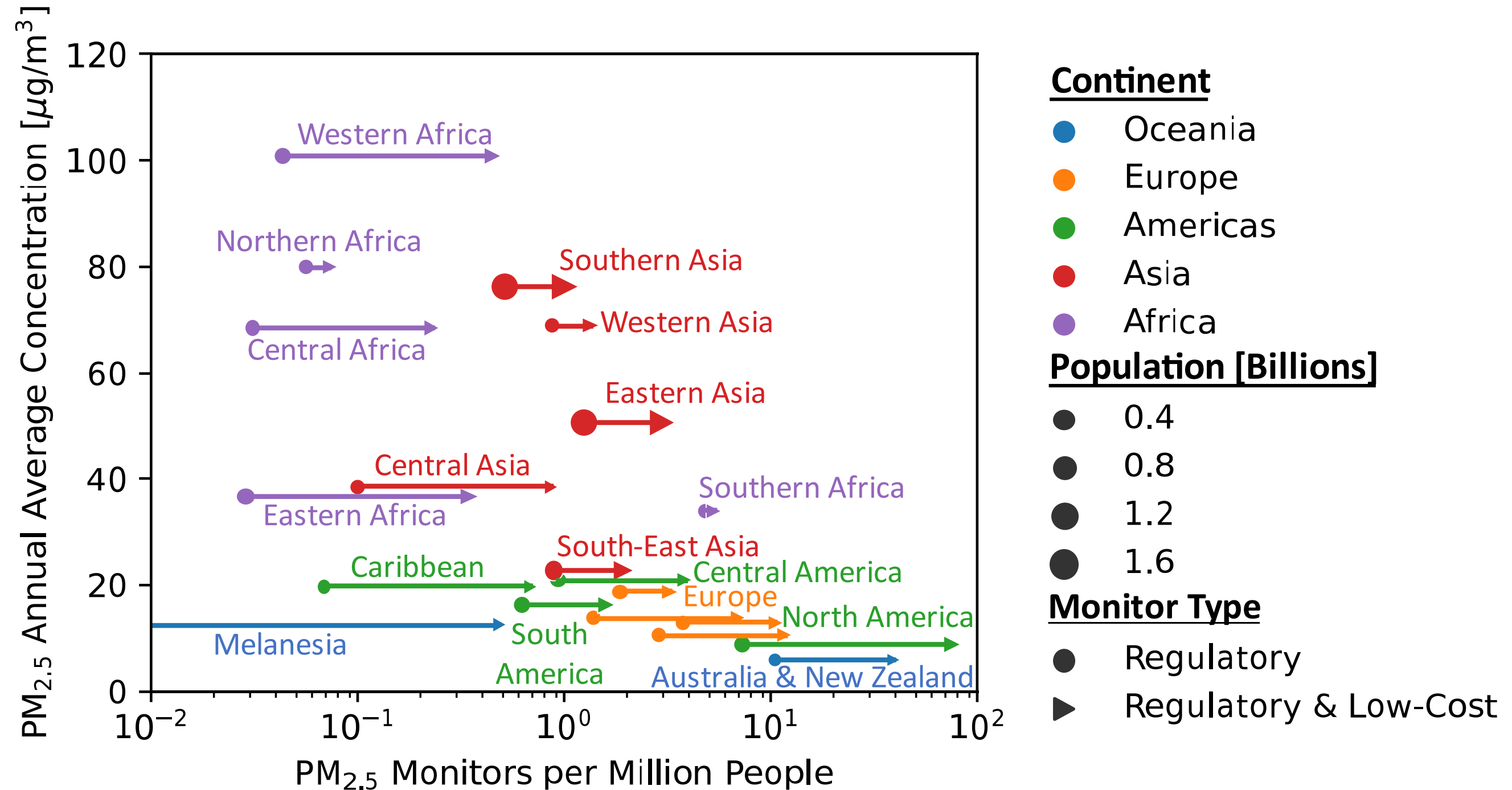
Kenya: 8.4% of deaths (30 thousand per year) due to air pollution



Source: Institute for Health Metrics and Evaluation (IHME).
GBD Compare Data Visualization. Seattle, WA:
IHME, University of Washington, 2020.
Available from <http://vizhub.healthdata.org/gbd-compare>.
(Accessed 30 Sep. 2024)
Reproduced according to Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

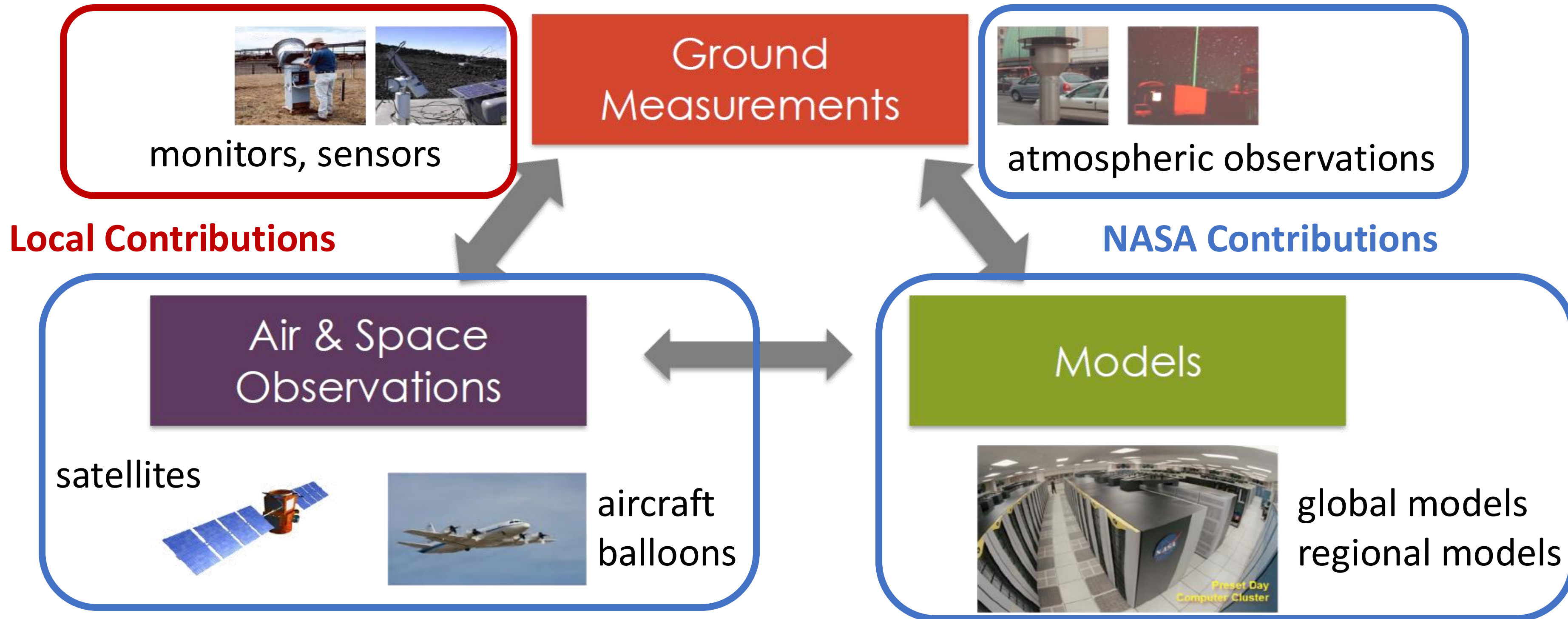
...but our knowledge of air quality is incomplete!

Analysis using
open air quality
data from
openAQ.org

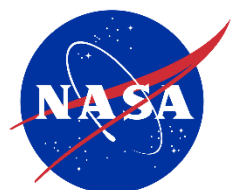


Source: Malings et al., 2024. "[Forecasting with the GEOS-CF System and Other NASA Resources to Support Air Quality Management](#)", in: Advances in Air Quality Research in Africa, Advances in Science, Technology & Innovation.

How can we measure and understand air quality?



Source: Gupta, P.; Follette-Cook, M. (2018). Satellite Remote Sensing of Air Quality. NASA Applied Remote Sensing Training Program (ARSET).
<https://appliedsciences.nasa.gov/join-mission/training/english/arset-satellite-remote-sensing-air-quality>

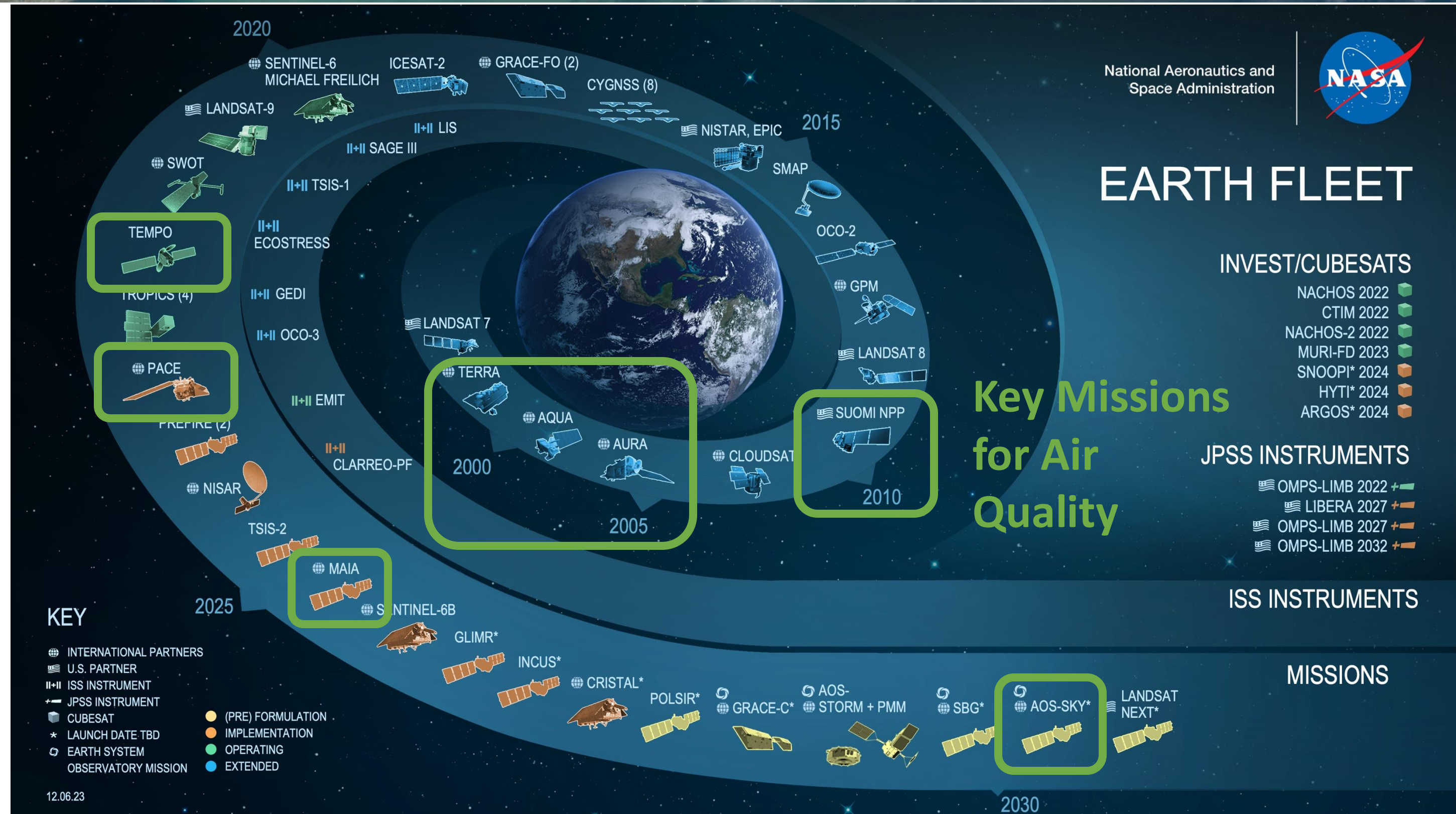


Partner



NASA Open Data

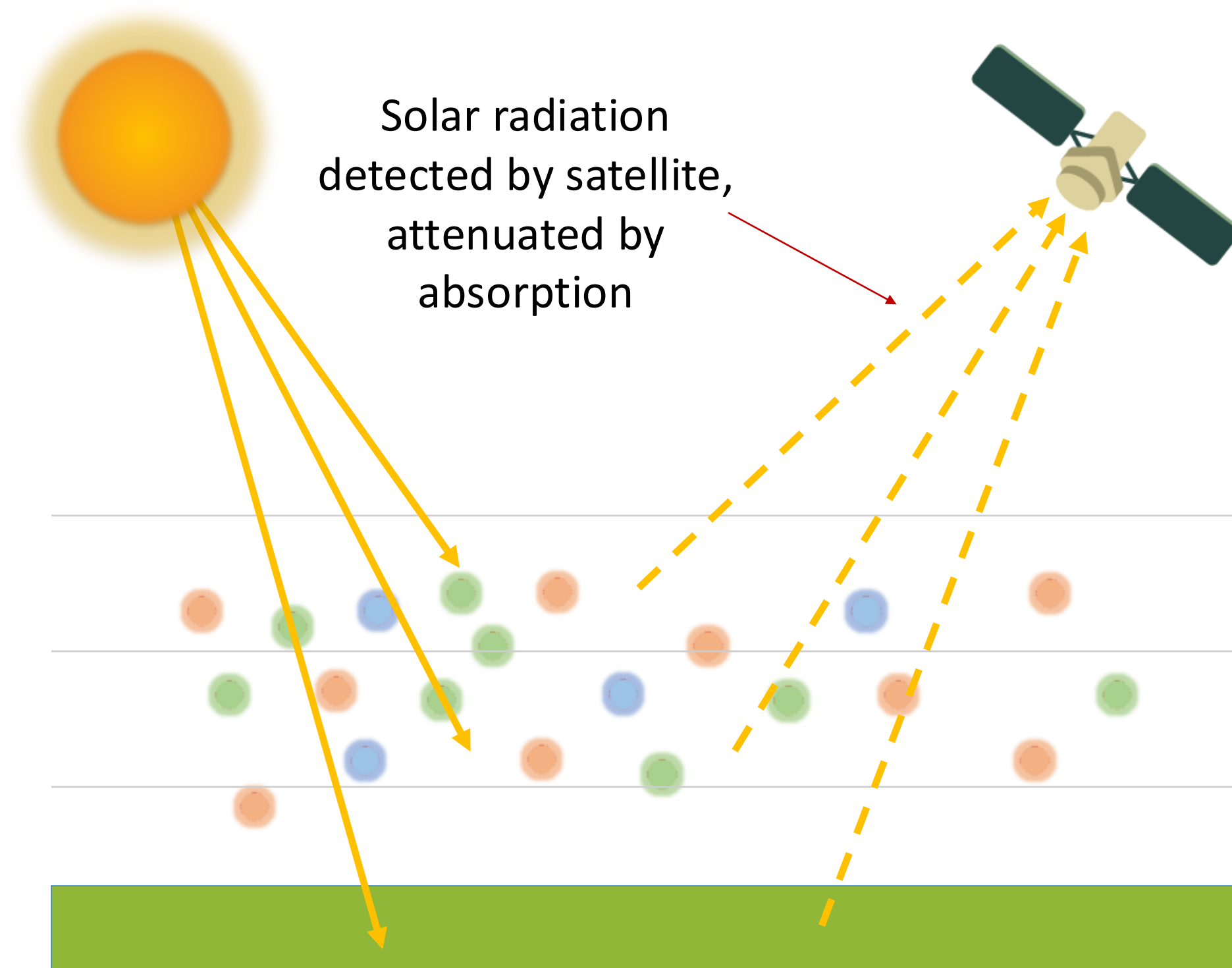
NASA's Earth Observation Fleet



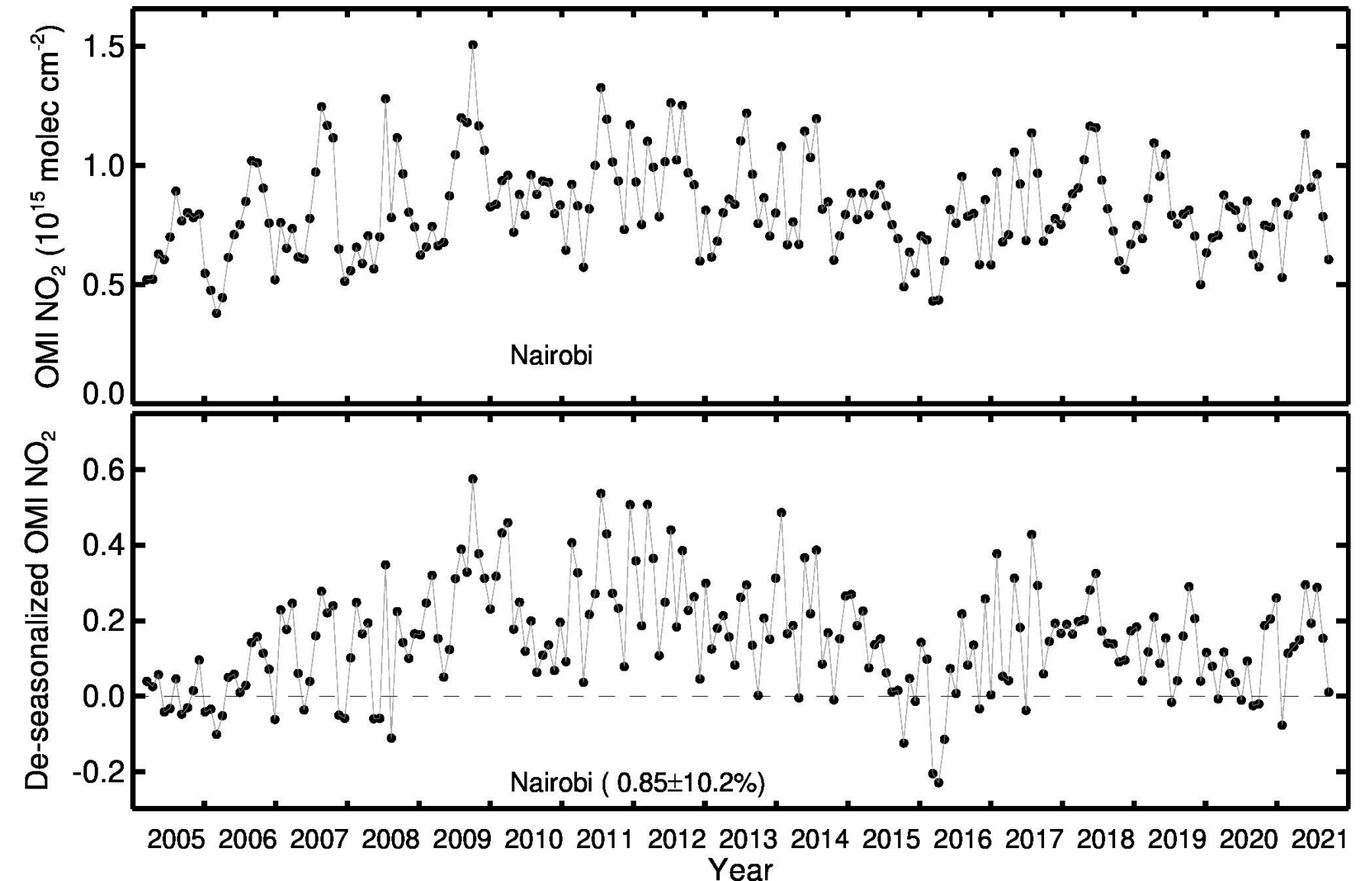
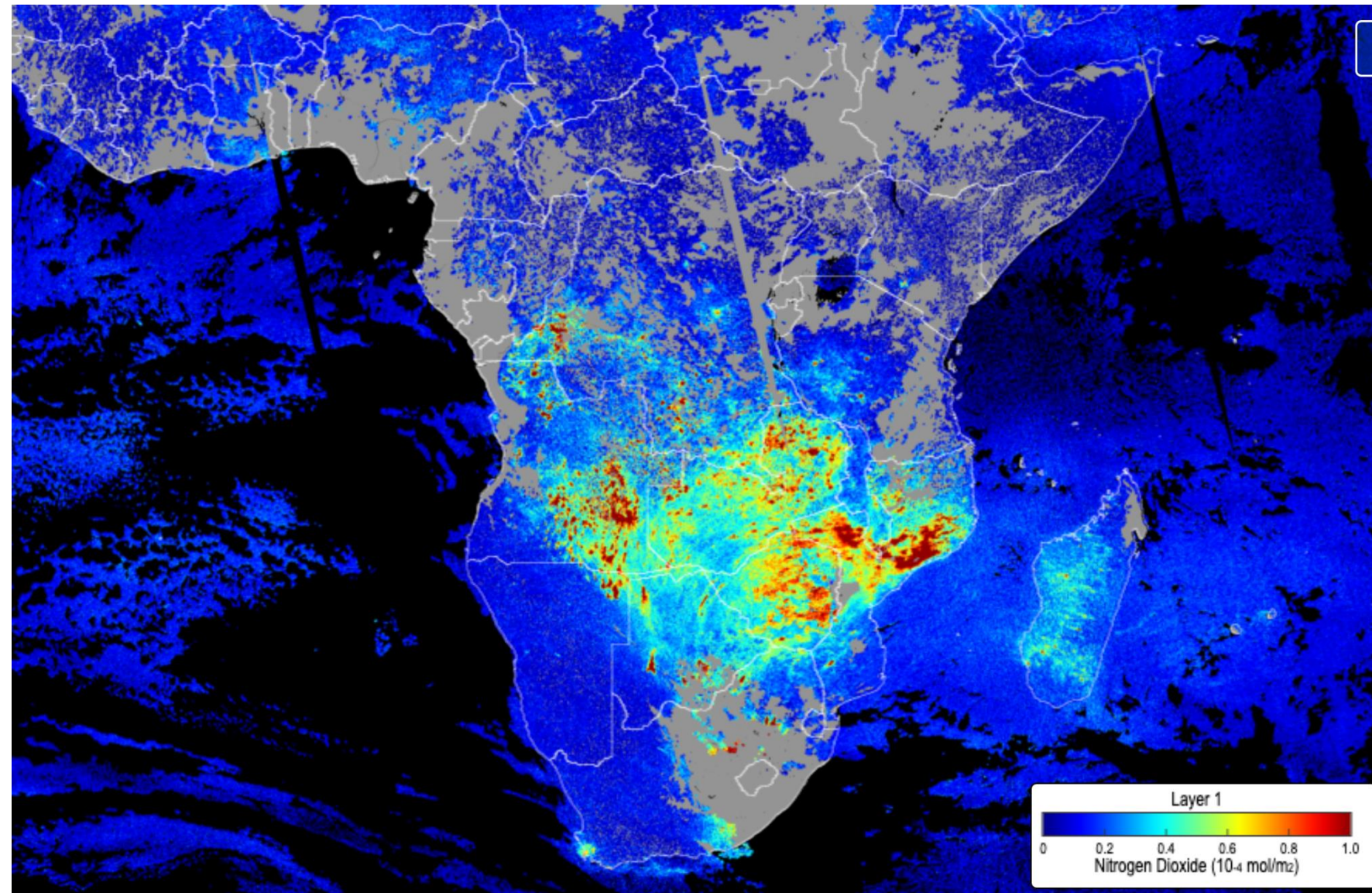
Source: [NASA Scientific Visualization Studio](#)

Air Quality Observation from Space

- In passive remote sensing (most common for air quality applications) satellites detect **solar radiation** backscattered from the Earth & atmosphere.
- Retrieval algorithms use physical principles to infer **geophysical quantities** from this data, such as:
 - **aerosol optical depth** for all aerosols in the atmosphere, using visible light.
 - **trace gas total column density** using the “spectral fingerprints” of many gases.
- Usually, there is little information on the **vertical distribution** of aerosols or trace gases, although limb view or active sensors can do this.

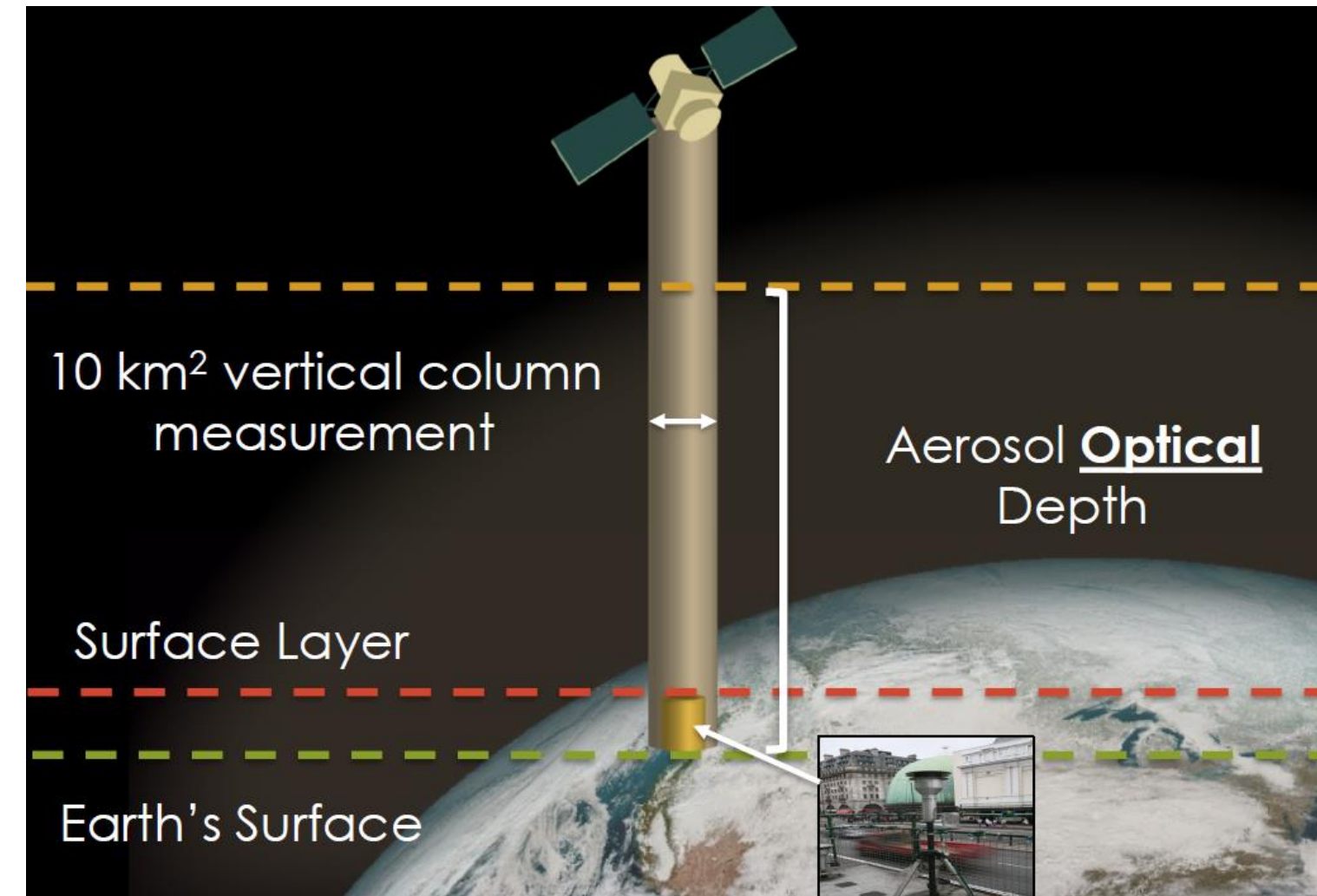
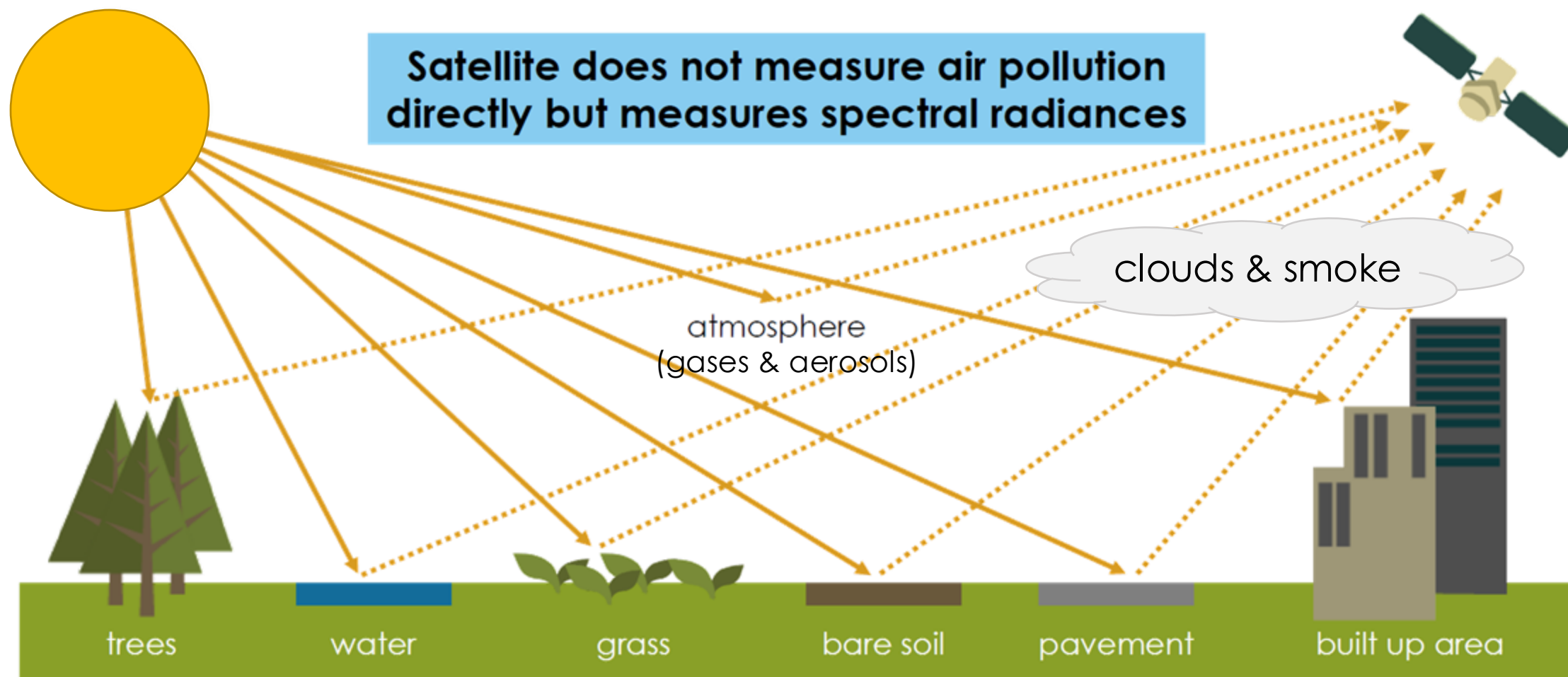


What satellite data CAN do for air quality



- **Examine a large area:** where are the hotspots? how is long-range transport happening?
- **Track changes over time:** how much has the average concentration over an area changed over time?
- **A picture is worth a million datapoints:** Anyone can understand a satellite photo of a smoke plume.

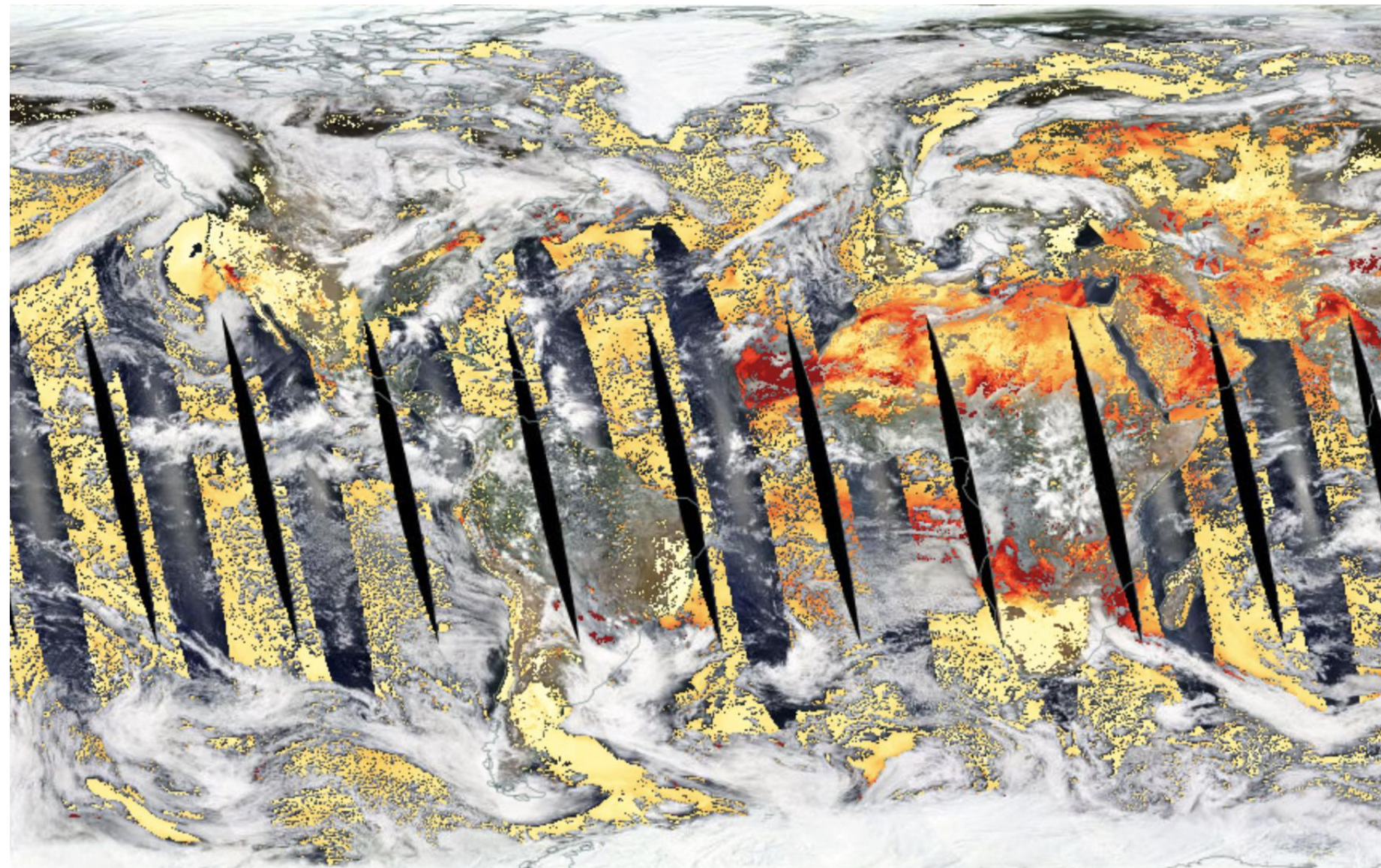
What satellite data CANNOT do for air quality



- **See at night:** satellites measure the properties of reflected sunlight passing through the atmosphere.
- **See through clouds:** most satellite measurements are blocked by cloud cover.
- **See what is happening at “nose level”:** satellites measure quantities in the whole atmosphere.
- **See at different times of day:** polar-orbiting satellites will observe a location once per day.

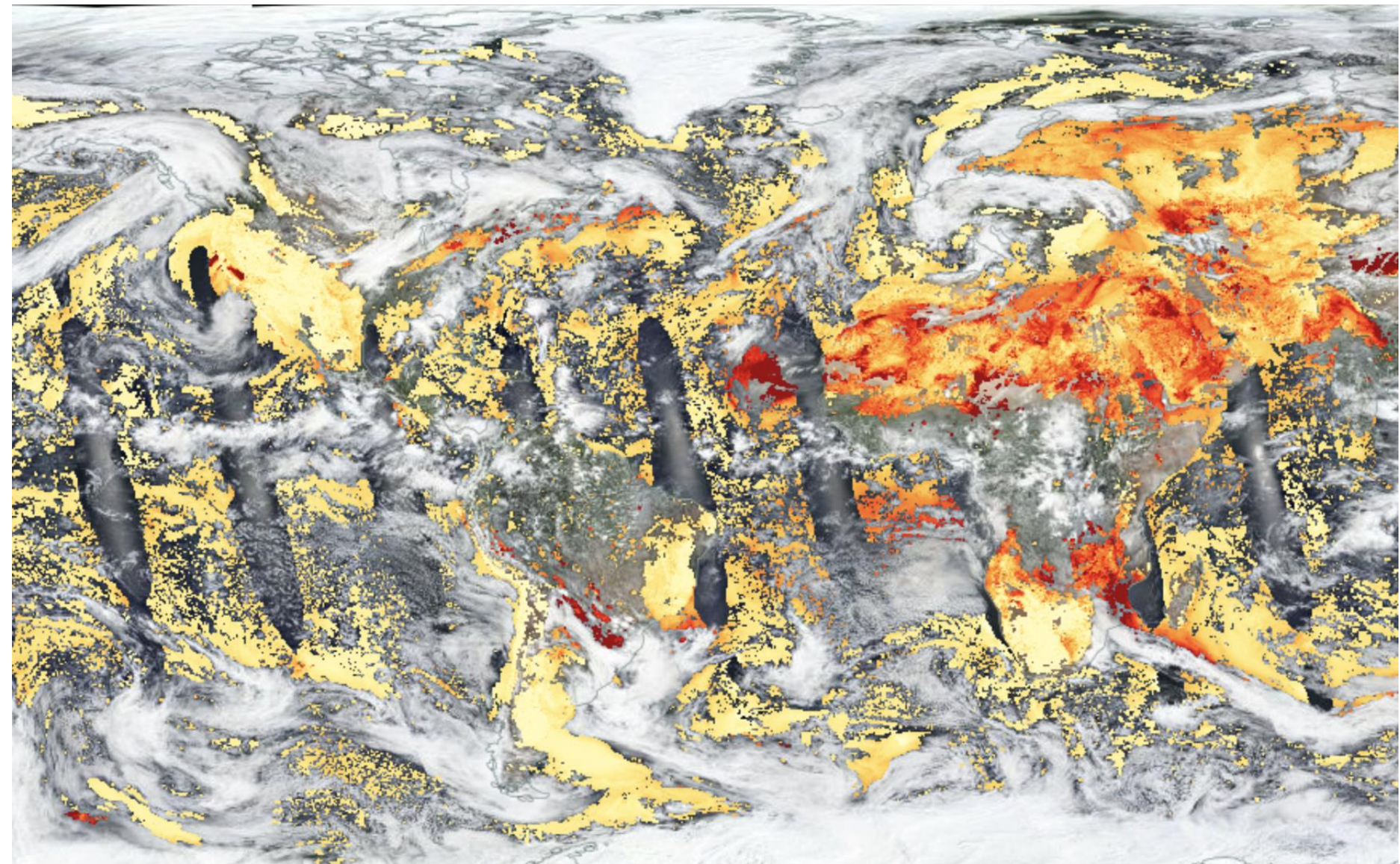
Aerosols – MODIS and VIIRS AOD

MODIS Aqua Deep Blue AOD (Launched 1999)



2 Satellites (Aqua, Terra)
1 km high-resolution AOD ([MAIAC](#))

VIIRS SNPP Deep Blue AOD (Launched 2011)



3 Satellites (SNPP, NOAA-20, NOAA-21)
0.75 km high-resolution AOD ([in development](#))

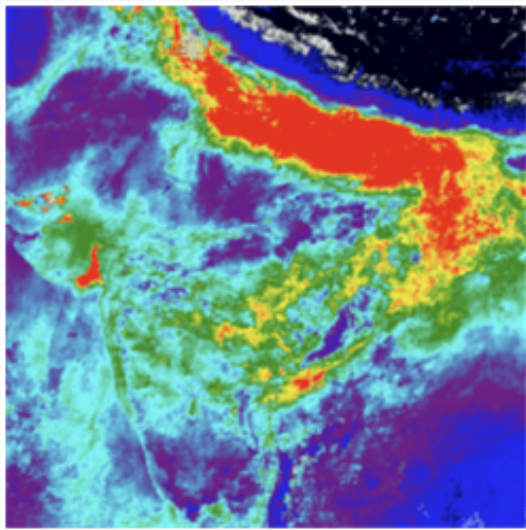
MODIS MAIAC AOD in Google Earth Engine

Earth Engine Data Catalog

Search / English

Home View all datasets Browse by tags Landsat MODIS Sentinel Publisher Community API Docs

MCD19A2.061: Terra & Aqua MAIAC Land Aerosol Optical Depth Daily 1km



Dataset Availability
2000-02-24T00:00:00Z–2024-09-25T23:55:00Z

Dataset Provider
[NASA LP DAAC at the USGS EROS Center](#)

Earth Engine Snippet

```
ee.ImageCollection("MODIS/061/MCD19A2_GRANULES")
```

Tags
aerosol aod aqua daily global maiac modis nasa terra usgs
mcd19a2

Source: <https://developers.google.com/earth-engine/datasets/catalog/>

Description

Bands

Terms of Use

Citations

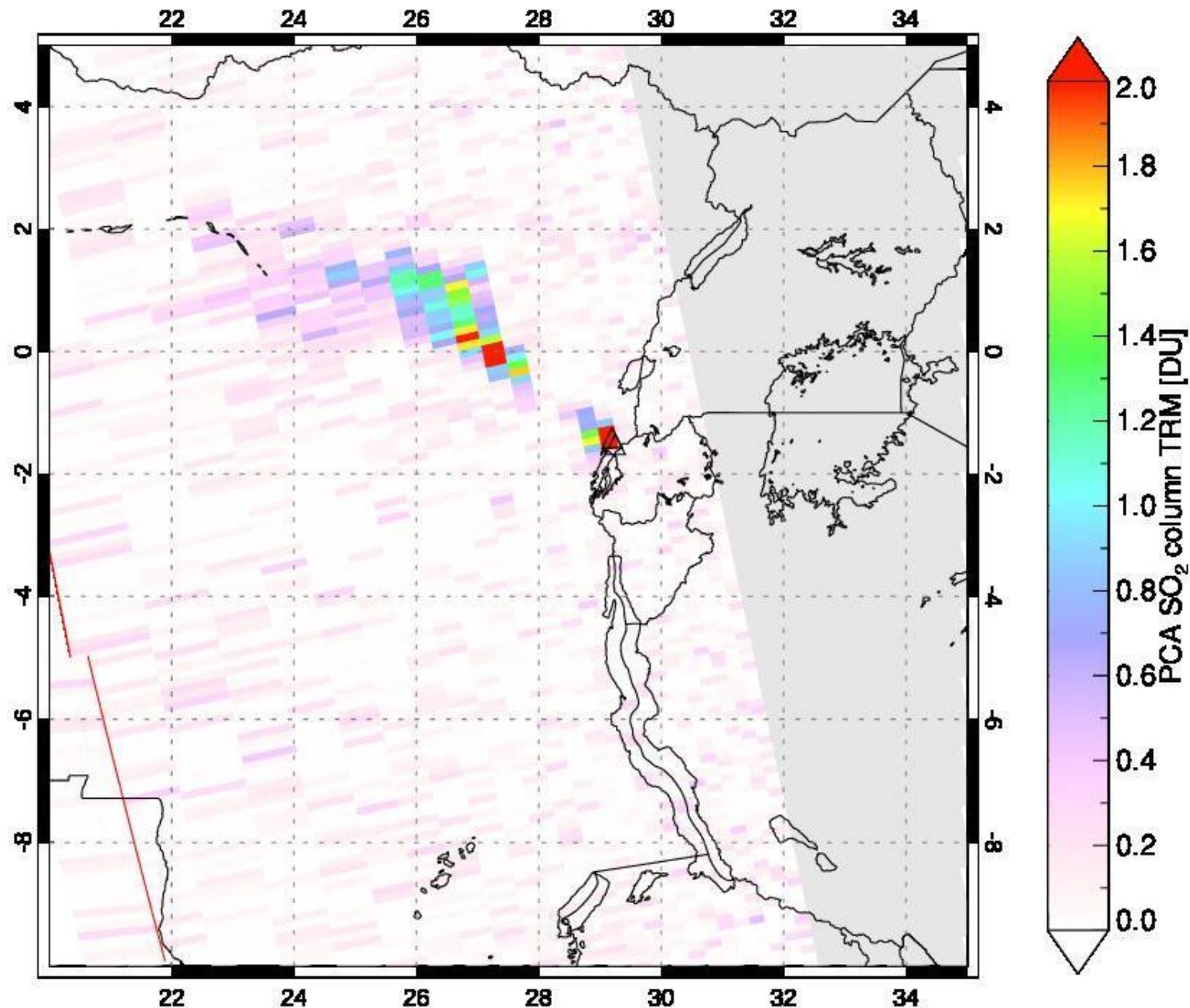
DOIs

The MCD19A2 V6.1 data product is a MODIS Terra and Aqua combined Multi-angle Implementation of Atmospheric Correction (MAIAC) Land Aerosol Optical Depth (AOD) gridded Level 2 product produced daily at 1 km resolution. For more information see the [MAIAC user guide](#).

NOTE: This product has been released with the caveat that the reprocessing for the full mission is expected to continue through summer 2023.

Gases – NASA's OMI and ESA's TROPOMI

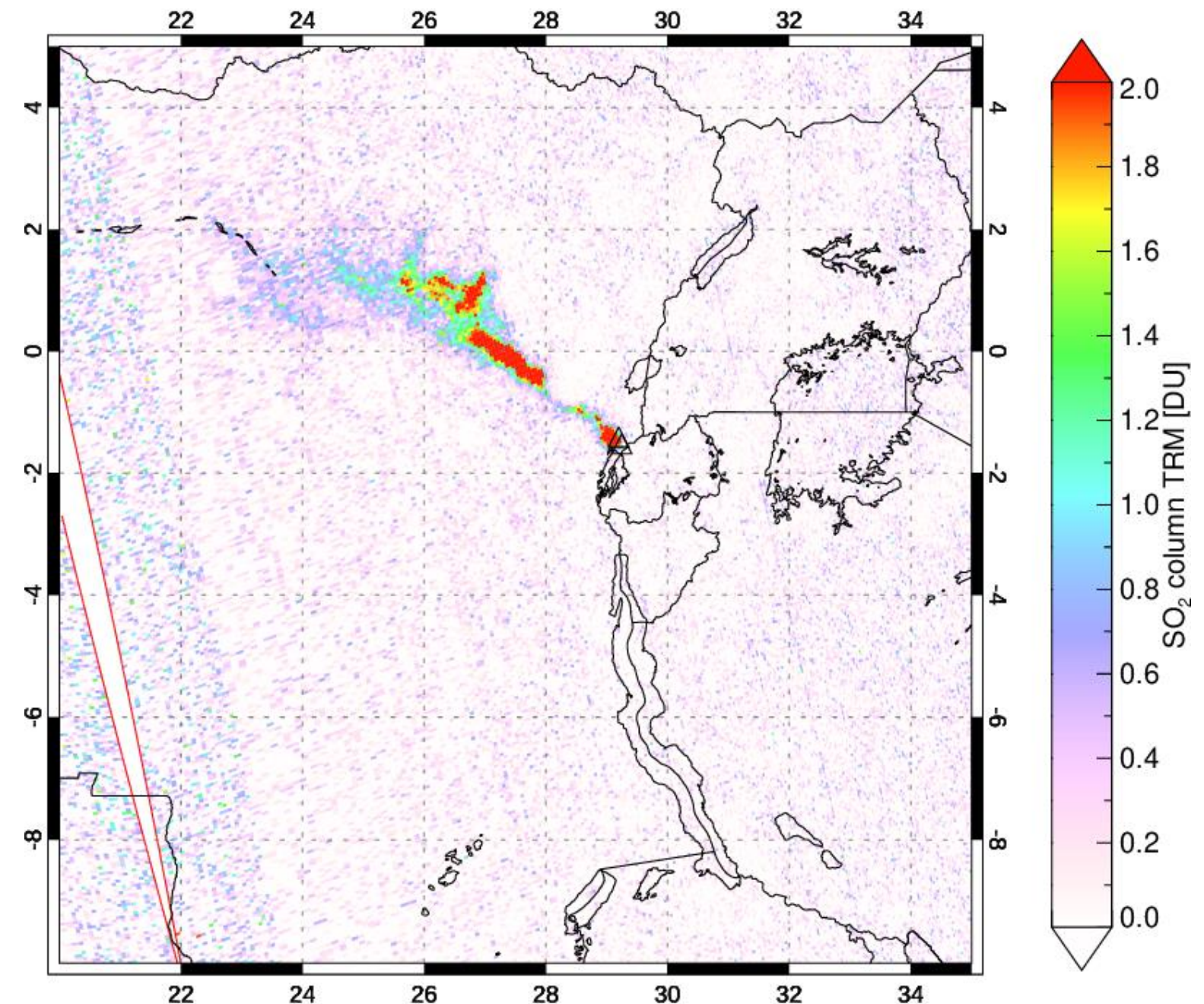
OMI SO₂, Nyiragongo, Sep. 10, 2024



Data: NASA Aura Project

NO₂, SO₂, HCHO, Ozone - 13 x 24 km

TROPOMI SO₂, Nyiragongo, Sep. 10, 2024



Data: BIRA-IASB/DLR/ESA/EU Copernicus Program

NO₂, SO₂, HCHO, Ozone, CO, CH₄ - 3.5 x 7 km

Source: NASA GSFC Global Sulfur Dioxide Monitoring Home Page <https://so2.gsfc.nasa.gov/>

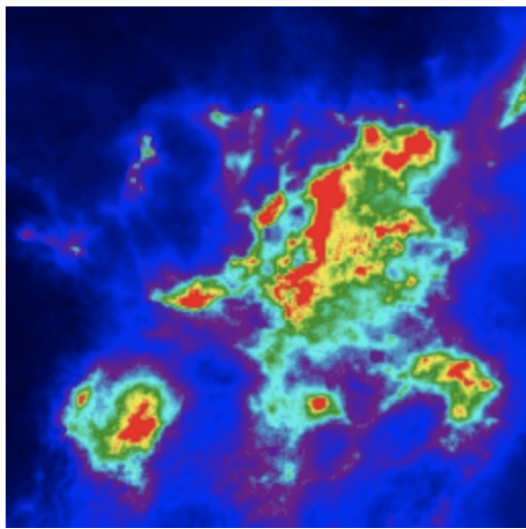
ESA's TROPOMI in Google Earth Engine

Earth Engine Data Catalog

Search / English

HomeView all datasetsBrowse by tagsLandsatMODISSentinelPublisherCommunityAPI Docs

Sentinel-5P OFFL NO2: Offline Nitrogen Dioxide



Dataset Availability
2018-06-28T10:24:07Z–2024-09-20T22:55:18Z

Dataset Provider
[European Union/ESA/Copernicus](#)

Earth Engine Snippet

```
ee.ImageCollection("COPERNICUS/S5P/OFFL/L3_NO2")
```

Tags
air-quality copernicus esa eu knmi nitrogen-dioxide no2 pollution s5p sentinel tropomi

Source: <https://developers.google.com/earth-engine/datasets/catalog/>

Description Bands Image Properties Terms of Use

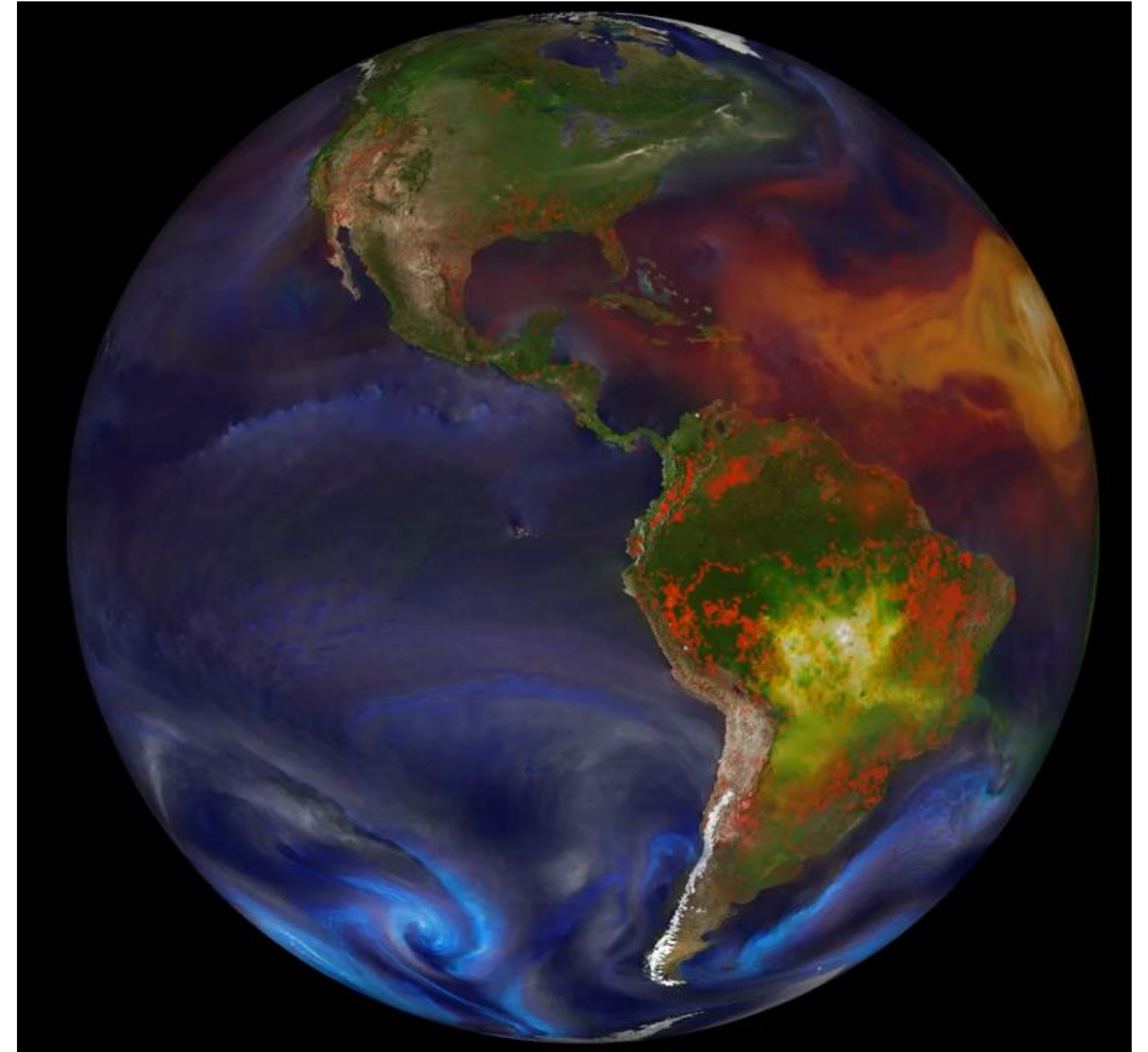
OFFL/L3_NO2

This dataset provides offline high-resolution imagery of NO₂ concentrations.

Nitrogen oxides (NO₂ and NO) are important trace gases in the Earth's atmosphere, present in both the troposphere and the stratosphere. They enter the atmosphere as a result of anthropogenic

NASA Global Modeling and Assimilation Office

GMAO's core mission is to enhance the value of NASA's observations to understand, analyze and predict changes in the physics, chemistry and biology of the Earth system



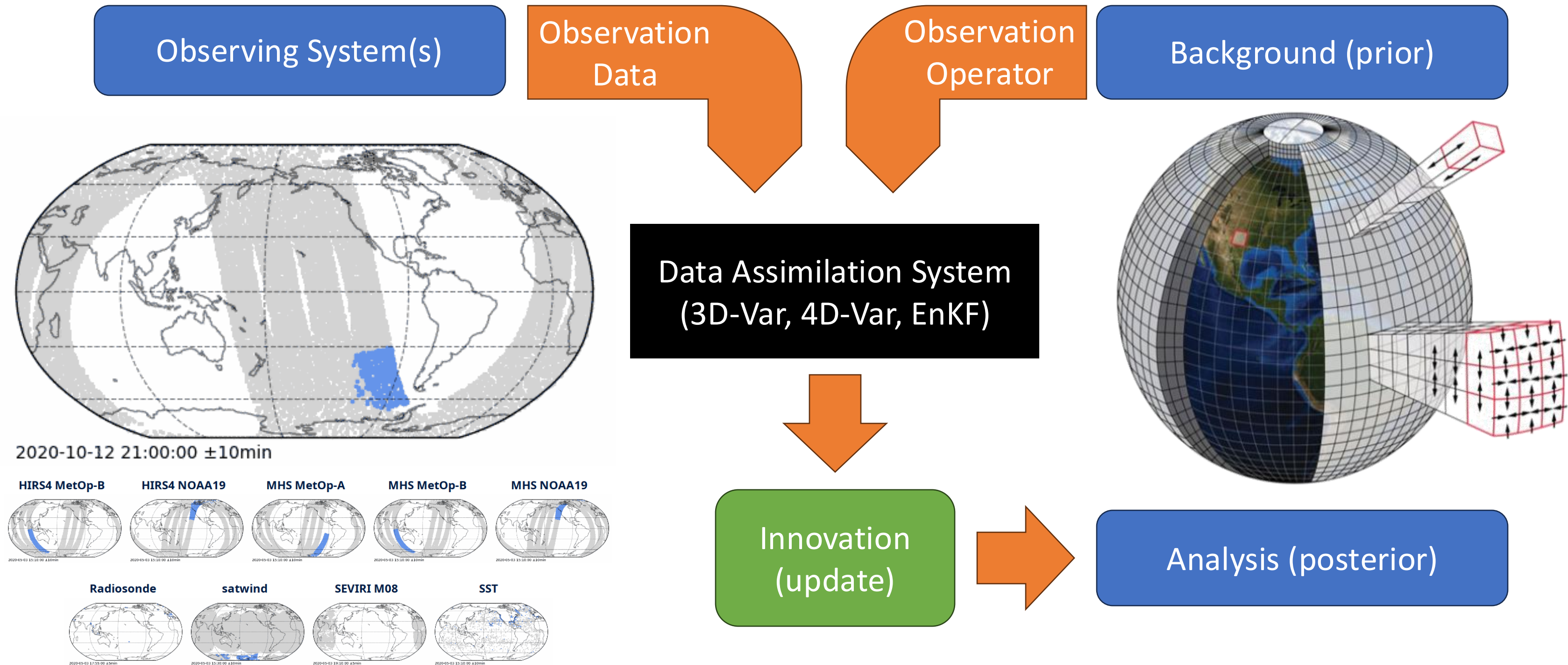
Source: NASA Scientific Visualization Studio <https://svs.gsfc.nasa.gov/30017>

Presentation for GEEDEVs Nairobi Community
carl.a.malings@nasa.gov

17 Oct. 2024

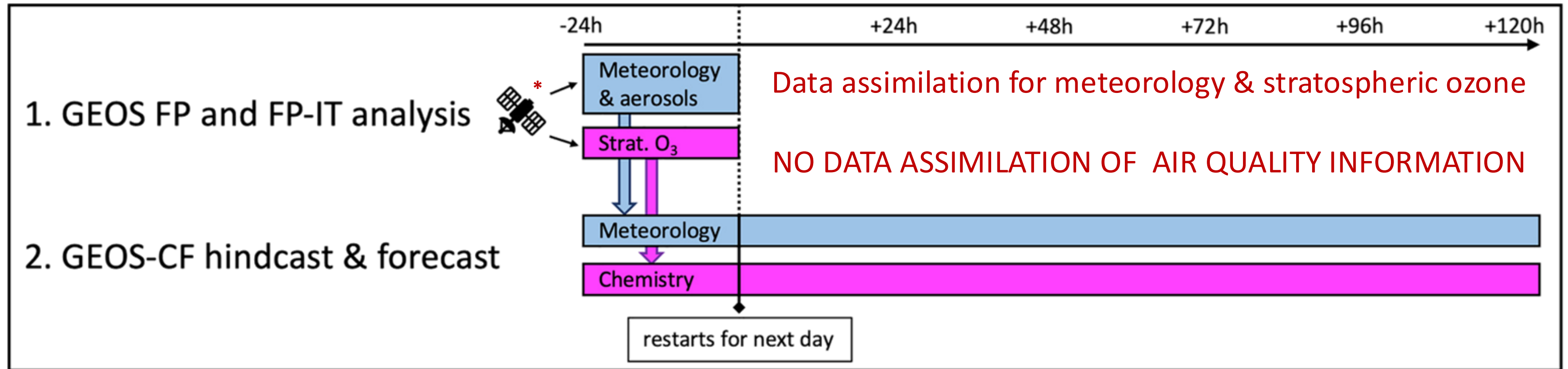
16

Data Assimilation

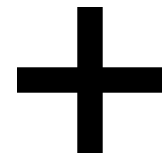
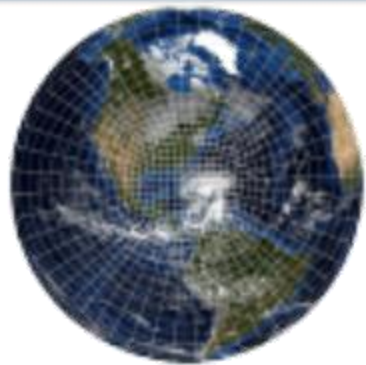


Source: Joint Center for Satellite Data Assimilation, <https://www.jcsda.org/>

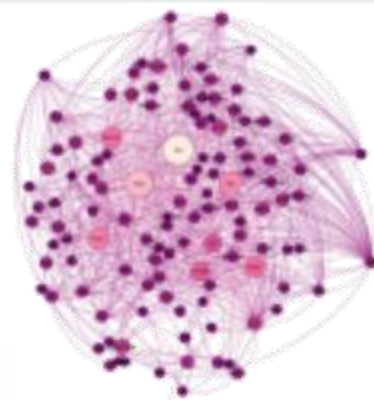
GEOS-CF Composition Forecast



GEOS
Meteorology



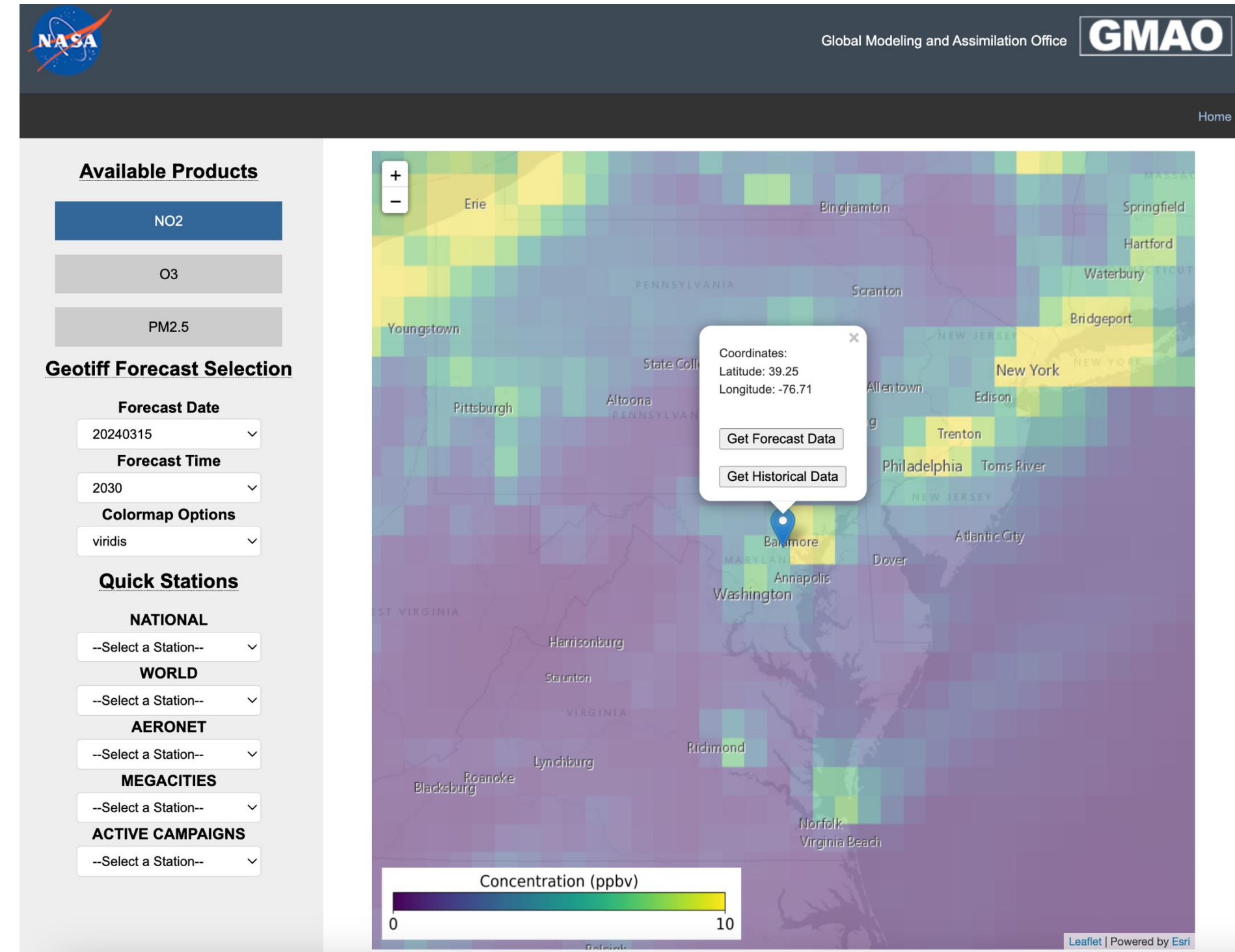
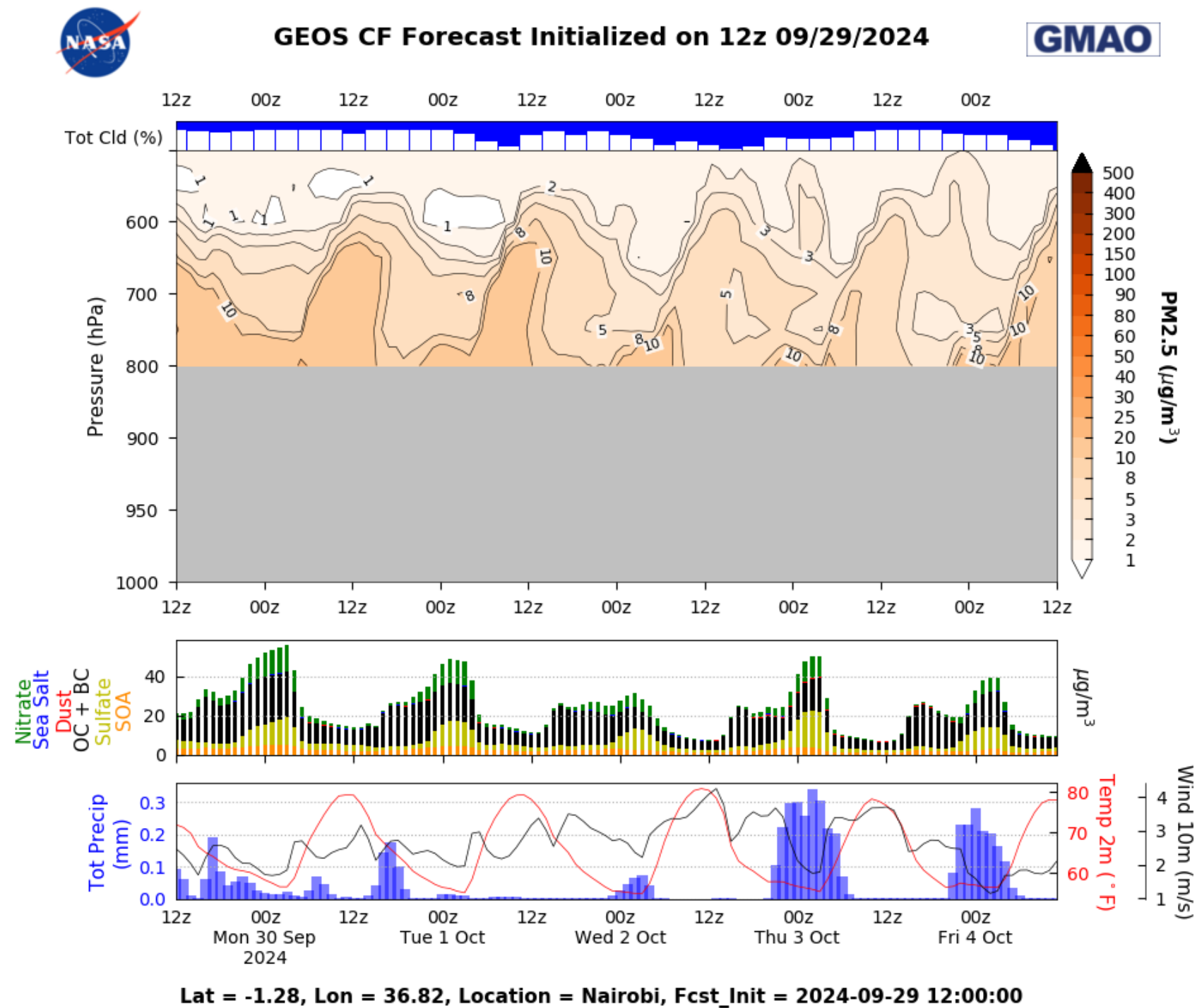
GEOS-Chem
Chemistry



- Combine GEOS meteorology with GEOS-Chem chemistry
- 250 chemical species
- Hourly temporal resolution
- 0.25 degree (25 km) spatial resolution
- Global coverage
- Daily 1-day hindcast and 5-day forecast

Source: Keller, C., et al. (2021) "Description of the NASA GEOS Composition Forecast Modeling System GEOS-CF v1.0". *Journal of Advances in Modeling Earth Systems*, 13:4.
<https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020MS002413>

NASA FLUID webpage



GEOS-CF Forecasts in Google Earth Engine

Earth Engine Data Catalog

Q GEOS-CF

English

Home

View all datasets

Browse by tags

Landsat

MODIS

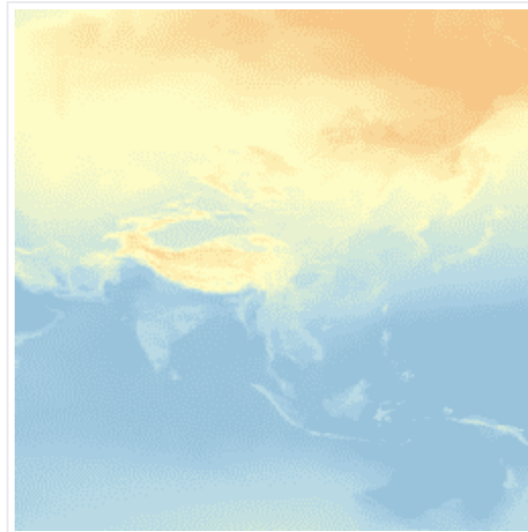
Sentinel

Publisher

Community

API Docs

GEOS-CF fcst tavg1hr v1: Goddard Earth Observing System Composition Forecast



Dataset Availability

2022-10-01T00:00:00Z–2024-09-27T12:00:00Z

Dataset Provider

[NASA / GMAO](#)

Earth Engine Snippet

```
ee.ImageCollection("NASA/GEOS-CF/v1/fcst/tavg1hr")
```



Tags

composition

forecast

geos

gmao

nasa

Source: <https://developers.google.com/earth-engine/datasets/catalog/>

Description

Bands

Image Properties

Terms of Use

Citations

DOIs

This dataset contains meteorological forecast (fcst) of time-averaged frequency data (tavg1hr). Use the 'creation_time' and 'forecast_time' properties to select data of interest. The Goddard Earth Observing System Composition Forecast (GEOS-CF) system is a high-resolution (0.25°) global constituent prediction system from NASA's [Global Modeling and Assimilation Office\(GMAO\)](#).

GEOS-CF offers a new tool for atmospheric chemistry research, with the goal to supplement NASA's broad range of space-based and in-situ observations. GEOS-CF expands on the GEOS weather and aerosol modeling system by introducing the [GEOS-Chem](#) chemistry module to provide hindcasts and 5-days forecasts of atmospheric constituents including ozone (O3), carbon monoxide (CO), nitrogen dioxide (NO2), sulfur dioxide (SO2), and fine particulate matter (PM2.5). The chemistry module integrated in GEOS-CF is identical to the offline GEOS-Chem model and readily benefits from the innovations provided by the GEOS-Chem community.



Presentation for GEEDEVs Nairobi Community
carl.a.malings@nasa.gov

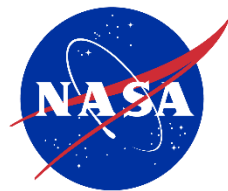
17 Oct. 2024

20

NASA Earthdata Search

The screenshot displays the NASA Earthdata Search web application. The top navigation bar includes the NASA logo, 'EARTHDATA', a 'Find a DAAC' dropdown, and links for 'Feedback' and a user profile 'Carl'. The main header reads 'EARTHDATA SEARCH'. Below this is a search bar with the placeholder text 'Search for collections or topics'. To the left of the search results is a sidebar with a 'Filter Collections' section. This section includes checkboxes for 'Available in Earthdata Cloud', 'Customizable', and 'Map Imagery'. It also has expandable sections for 'Keywords', 'Platforms', and 'Instruments' (currently showing '1 Selected'). Under 'Instruments', a list of instruments is shown with their respective granule counts: AIRS (138), AMSR2 (74), AMSR-E (100), Anemometer (83), Aquarius_Radiometer (138), Aquarius_Scatterometer (161), ASTER (56), ATMS (66), AVHRR (132), and AVHRR-3 (126). The main content area shows '69 Matching Collections'. The first two results are for Sentinel-5P TROPOMI Tropospheric NO2 data (V2 and V1) at GES DISC, each with 26,882 and 9,849 granules respectively, and a description of the data's spatial resolution and start date. The third result is for Sentinel-5P TROPOMI Sulphur Dioxide SO2 data (V2) at GES DISC, with 15,385 granules. Each result includes a small thumbnail map and a link to the 'Earthdata Cloud'. On the right side of the interface is a large map of the Americas, showing North and Central America. The map includes a scale bar (1000 km / 500 mi) and a vertical toolbar with various map controls like zoom, pan, and layers. The footer of the page contains version information 'v23.2.5-2', search time '3.0s', and links to NASA Official, FOIA, Privacy Policy, and USA.gov. A link for 'Earthdata Access: A Section 508 accessible alternative' is also present.

<https://search.earthdata.nasa.gov/search>



Partner



Data Fusion in Google Earth Engine

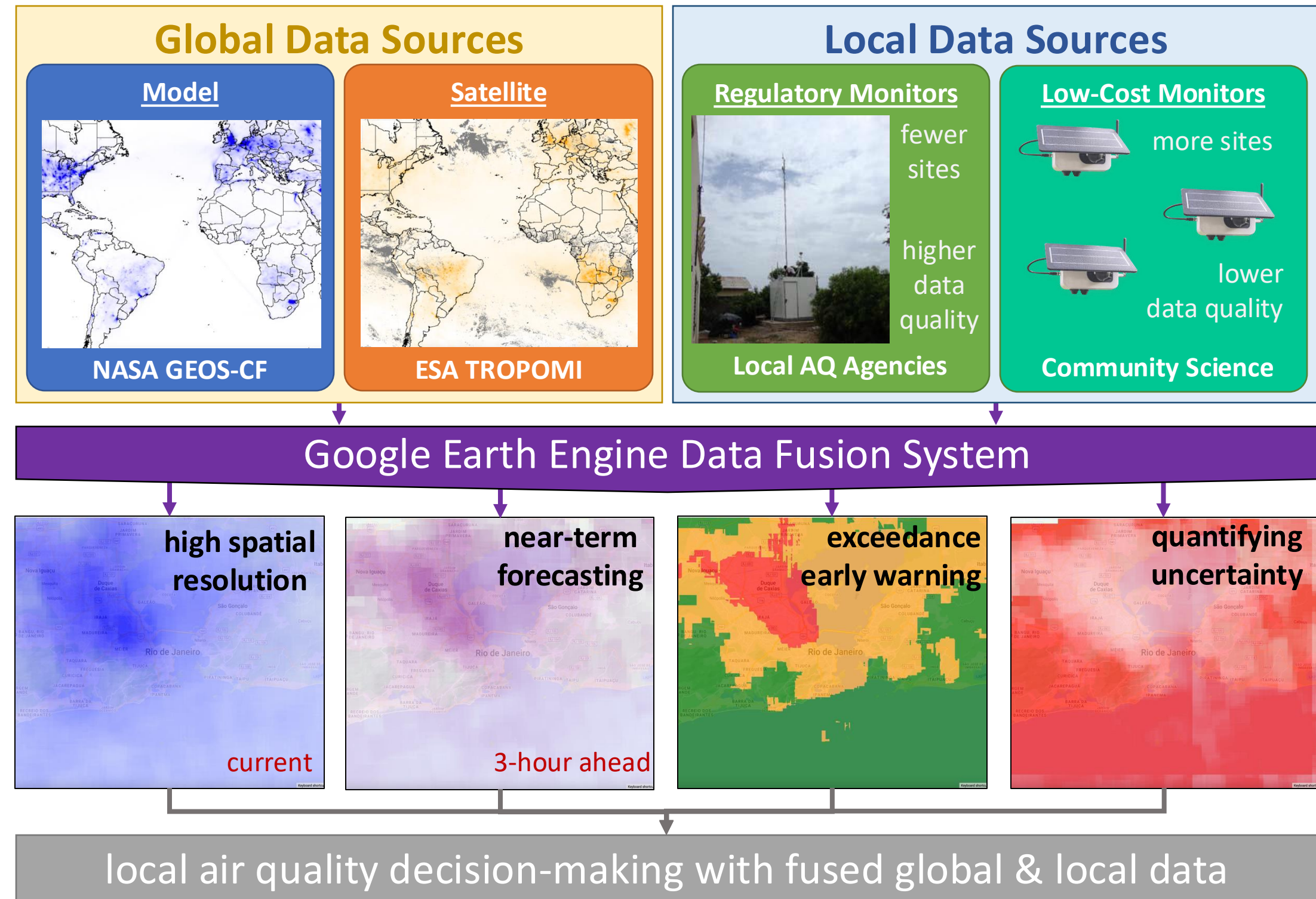
The goal of our project is to...

...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**...

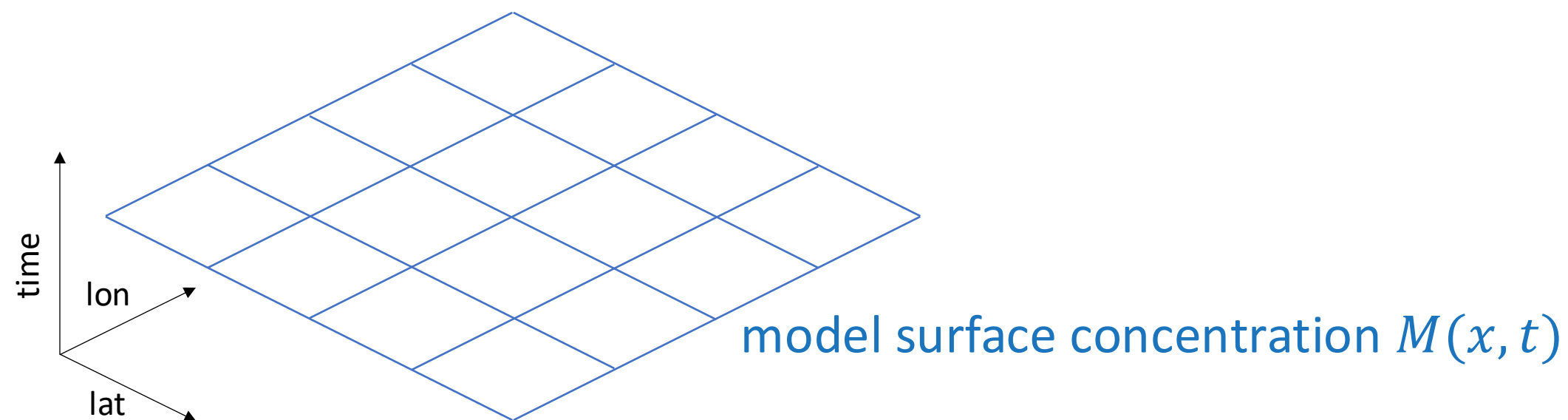
...freely accessible by air quality managers worldwide, facilitating their **decision-making** processes.



Source: NASA GMAO Science Snapshot [“Google Earth Engine Data Fusion Tool to support Air Quality Managers”](#)

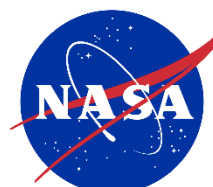
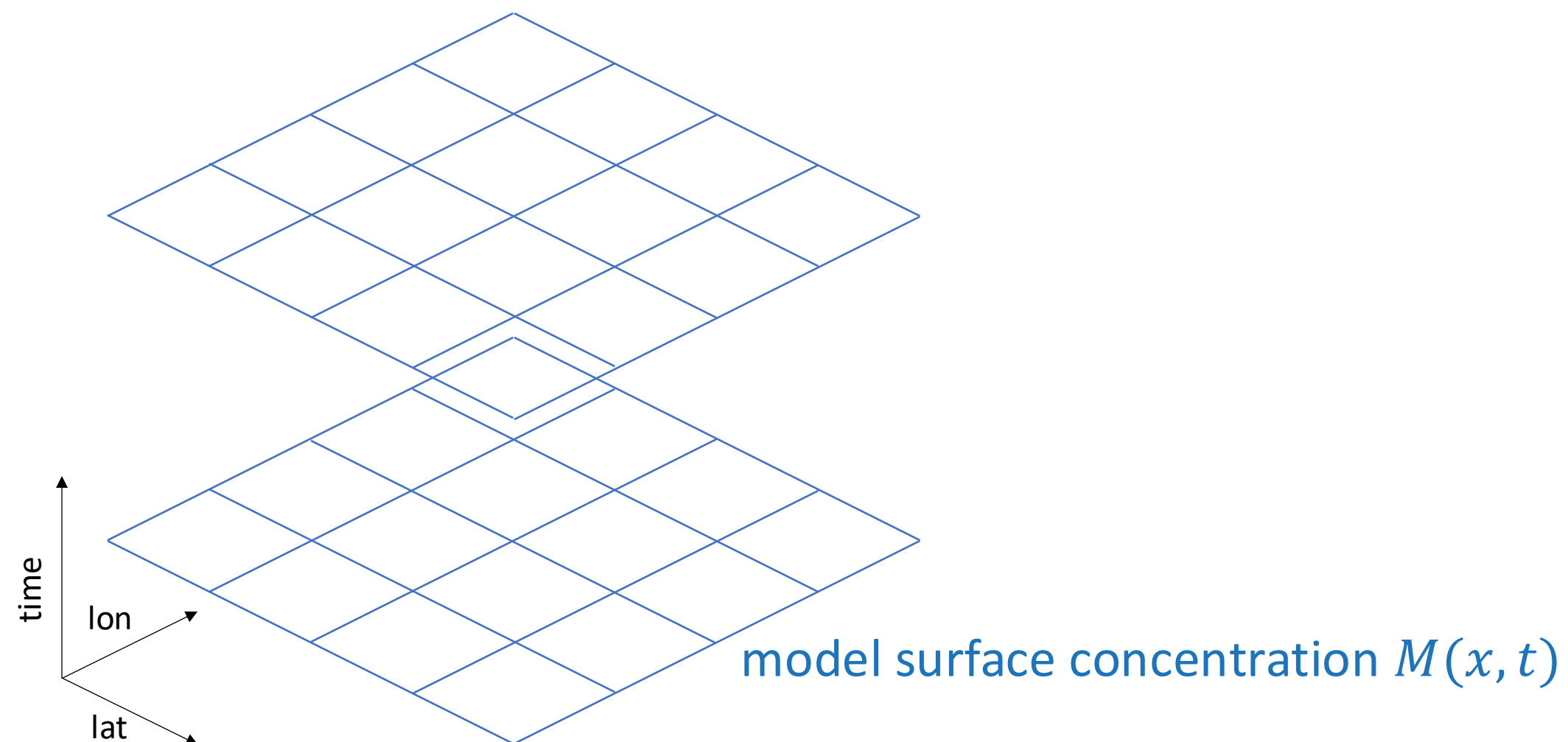
Phase 1: Model Only

Phase	Estimate
1	forecast model (GEOS-CF)



Phase 1: Model Only

Phase	Estimate
1	forecast model (GEOS-CF)



Partner

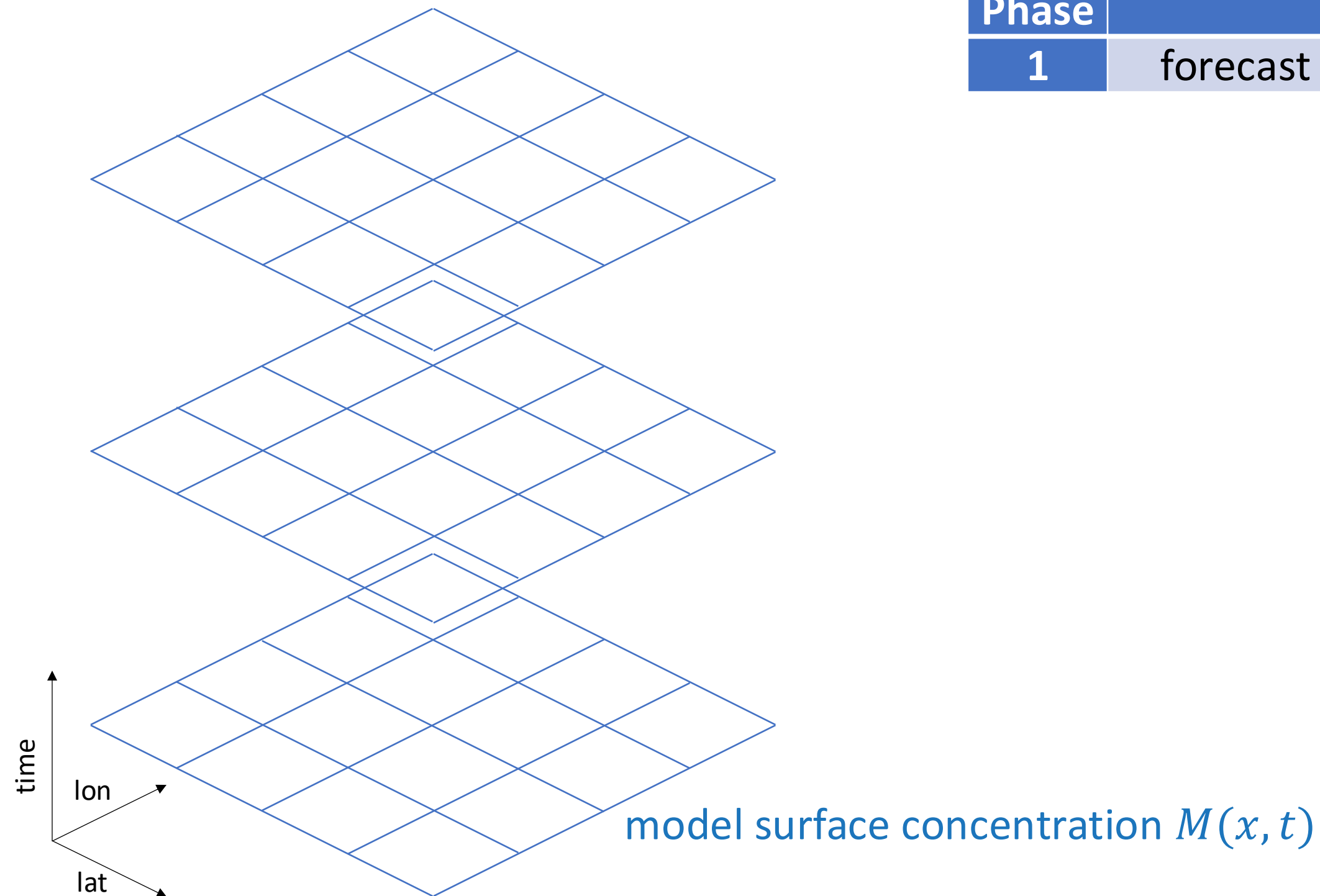


Presentation for GEEDEVs Nairobi Community
carl.a.malings@nasa.gov

17 Oct. 2024

Phase 1: Model Only

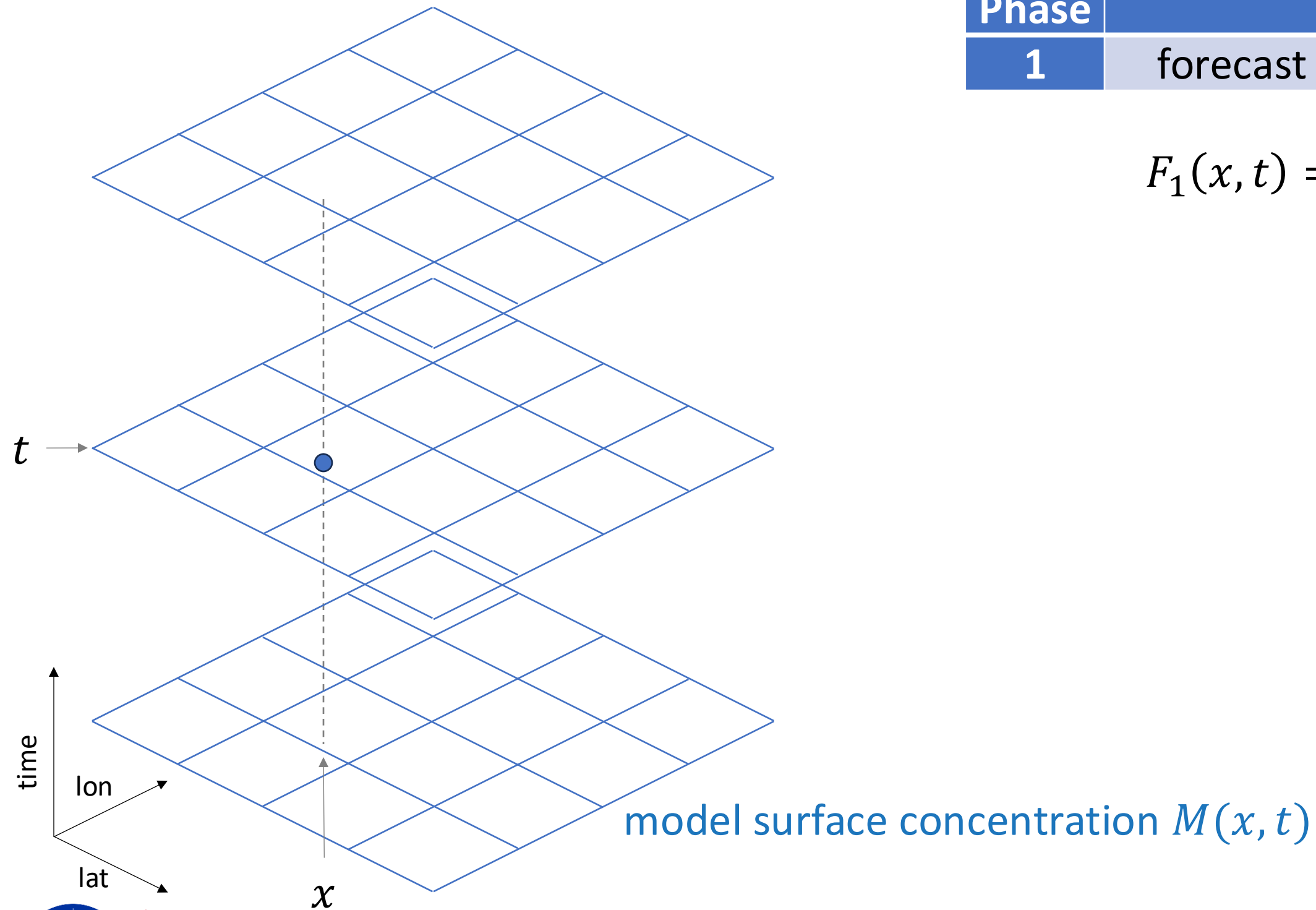
Phase	Estimate
1	forecast model (GEOS-CF)



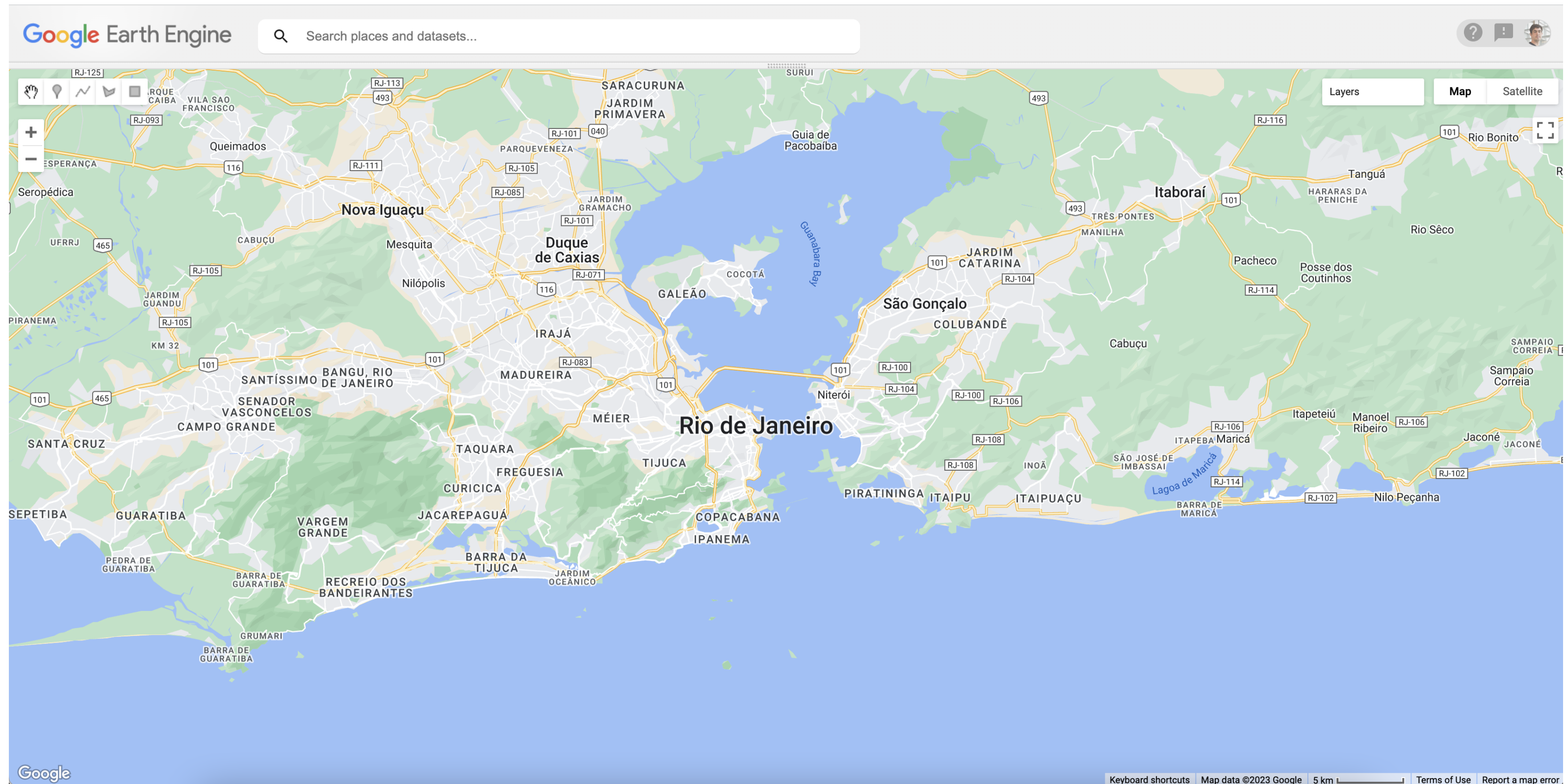
Phase 1: Model Only

Phase	Estimate
1	forecast model (GEOS-CF)

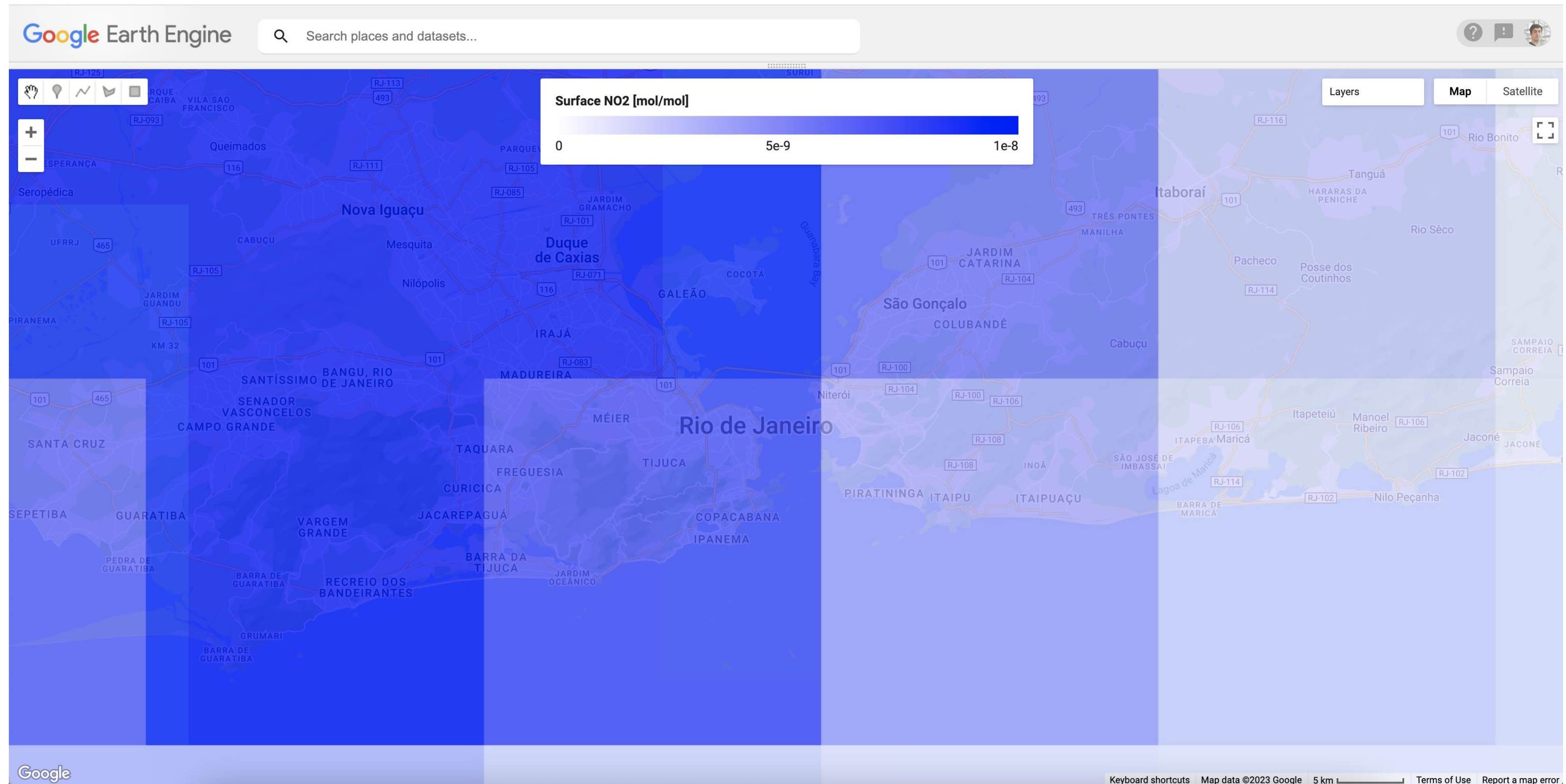
$$F_1(x, t) = M(x, t)$$



Phase 1: Model Only

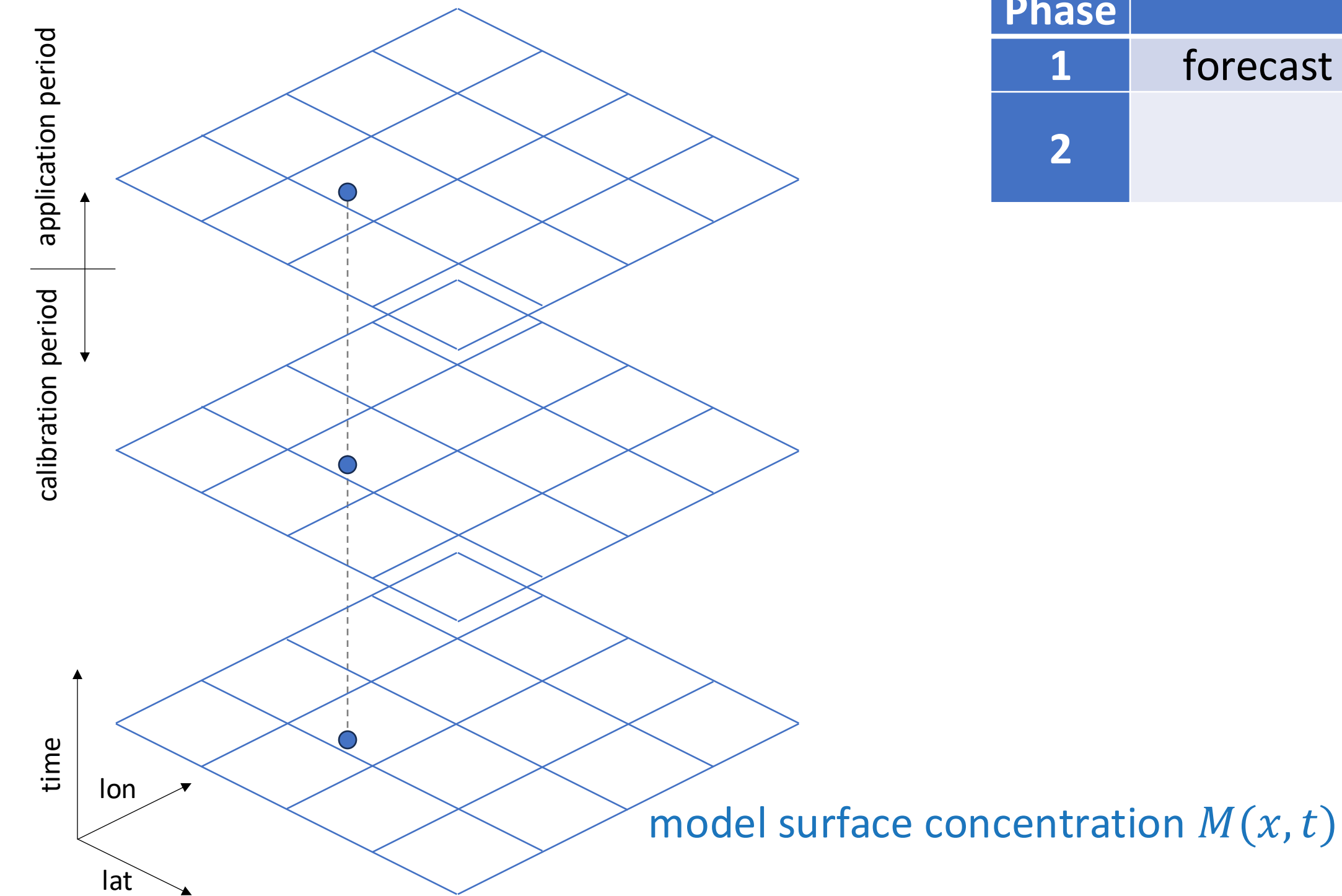


Phase 1: Model Only



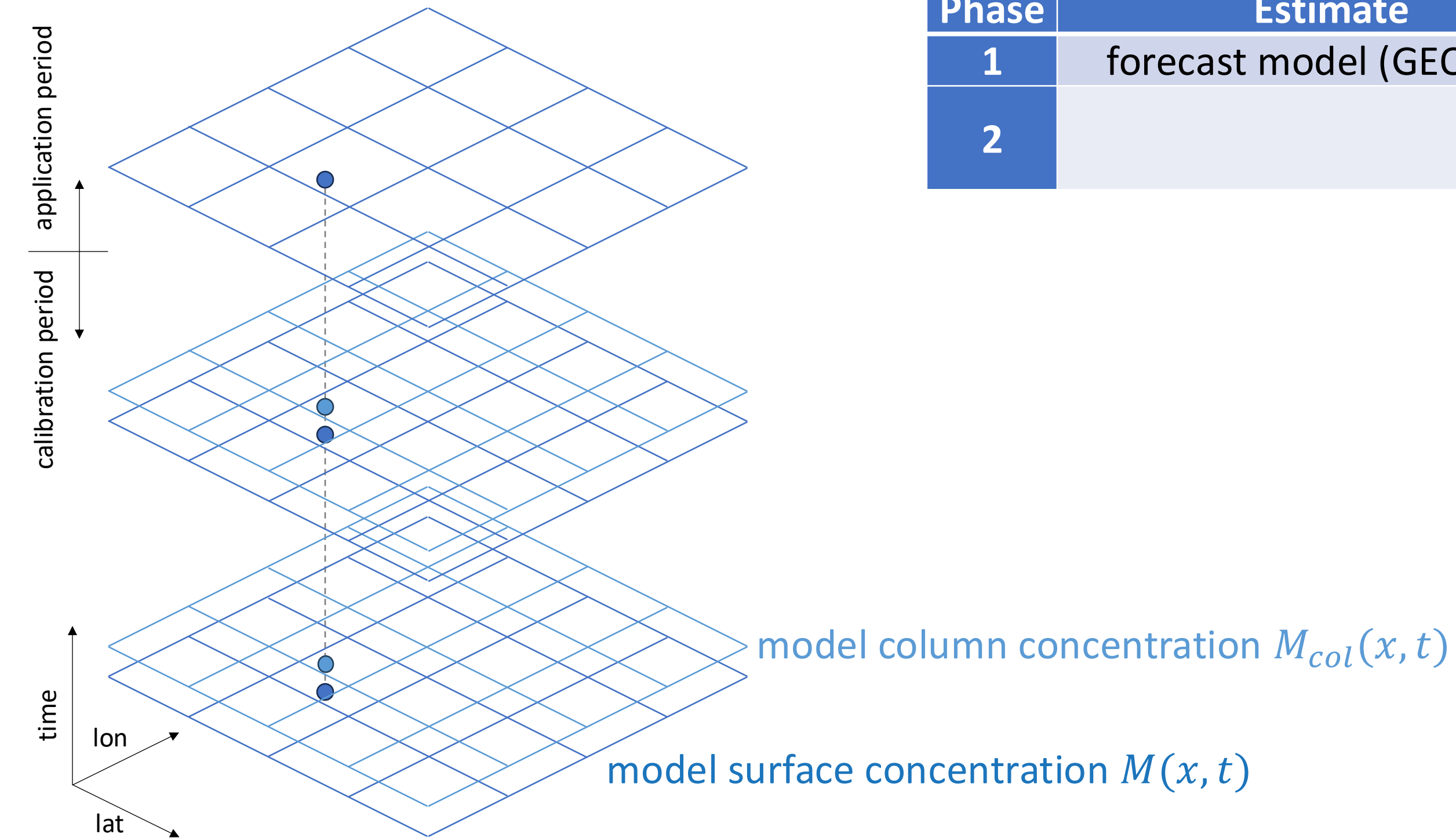
Phase 2: Model & Satellite

Phase	Estimate	Uncertainty
1	forecast model (GEOS-CF)	cell-to-cell variability of model
2		



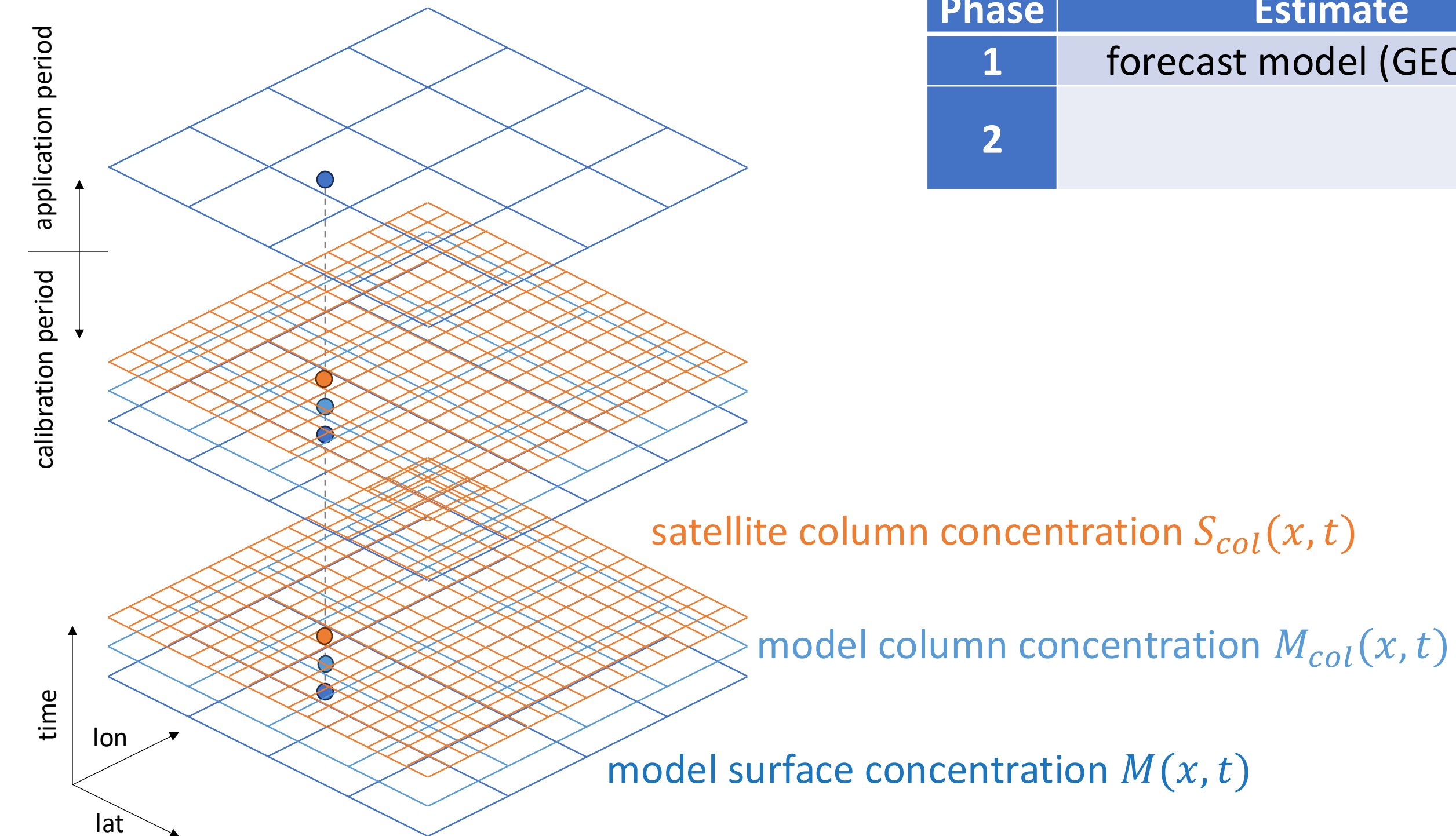
Phase 2: Model & Satellite

Phase	Estimate	Uncertainty
1	forecast model (GEOS-CF)	cell-to-cell variability of model
2		

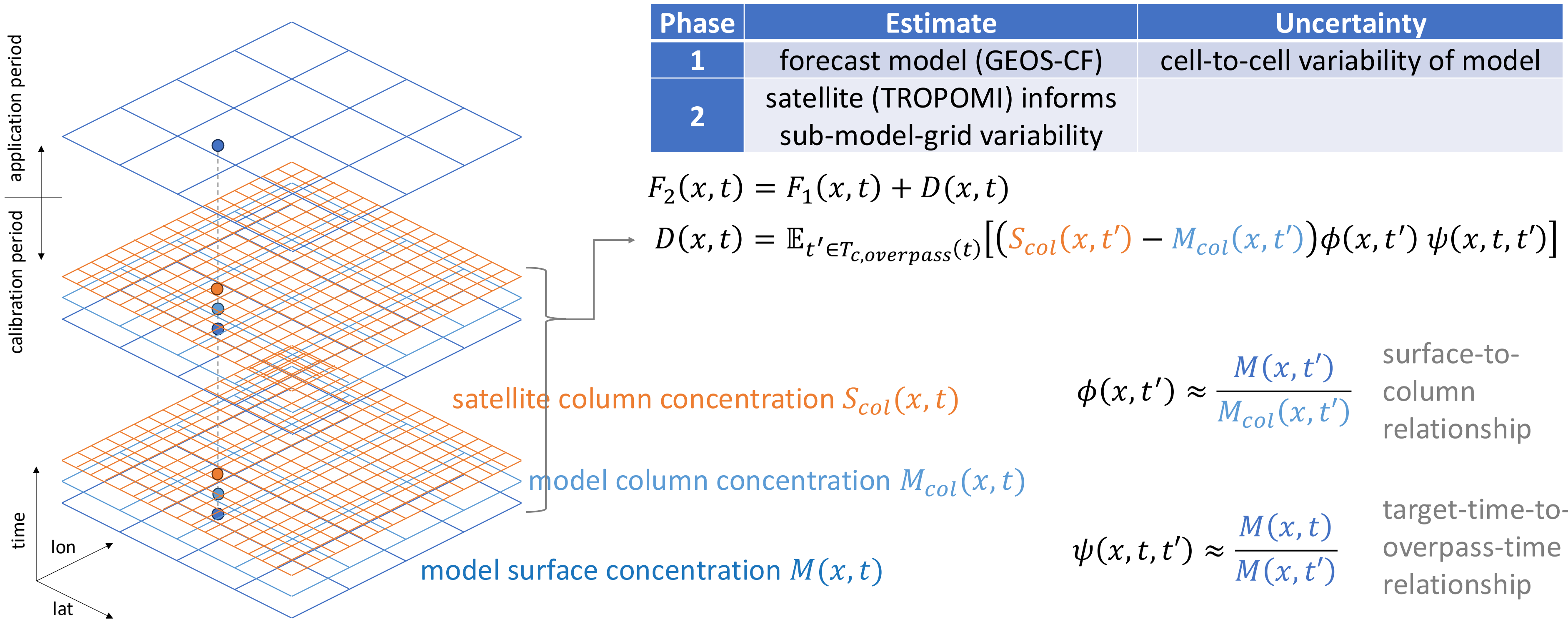


Phase 2: Model & Satellite

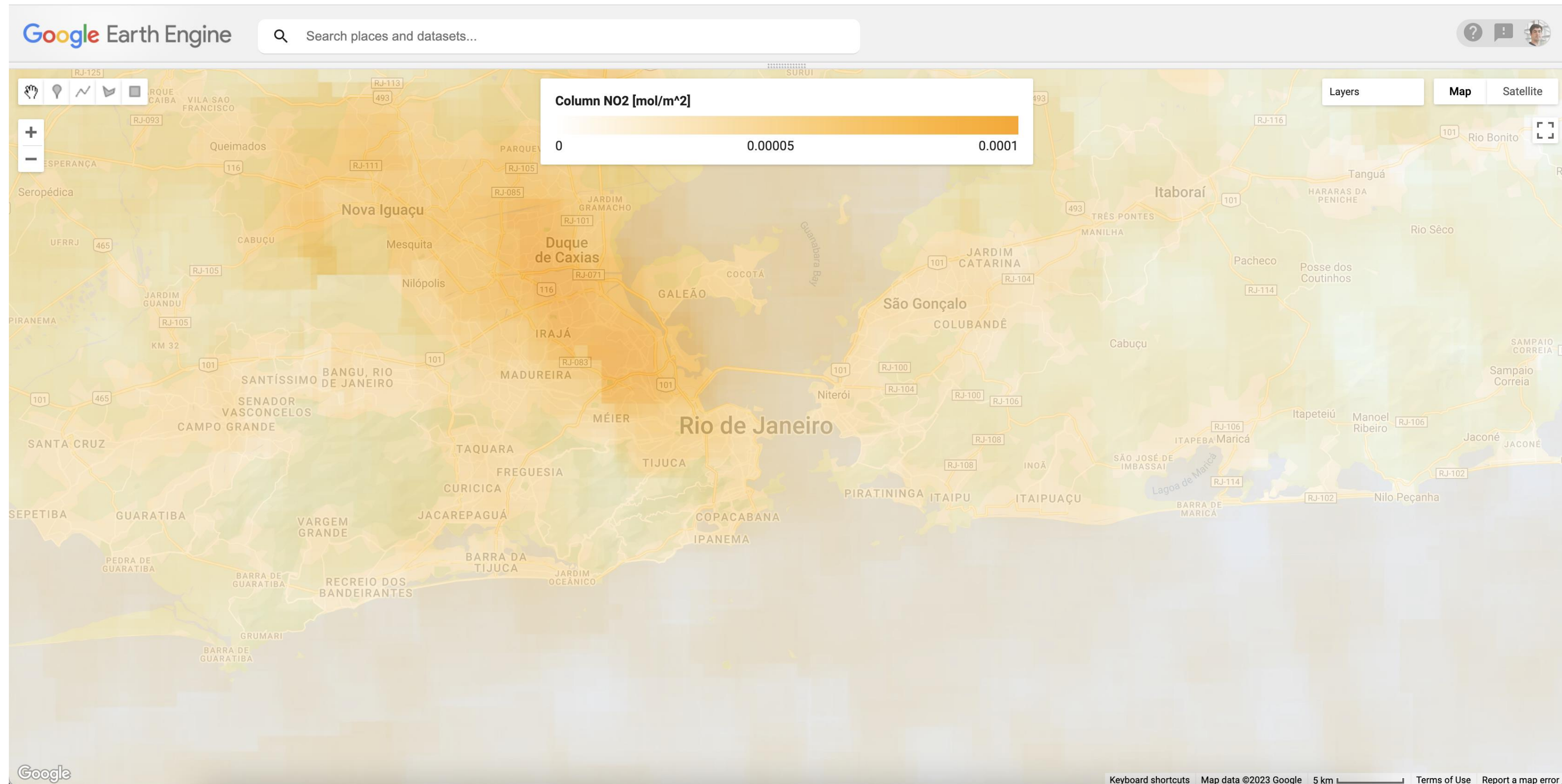
Phase	Estimate	Uncertainty
1	forecast model (GEOS-CF)	cell-to-cell variability of model
2		



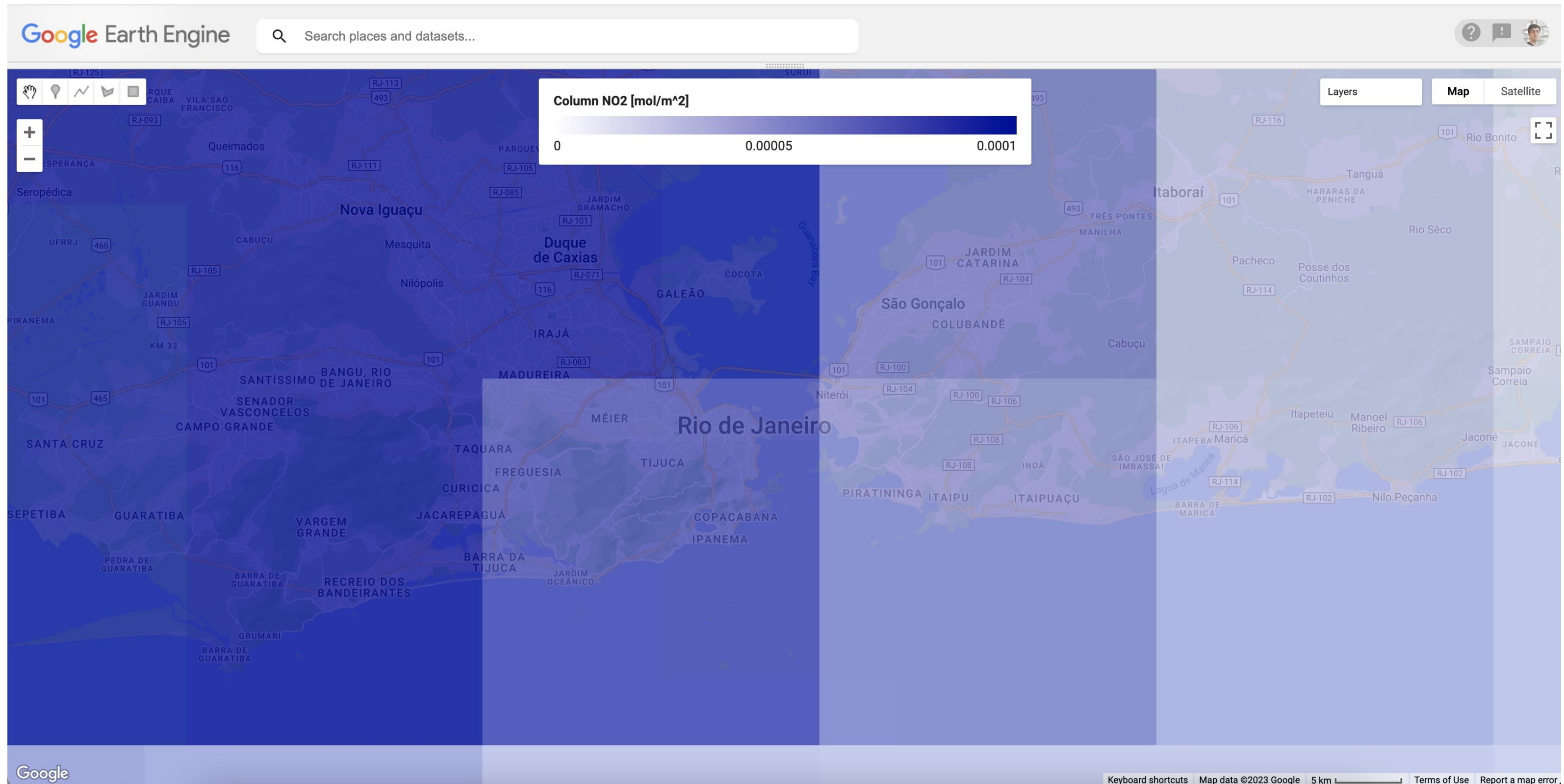
Phase 2: Model & Satellite



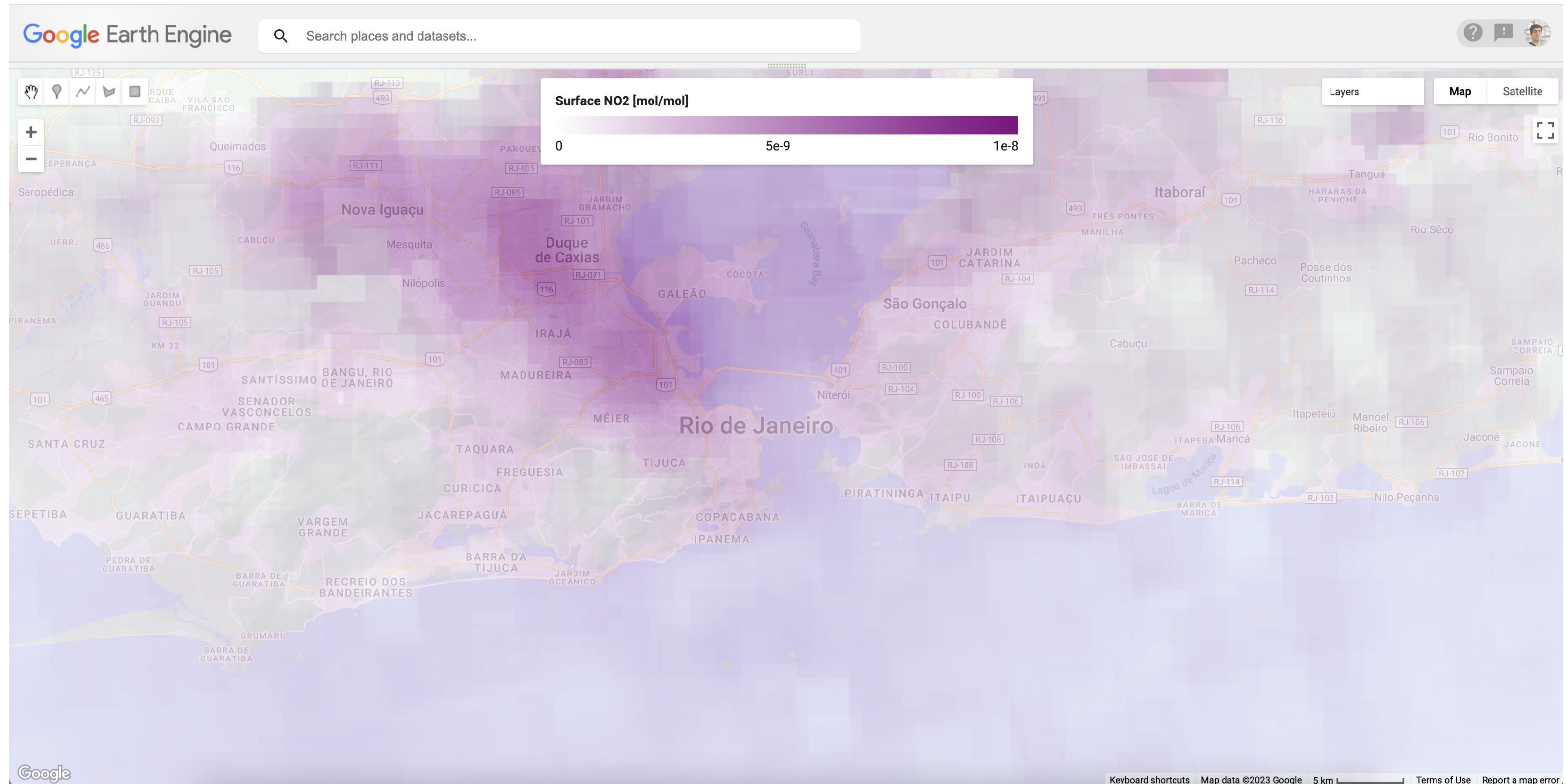
Phase 2: Model & Satellite



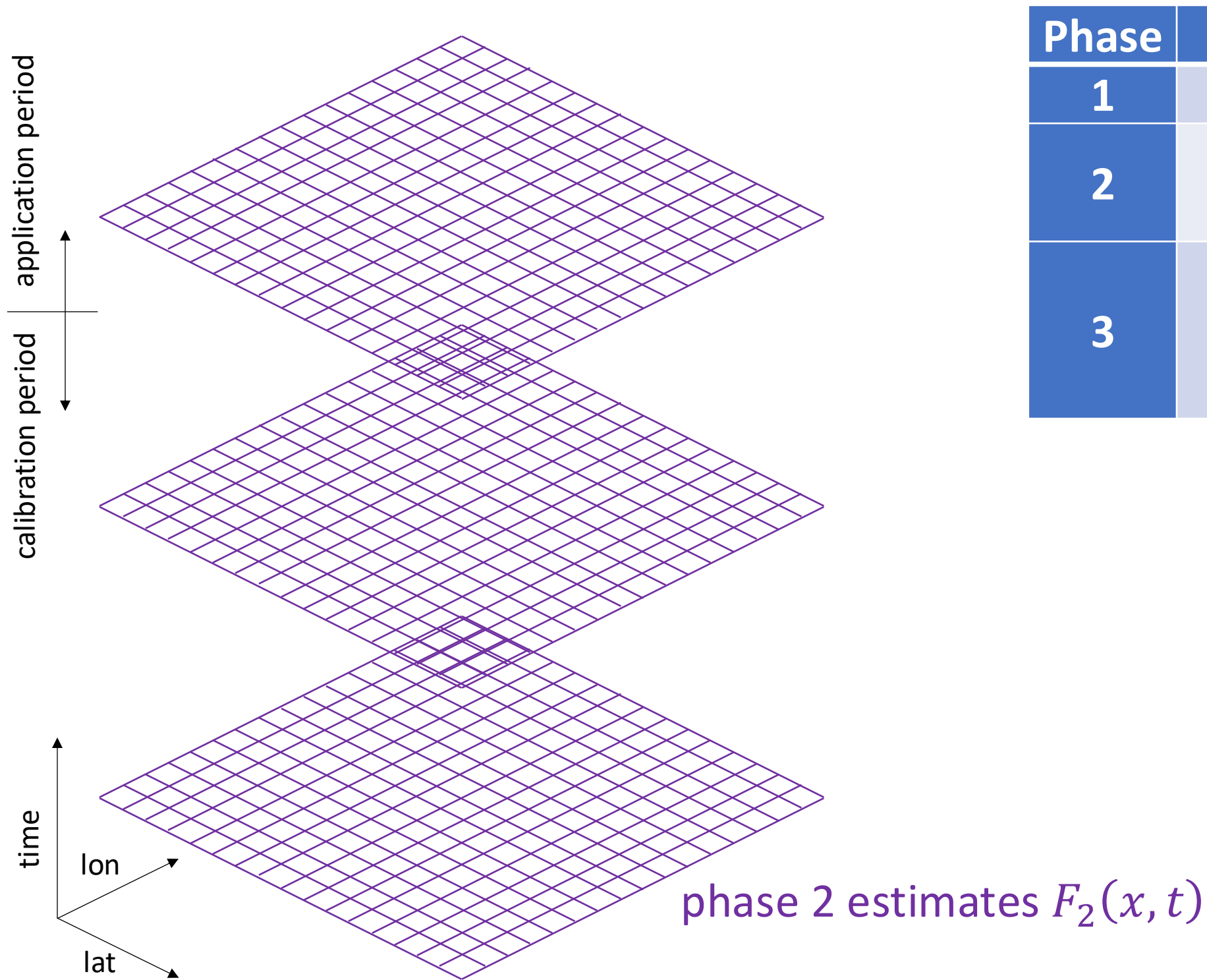
Phase 2: Model & Satellite



Phase 2: Model & Satellite

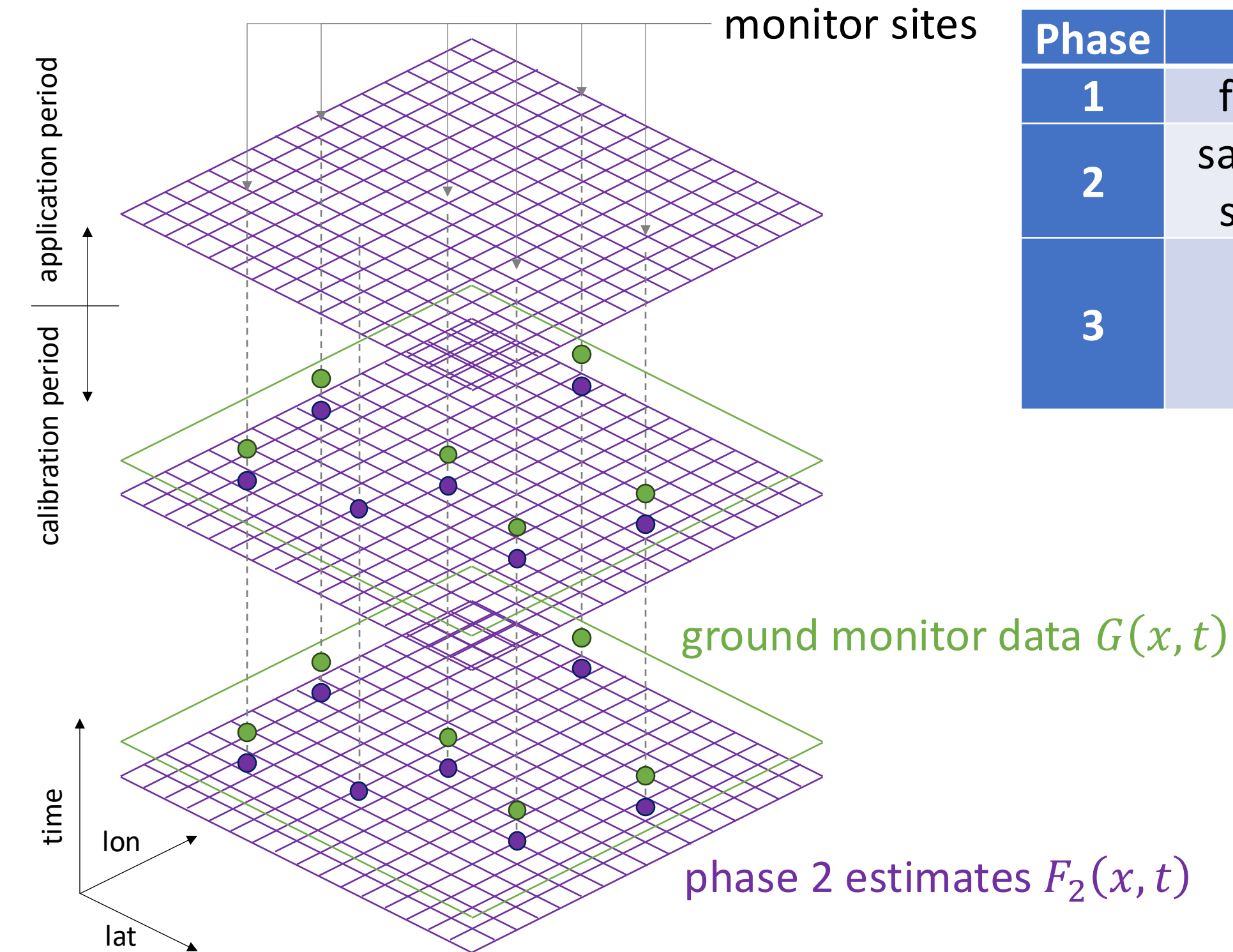


Phase 3: Model & Satellite & Ground



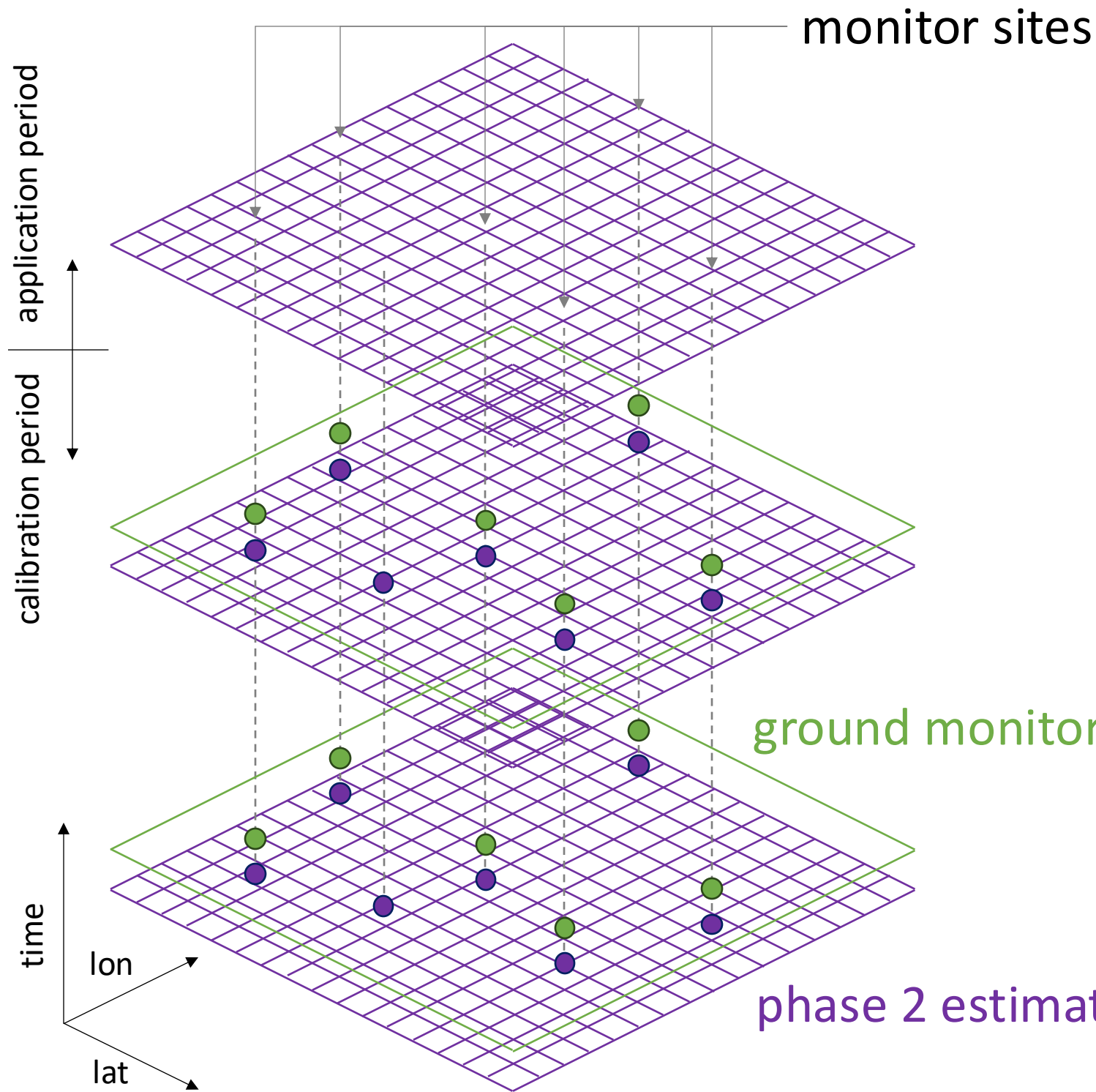
Phase	Estimate	Uncertainty
1	forecast model (GEOS-CF)	cell-to-cell variability of model
2	satellite (TROPOMI) informs sub-model-grid variability	satellite-to-model and surface-to-column ratios vary over time
3		

Phase 3: Model & Satellite & Ground



Phase	Estimate	Uncertainty
1	forecast model (GEOS-CF)	cell-to-cell variability of model
2	satellite (TROPOMI) informs sub-model-grid variability	satellite-to-model and surface-to-column ratios vary over time
3		

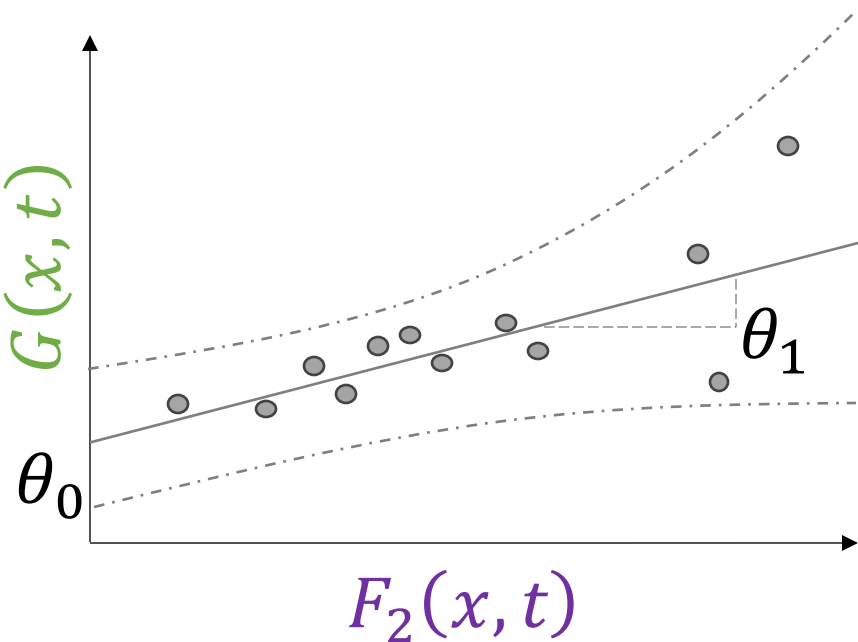
Phase 3: Model & Satellite & Ground



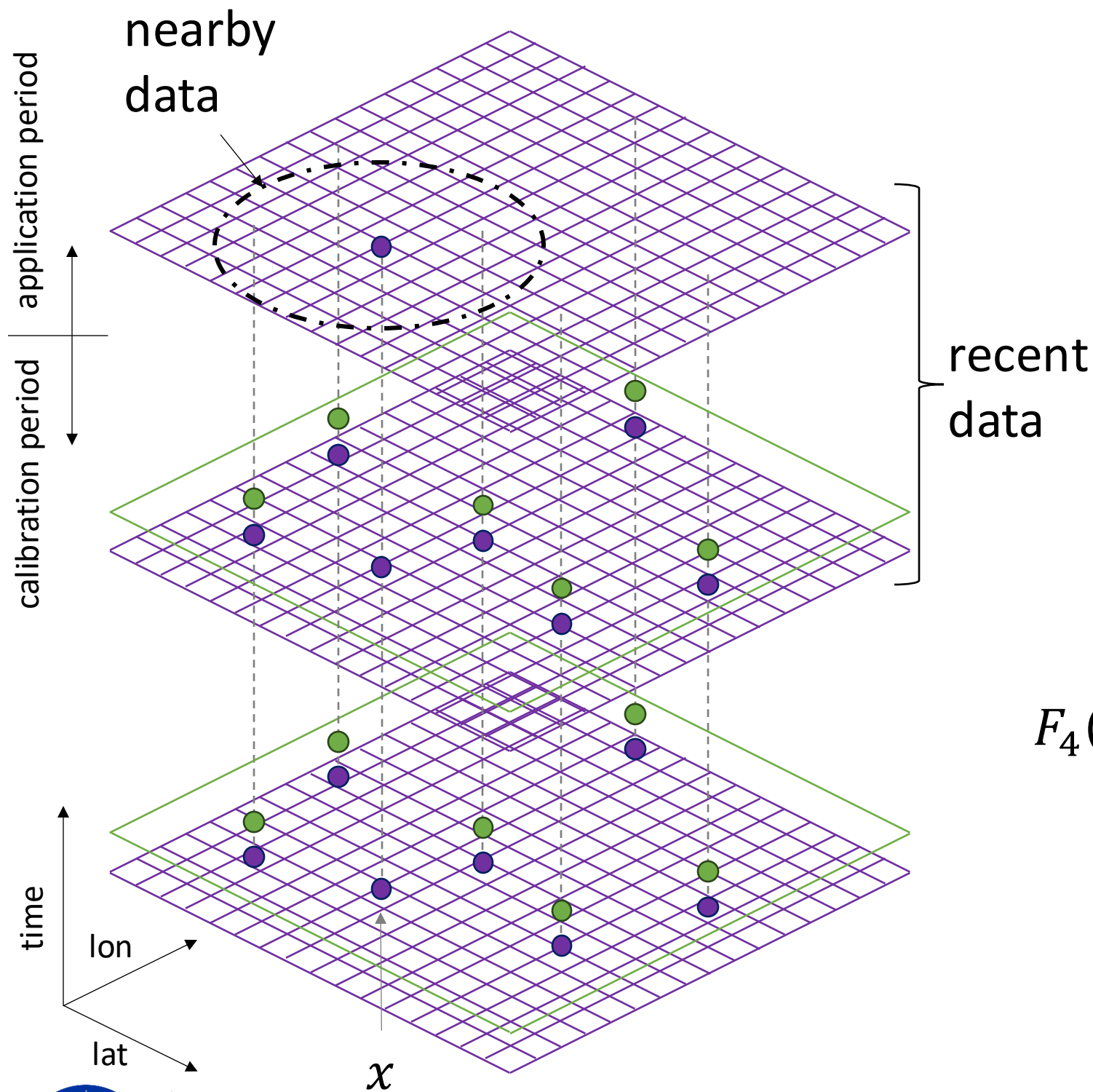
Phase	Estimate	Uncertainty
1	forecast model (GEOS-CF)	cell-to-cell variability of model
2	satellite (TROPOMI) informs sub-model-grid variability	satellite-to-model and surface-to-column ratios vary over time
3	phase 2 corrected to match surface monitor data	

$$F_3(x, t) = \theta_1 F_2(x, t) + \theta_0$$

$$\theta_0, \theta_1 = \mathbb{LR}_{t' \in T_c(t), x' \in X_c(x)} [G(x', t') \sim F_2(x', t')]$$

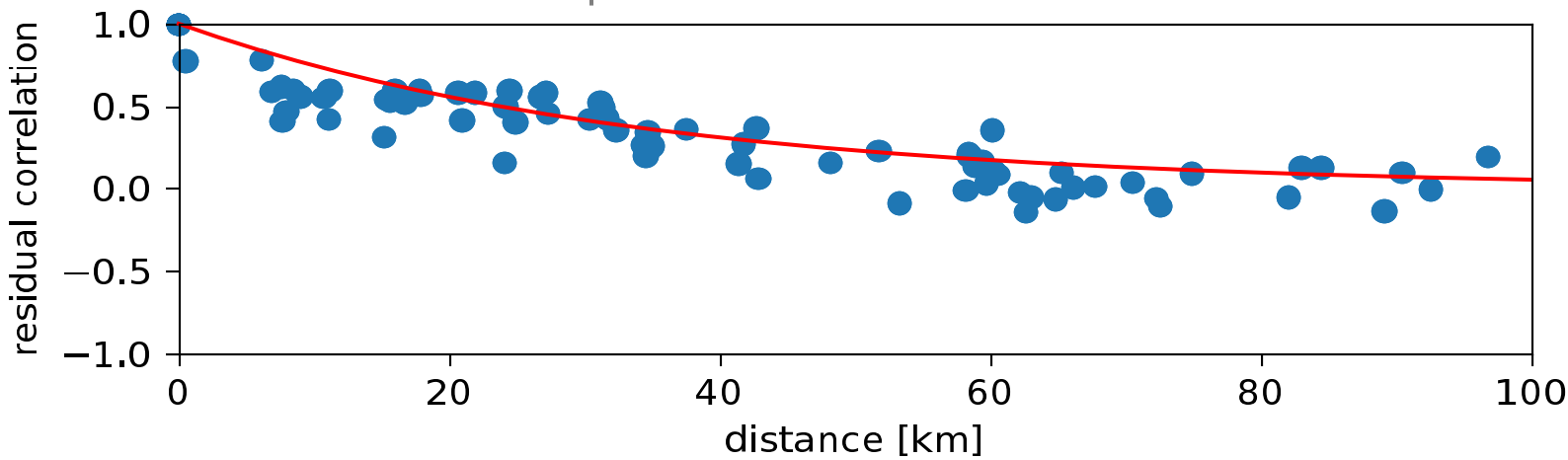


Phase 4: Residual Kriging

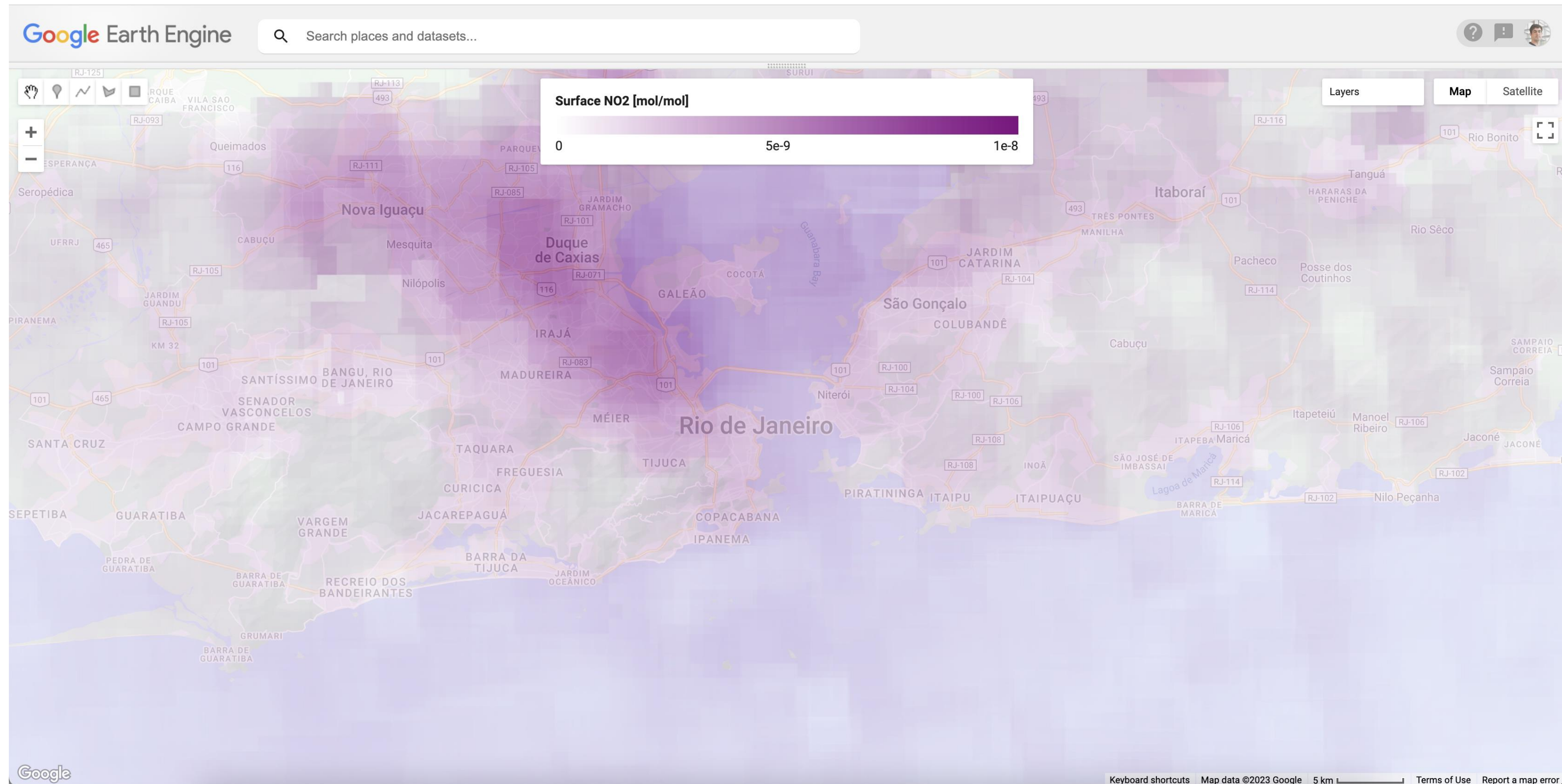


Phase	Estimate	Uncertainty
1	forecast model (GEOS-CF)	cell-to-cell variability of model
2	satellite (TROPOMI) informs sub-model-grid variability	satellite-to-model and surface-to-column ratios vary over time
3	phase 2 corrected to match surface monitor data	uncertain regression parameters between phase 2 output and surface monitor data
4	update phase 3 based on recent surface monitor data	

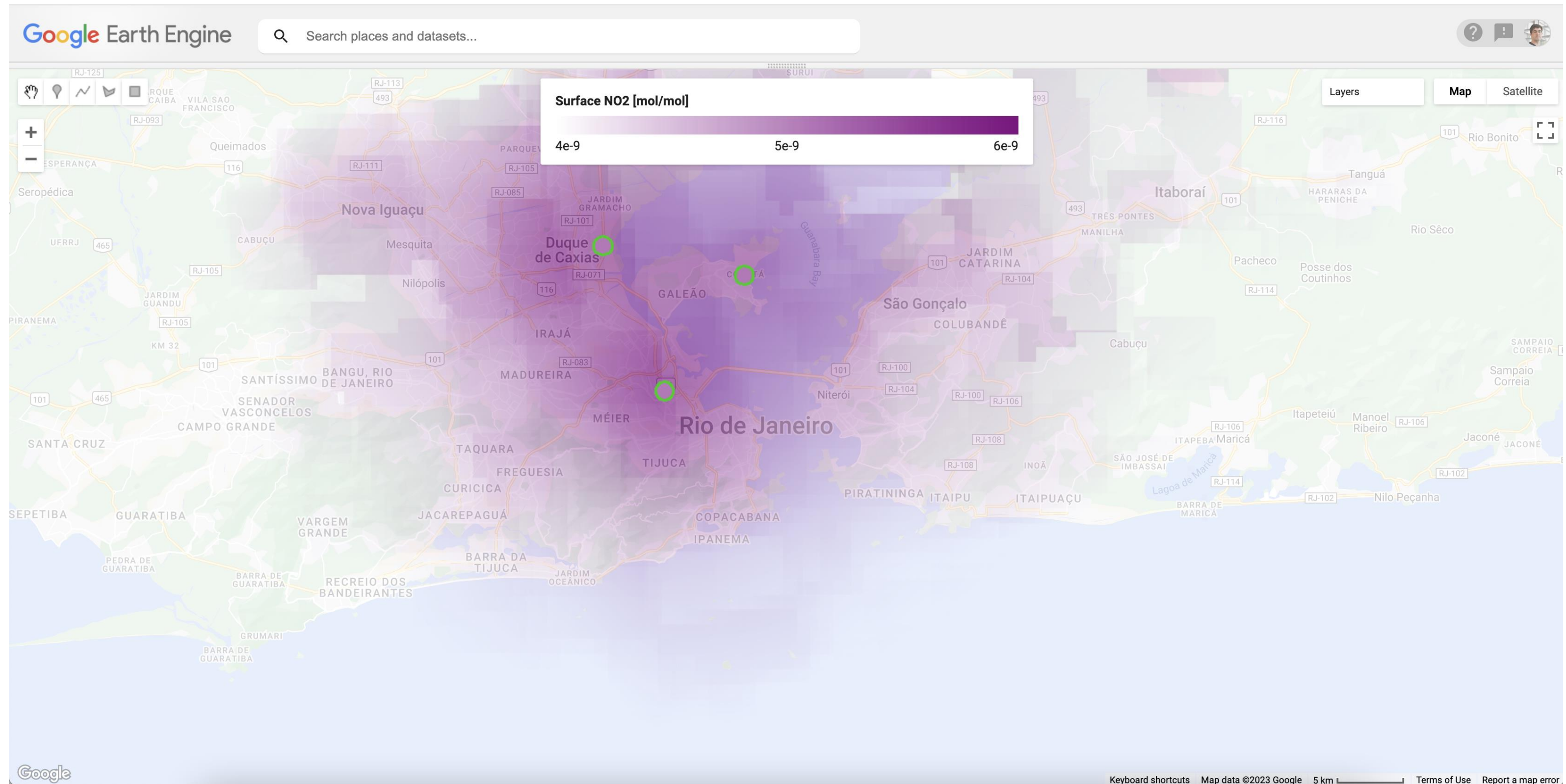
$$F_4(x, t) = F_3(x, t) + \sum_{x' \in X_n(x), t' \in T_n(t)} K(x, x', t, t') [G(x', t') - F_3(x', t')]$$



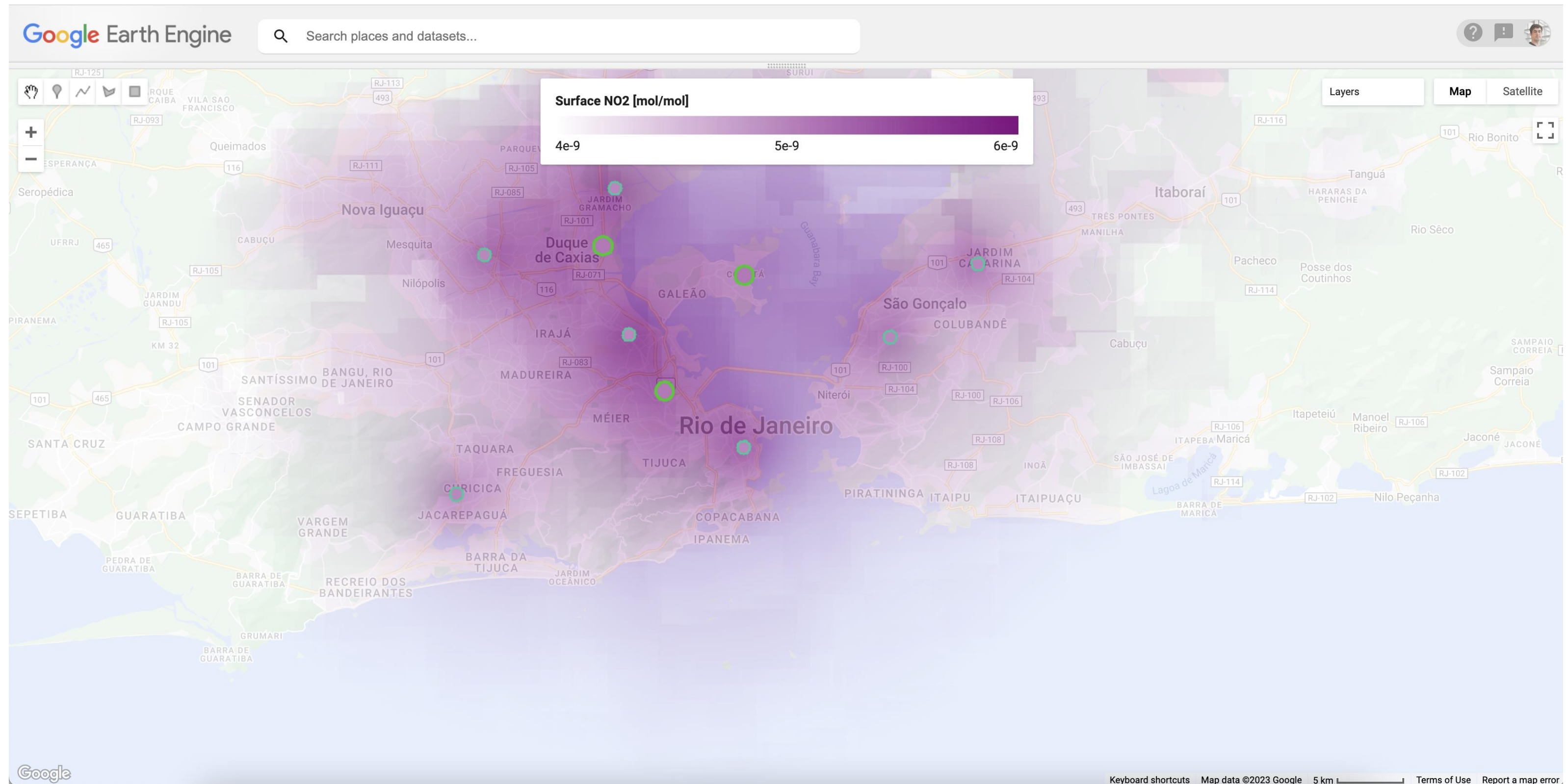
Phase 2: Model & Satellite



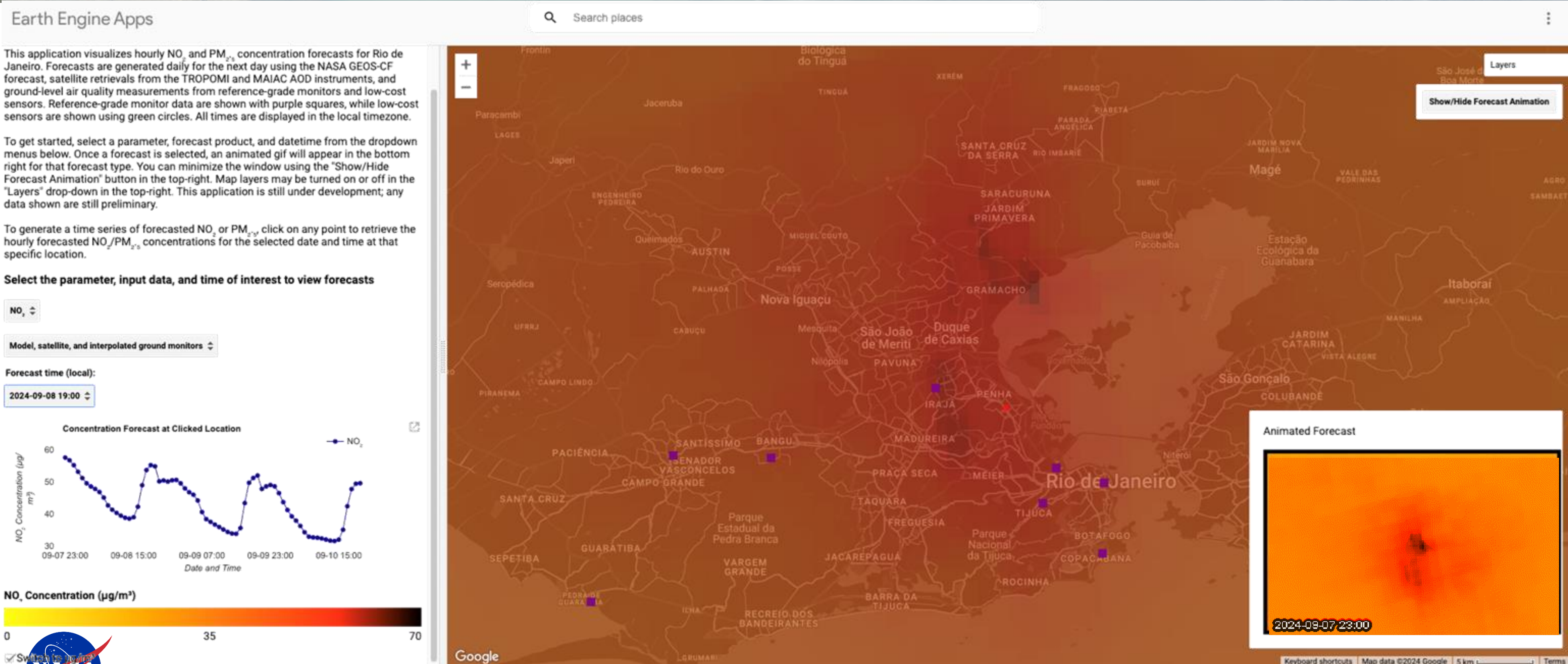
Phase 3 & Phase 4 (simulation)



Phase 3 & Phase 4 (simulation)



Prototype User Interface



Pros and Cons of Google Earth Engine

- **Data Availability and Accessibility**

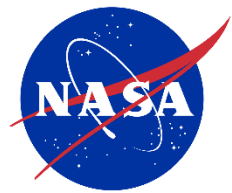
- + Many air-quality-relevant Earth observations datasets are already available
- + Many interoperability issues (resolution, coordinate systems) are solved “under the hood”
- In-situ air quality data are yet not available/accessible

- **Computation**

- + Efficient computation over large spatial and (to a lesser extent) temporal scales
- Interfacing with in-situ data (e.g., extracting raster data at vector coordinates) is inefficient
- Costs associated with operational computations introduced since the project started

- **User Interface**

- + Ready-made interface options allow for most common display functions
- Some unforeseen issues (display scale factors, automatic local time adjustment)



Partner



Further Resources

NASA Applied Remote Sensing Training Program (ARSET)

<https://appliedsciences.nasa.gov/arset>

ARSET provides accessible, relevant, and cost-free training on remote sensing satellites, sensors, methods, and tools.

Our trainings are:

- Online and in-person
- Open to everyone
- Live, instructor-led, or self-guided
- Provided at no cost, with materials and recordings available from our website
- Often multi-lingual
- Tailored to those with a range of experience in remote sensing, from **introductory** to **advanced**



Training Themes



AGRICULTURE



CLIMATE & RESILIENCE



DISASTERS



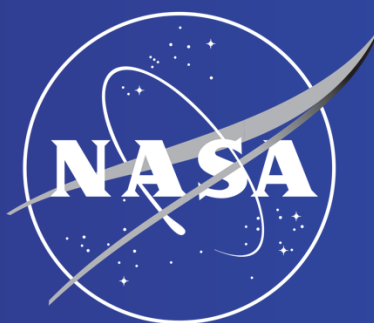
ECOLOGICAL CONSERVATION



HEALTH & AIR QUALITY



WATER RESOURCES



EARTHDATA Offers

The Air Quality Data Pathfinder

for Your Research & Applications

Air pollution is one of the largest global environmental and health threats. NASA provides data resources to better understand the movement of pollutants and the impact of events leading to poor air quality. This Pathfinder helps you access, and leverage data acquired from NASA's satellite, airborne, and ground-based missions and campaigns.

Available Data Types:

- Aerosols
- Trace Gases (e.g., Nitrogen Dioxide, Sulfur Dioxide, Carbon Monoxide, etc.)
- Weather (e.g., Air Temperature, Clouds, Precipitation, etc.)
- Land Surface (e.g., Soil Moisture, Surface Reflectance, Topography, etc.)
- Human Dimensions

Data are from satellites, airborne and ground-based platforms, and models, including:

- | | |
|------------|-----------|
| • AIRS | • OMPS |
| • AMSR2 | • SMAP |
| • GPM | • TROPOMI |
| • MODIS | • VIIRS |
| • OLI/TIRS | • GEOS |
| • OMI | • MERRA-2 |



Visit the EARTHDATA
Air Quality Data Pathfinder
for more information:

- Commonly Used Datasets for Air Quality Research and Applications
- Tools for Using Data
- Resources for Applying and Connecting NASA Data
- GIS Resources
- Tips for Getting Help and Connecting with NASA experts
- Tutorials and more!



Health and Air Quality Applied Science Team

<https://haqast.org/>

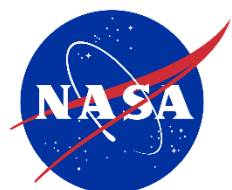
“Our goal is to use NASA’s data and satellites to pursue cutting edge applied research in order to keep you healthy and safe.”

- Use NASA satellite & other data to help solve real-world public health and air quality problems.
- Work around the world on diverse issues related to health and air quality.
- Collaborate with public stakeholders to help guide long-term research.
- “Tiger Teams” pursue short-term, high-impact projects in small groups.



Getting started with NASA satellite data for health and air quality:

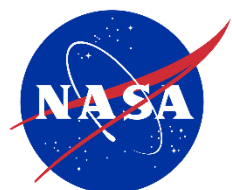
<https://haqast.org/getting-started/>



Partner



Thank you!



Partner



Backup Slides

Satellite and Instrument Characteristics to Consider

Temporal Resolution

Polar Orbit: ≥ 1 day (days, weeks)

Geostationary Orbit: < 1 day
(minutes, hours)

Geostationary Orbit

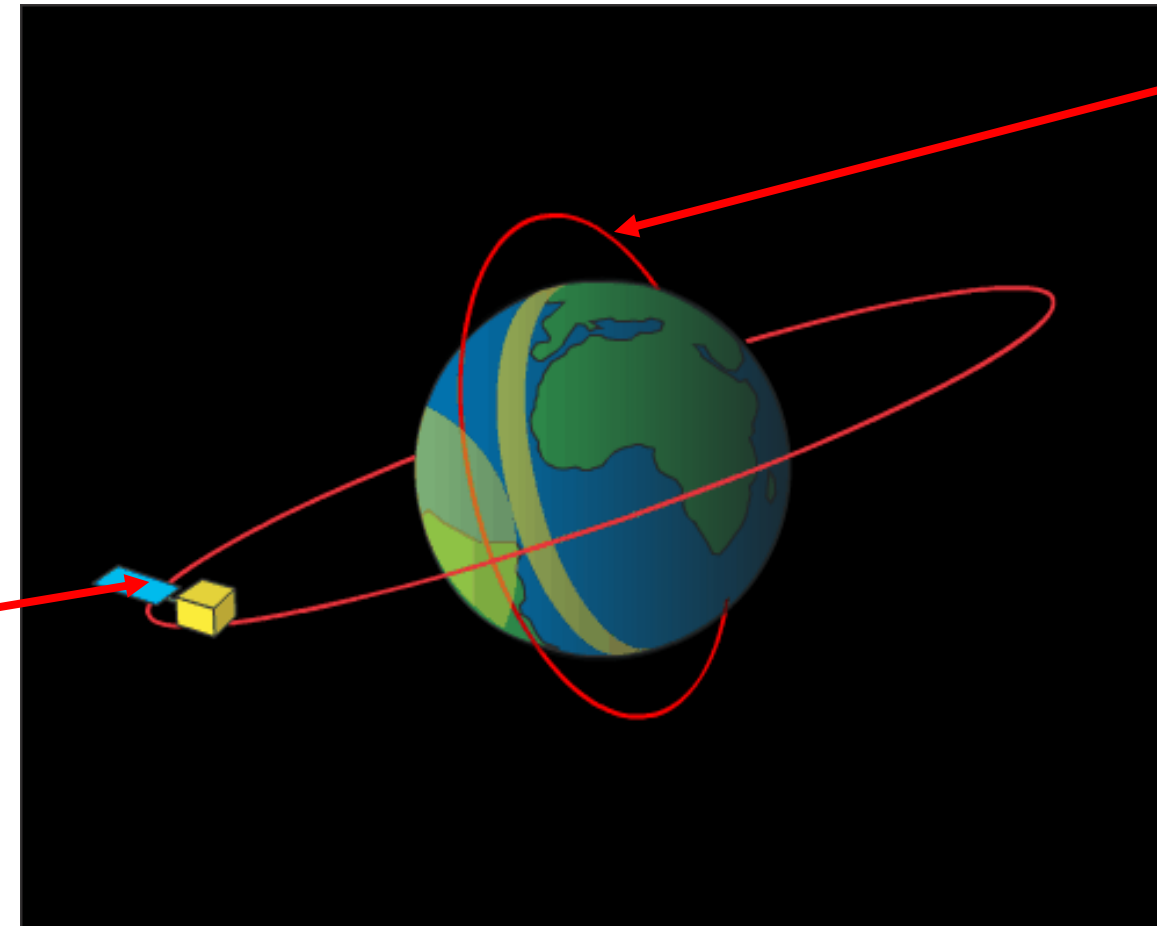
Coverage

Polar Orbit (swath):

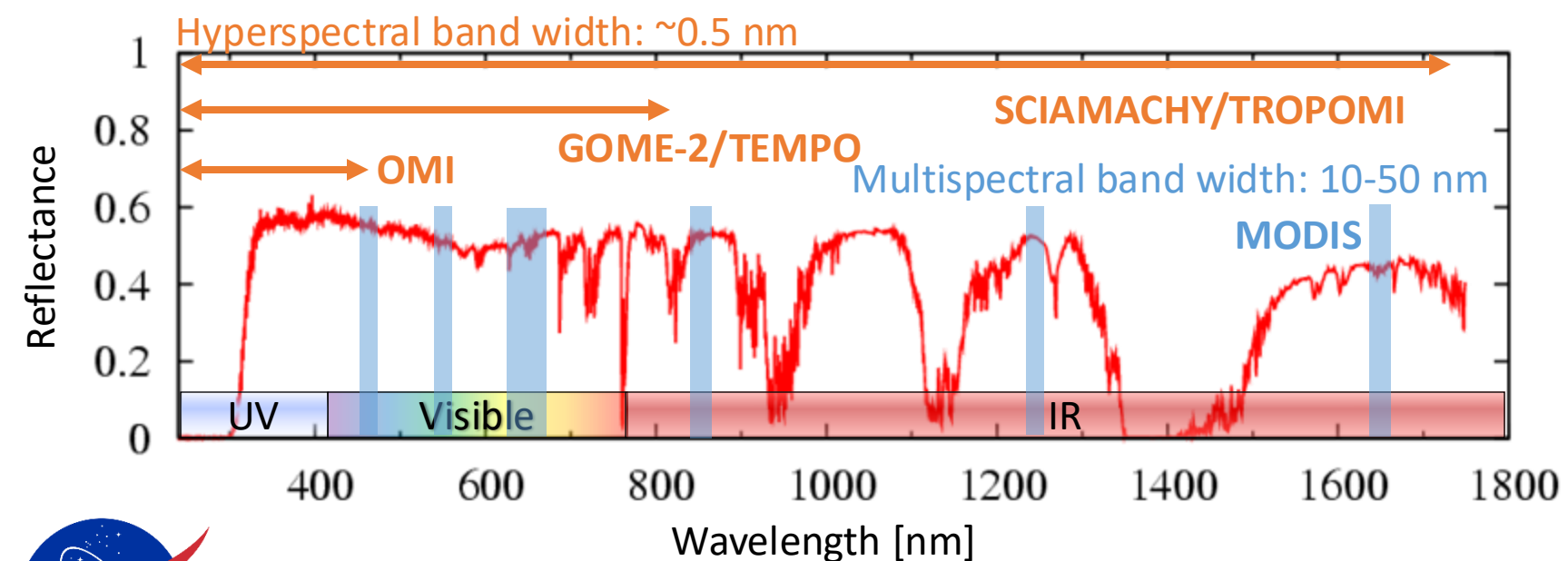
global coverage over many orbits

Geostationary Orbit (field of regard):

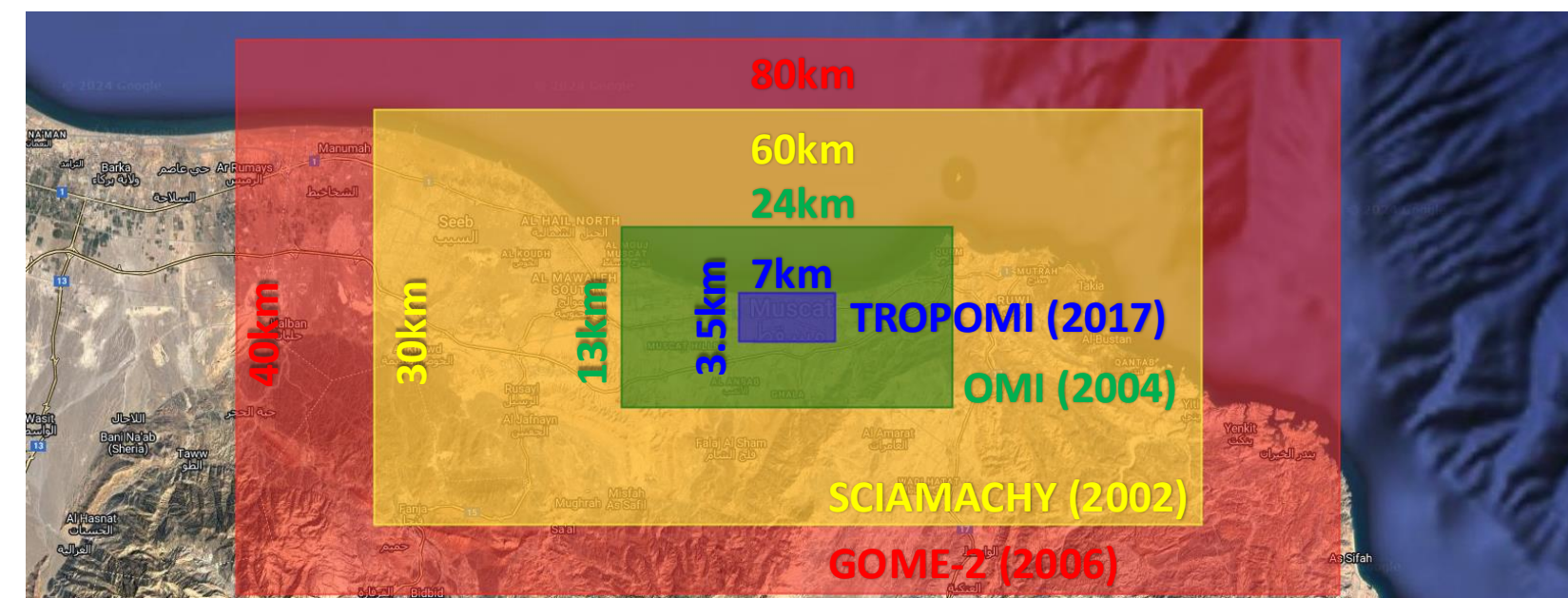
hemisphere or regional coverage



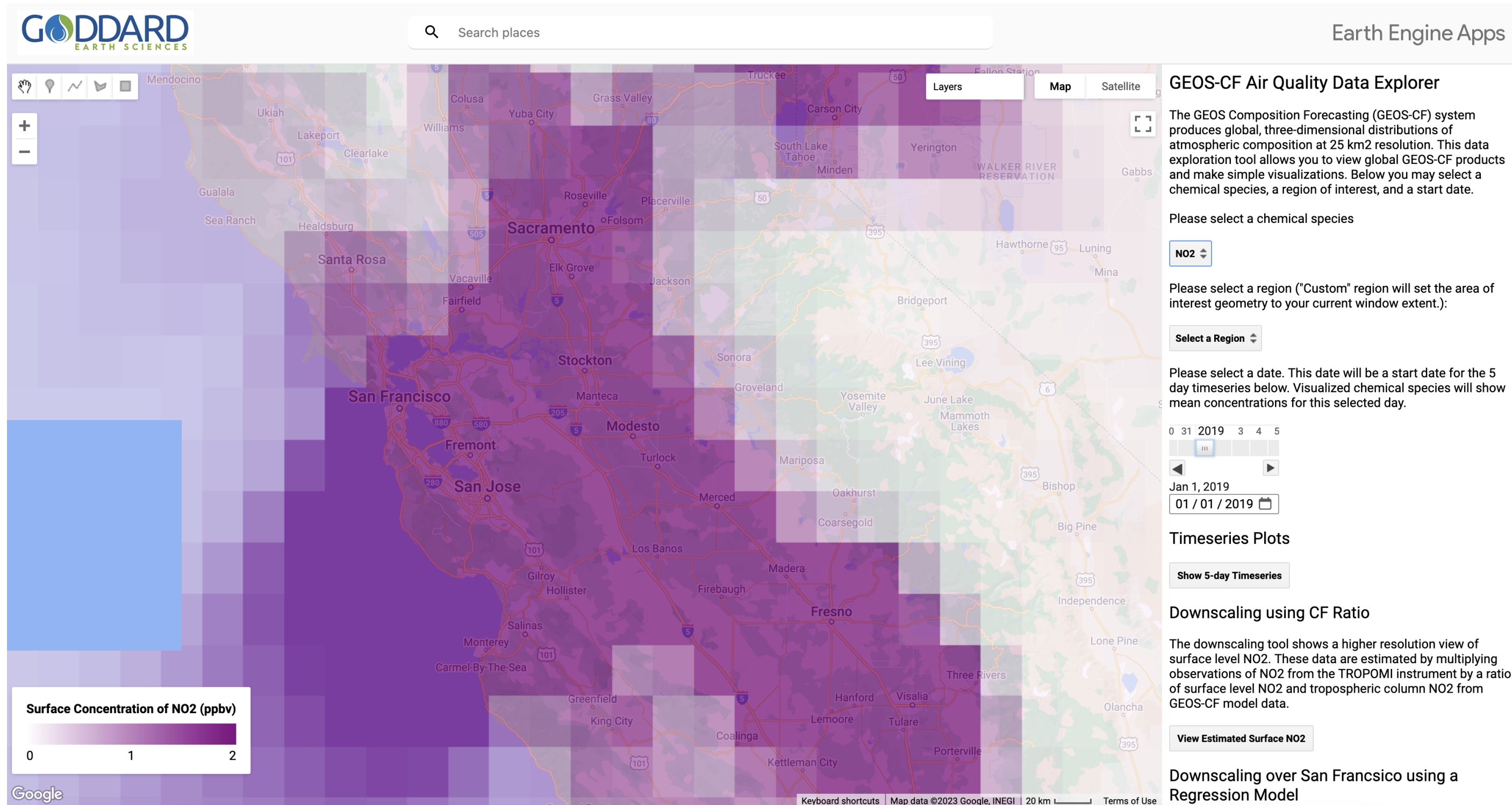
Spectral & Radiometric Resolution



Spatial Resolution



GMAO products in Google Earth Engine



Part of the NASA-Google Partnership, some GMAO products are being ingested into Google Earth Engine:

- **GEOS-CF**
- **MERRA-2**

Google Earth Engine and Google Cloud tools can apply statistical methods and machine learning techniques to downscale and fuse data products, improving biases at sub-grid scale.

Source: [NASA GMAO Science Snapshot "A New Generation of GMAO Apps in Google Earth Engine"](#)

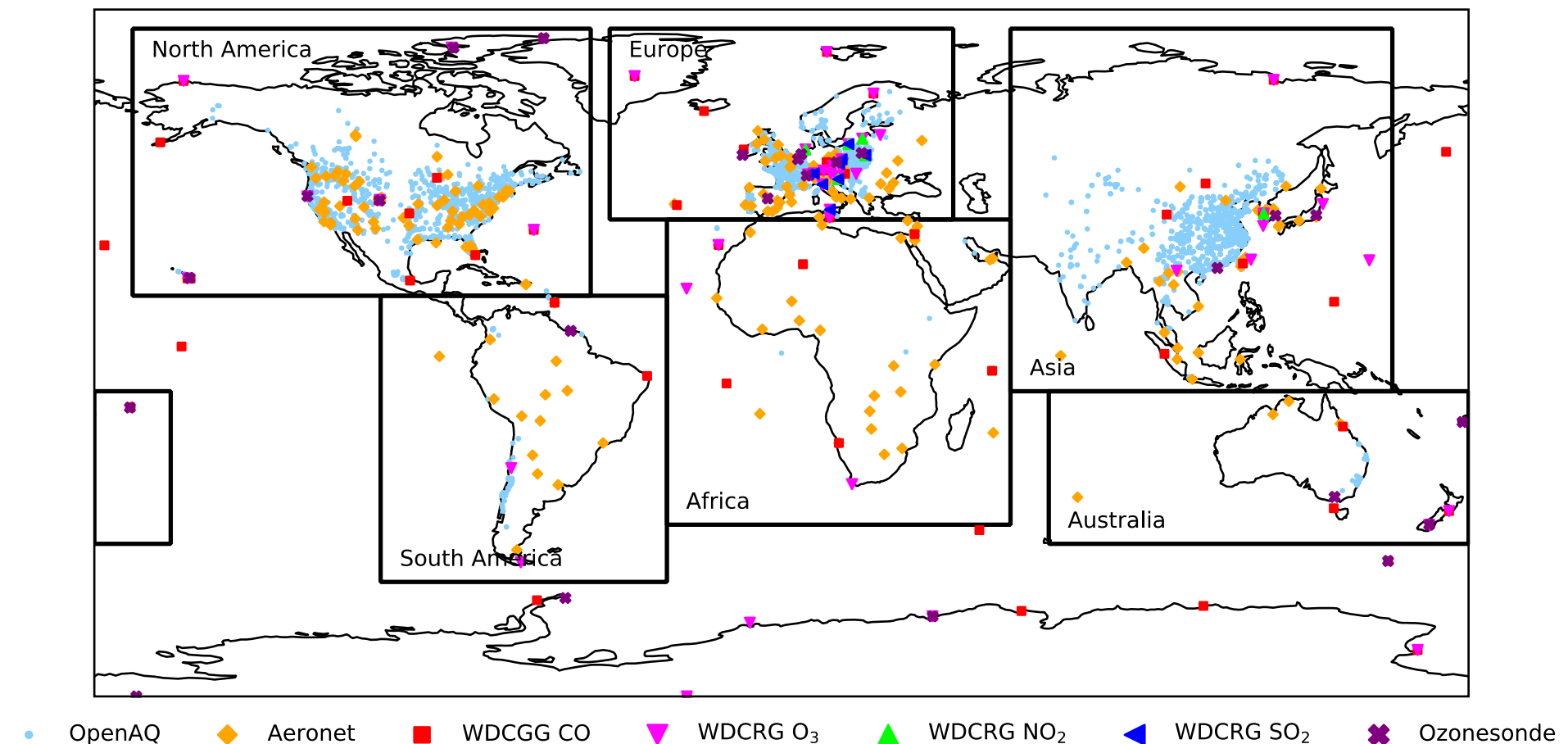
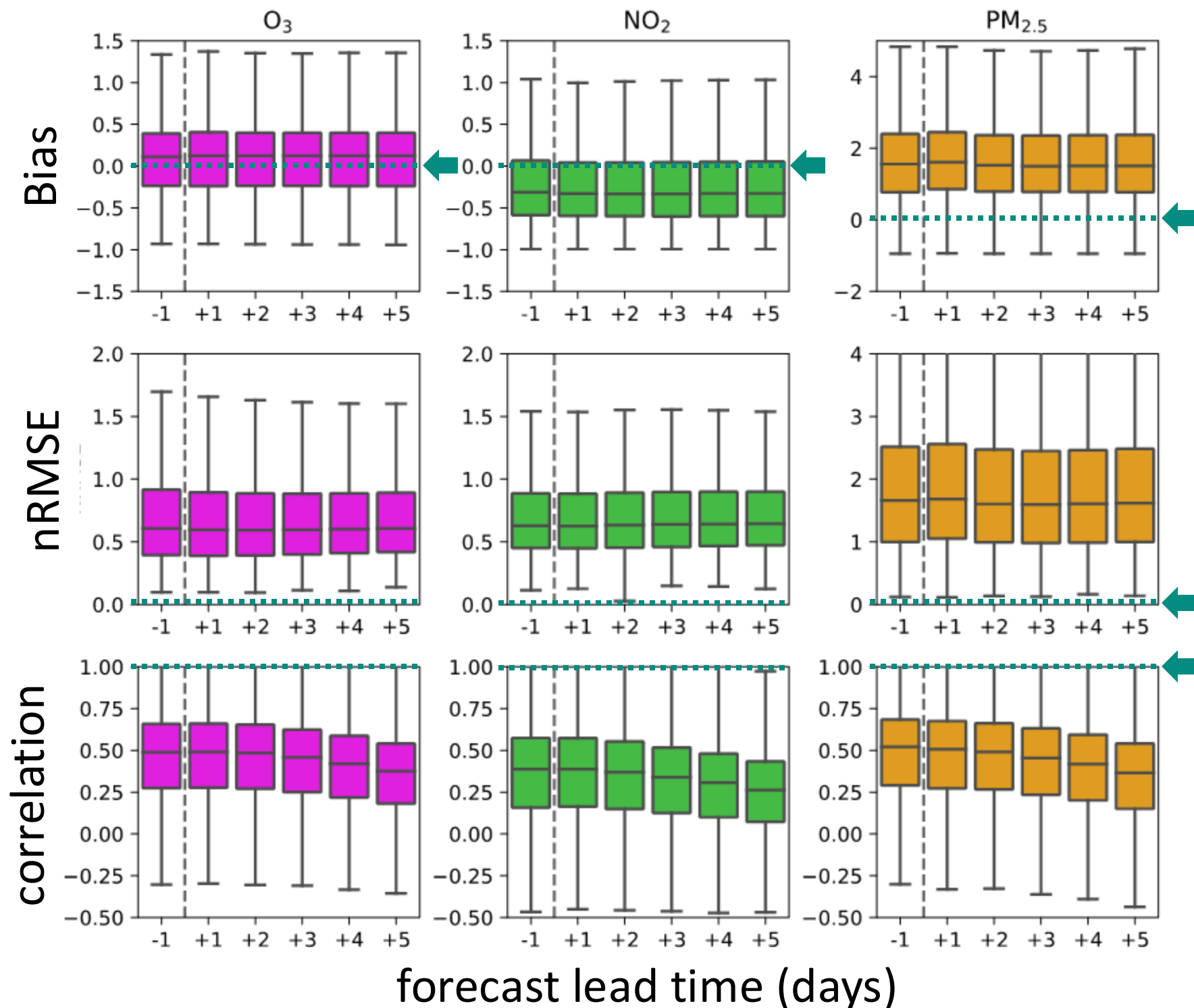
Presentation for GEEDSVS Nairobi Community
carl.a.malings@nasa.gov

NASA Global Earth Observing System (GEOS)

	GEOS FP	GEOS-CF	MERRA-2
Type	Analysis + Forecast	Replay + Forecast	Reanalysis
Domain	Global	Global	Global
Spatial Resolution	Simulation: ~12 km Output: ~25 km (0.25°x0.312°)	~25 km (0.25°x0.312°)	~50km (0.5°x0.625°)
Temporal Resolution	2-D data: Hourly 3-D data: Every 3 h	15 min, Hourly	Hourly, Daily, Monthly
Vertical Levels	72 (near surface-0.1 hPa)	72 (near surface-0.1 hPa)	72 (near surface-0.1 hPa)
Output available	Analysis: 2014 – Present Forecast: ~20 days	Replay: 2018 – Present Forecast: 2019 – Present (aqc collection) ~14 days (all collections)	1980-Present
Initialization	Daily 10-day forecast at 00Z Daily 5-day forecast at 12Z	Daily 5-day forecast at 12Z	~1-2 months behind real time
Data Assimilation	Yes	No	Yes
File Specification Doc	https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi1203.pdf *	https://gmao.gsfc.nasa.gov/pubs/docs/Knowland1204.pdf *	https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf *

Source: Gupta, P.; Follette-Cook, M.; Parrington, M.; Stewart, C. (2021). *Introduction and Access to Global Air Quality Forecasting Data and Tools*. NASA ARSET.
<https://appliedsciences.nasa.gov/join-mission/training/english/arset-introduction-and-access-global-air-quality-forecasting-data-and>

GEOS-CF Performance



PM_{2.5} is biased high
 NO₂ is biased low
 SO₂ is biased very high (not shown)
 Little variation in the skill scores out to 5 days

Slide Credit: K. Emma Knowland, NASA GMAO

Source: Keller, C., et al. (2021) "Description of the NASA GEOS Composition Forecast Modeling System GEOS-CF v1.0". *Journal of Advances in Modeling Earth Systems*, 13:4.
<https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020MS002413>

Why quantify uncertainty in air quality forecasting?

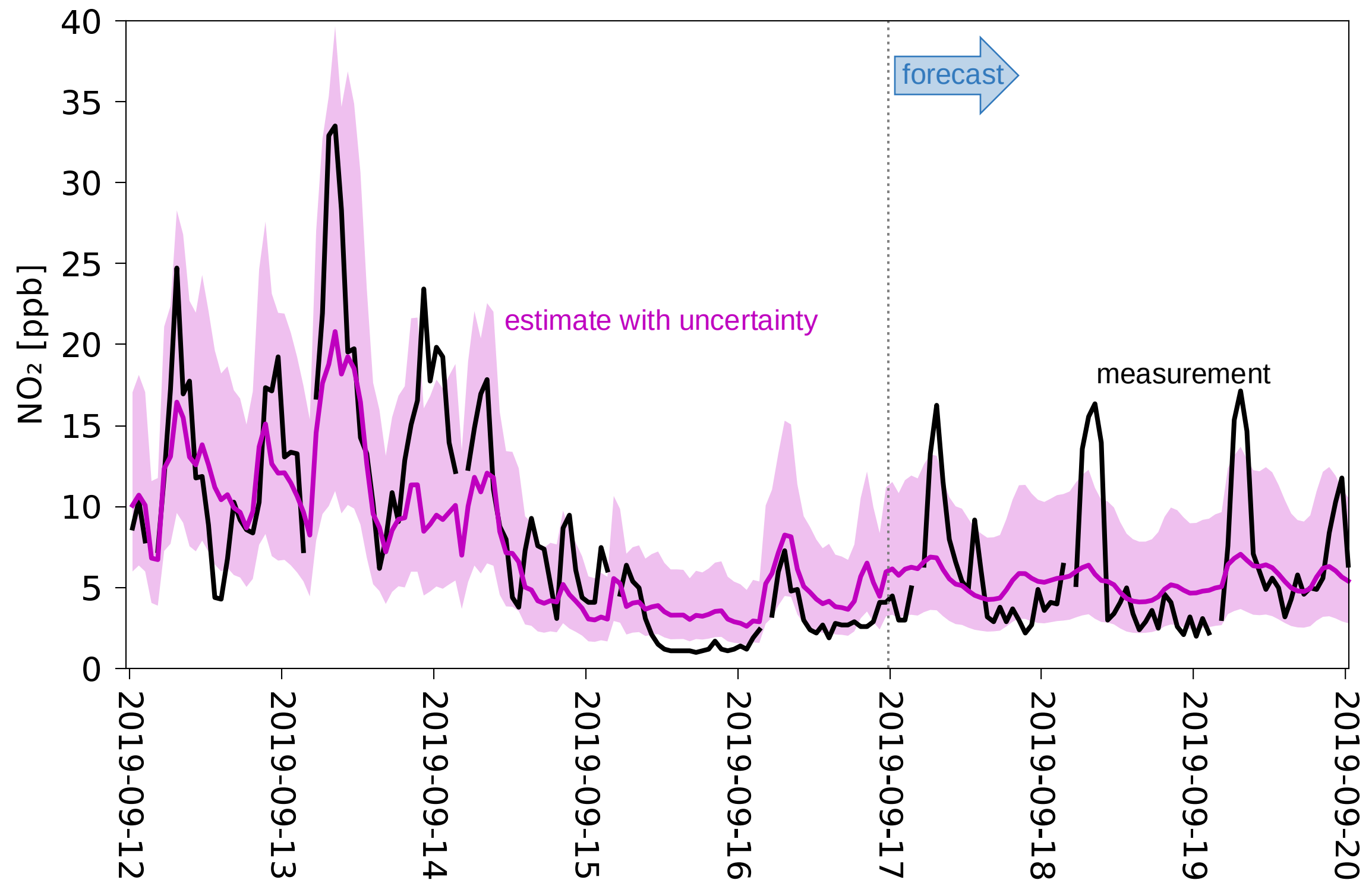
Provide a prior estimate of the relative confidence in a forecast

Convey probabilities of specific events, e.g., exceedance of standards

Identify a range of likely outcomes

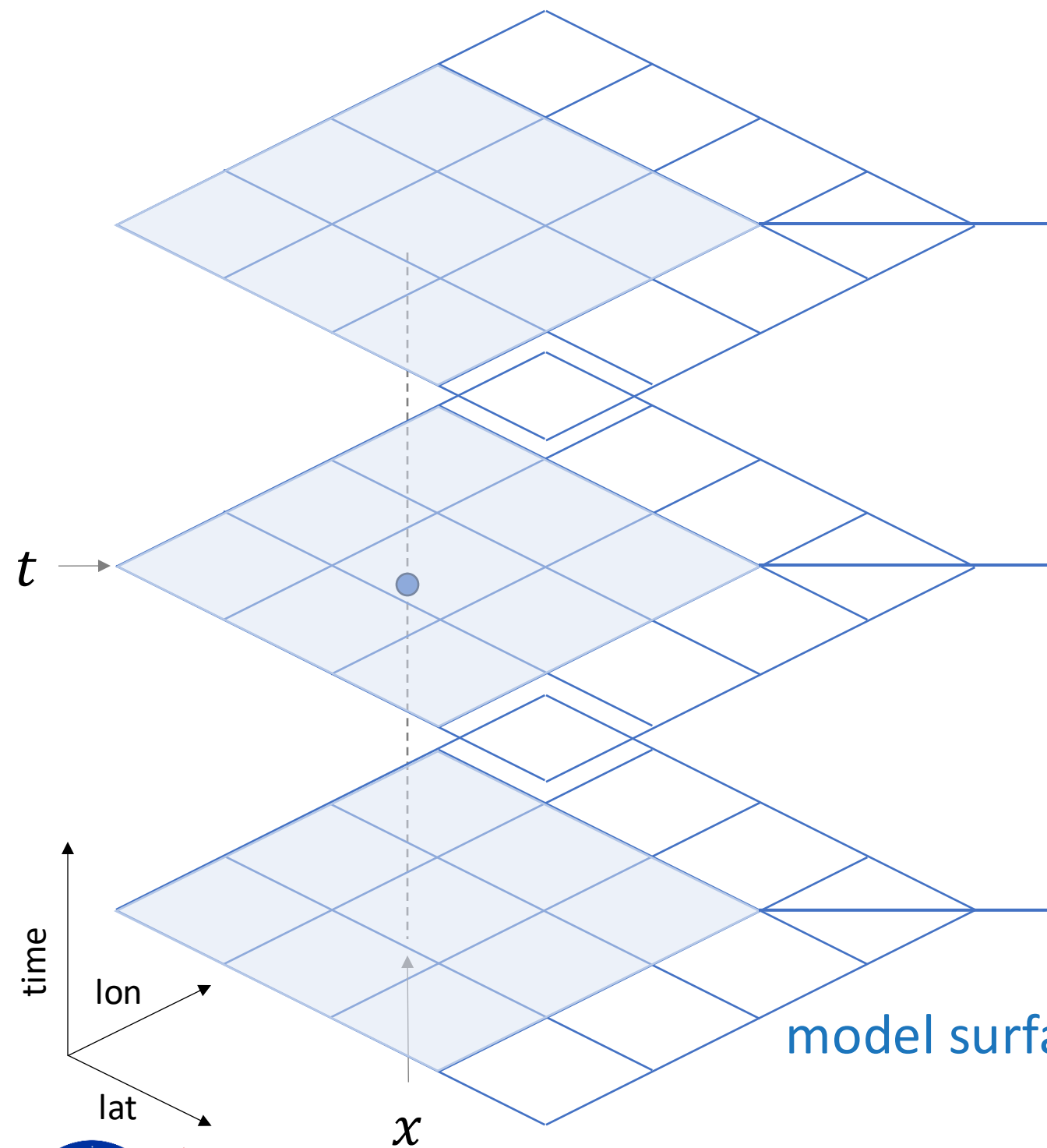
Quantify the impacts of different data sources in reducing uncertainties

Identify the potential to reduce uncertainties through additional data collection



Phase 1: Model Only

Phase	Estimate	Uncertainty
1	forecast model (GEOS-CF)	cell-to-cell variability of model



$$F_1(x, t) = M(x, t)$$

data fusion
uncertainty
(variance) at
phase 1

$$V_1(x, t) = V_{F1}(x, t, \tau)$$

uncertainty due to forecasting by τ
ahead; evaluate using past forecasts

$$+ V_M(x, t)$$

uncertainty due to model internal
variability

$$+ V_{B1}(x, t)$$

uncertainty due to model bias

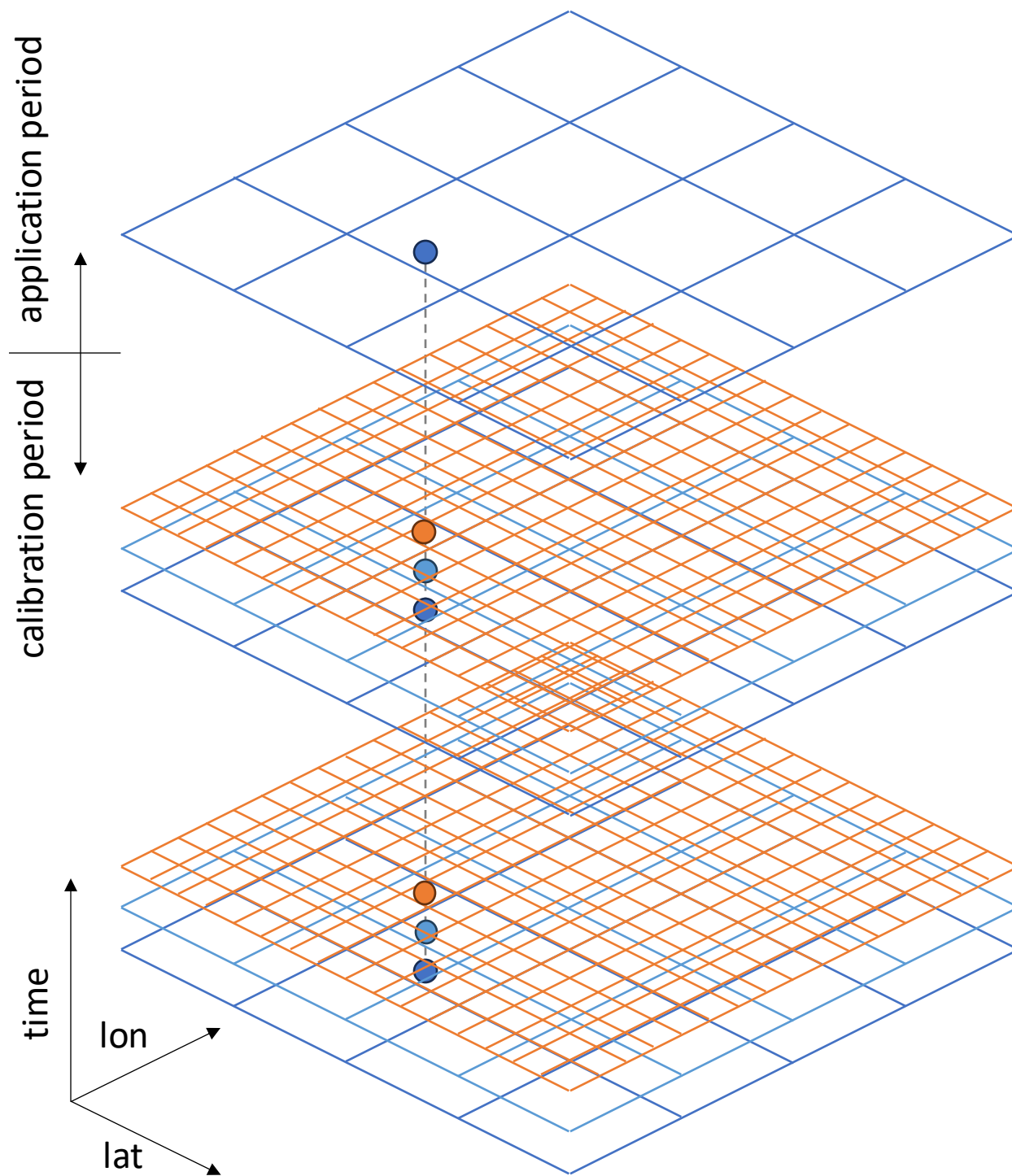
$$+ V_{R1}(x, t)$$

uncertainty due to spatial
representativity (model scale)

$$V_M(x, t) \approx \mathbb{E}_{x' \in X_n(x), t' \in T_n(t)} \left[(M(x', t') - M(x, t))^2 \right]$$

model surface concentration $M(x, t)$

Phase 2: Model & Satellite



Phase	Estimate	Uncertainty
1	forecast model (GEOS-CF)	cell-to-cell variability of model
2	satellite (TROPOMI) informs sub-model-grid variability	satellite-to-model and surface-to-column ratios vary over time

$V_2(x, t) = V_{F2}(x, t, \tau)$ ← uncertainty due to forecasting by τ ahead

$+V_M(x, t)$ ← uncertainty due to model internal variability

$+V_D(x, t)$ ← uncertainty in satellite-to-model differences

$+2V_{MD}(x, t)$ ← co-variance of satellite-to-model differences with model outputs

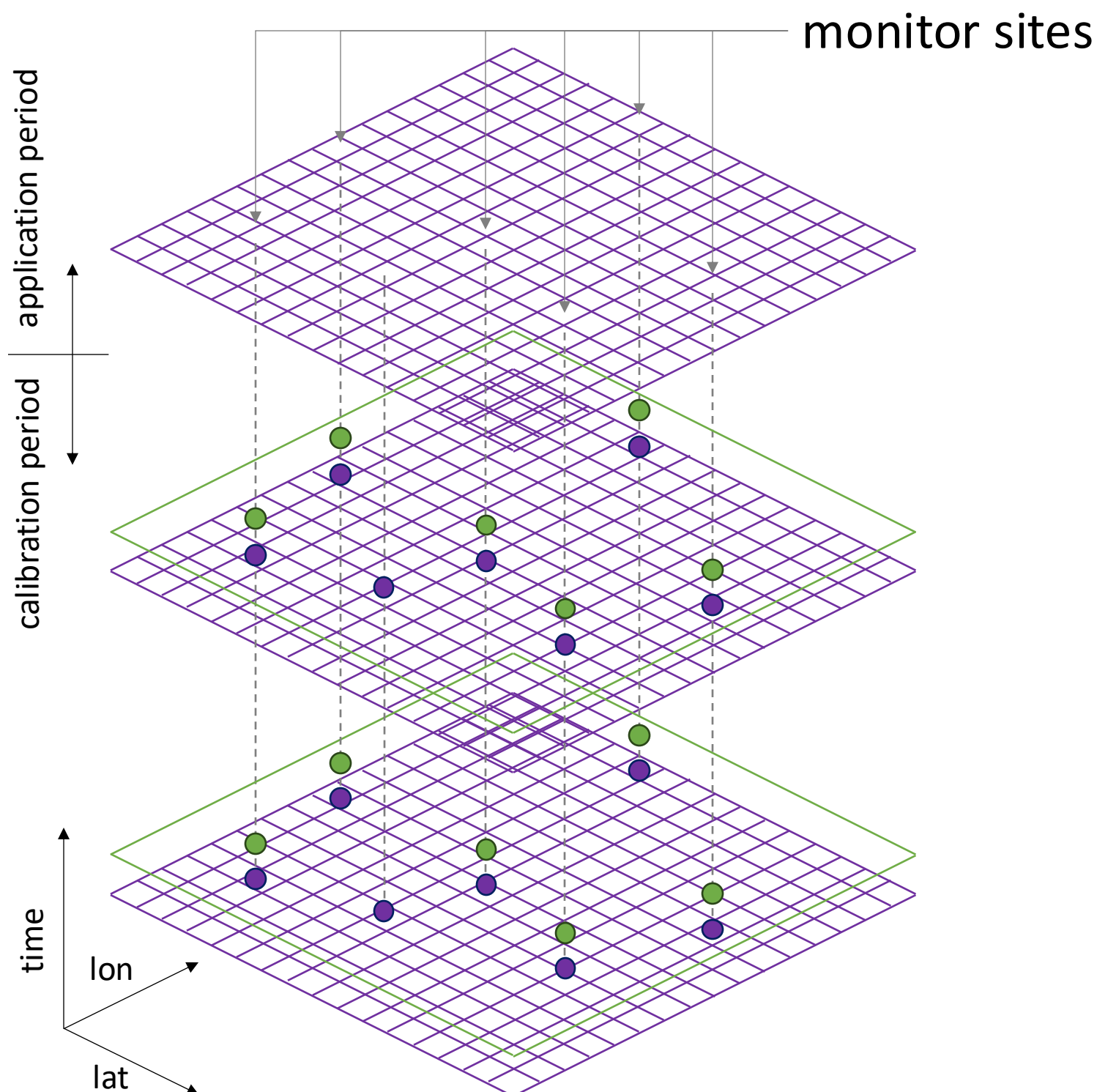
$+V_{B2}(x, t)$ ← uncertainty due to model & satellite bias

$+V_{R2}(x, t)$ ← uncertainty due to spatial representativity (satellite scale)

$$V_D(x, t) \approx \mathbb{V}_{t' \in T_{c, overpass}(t)} [(S_{col}(x, t') - M_{col}(x, t')) \phi(x, t') \psi(x, t, t')]$$

$$V_{MD}(x, t) \approx \mathbb{E}_{x' \in X_n(x), t' \in T_n(t)} [(M(x', t') - M(x, t))(D(x', t') - D(x, t))]$$

Phase 3: Model & Satellite & Ground

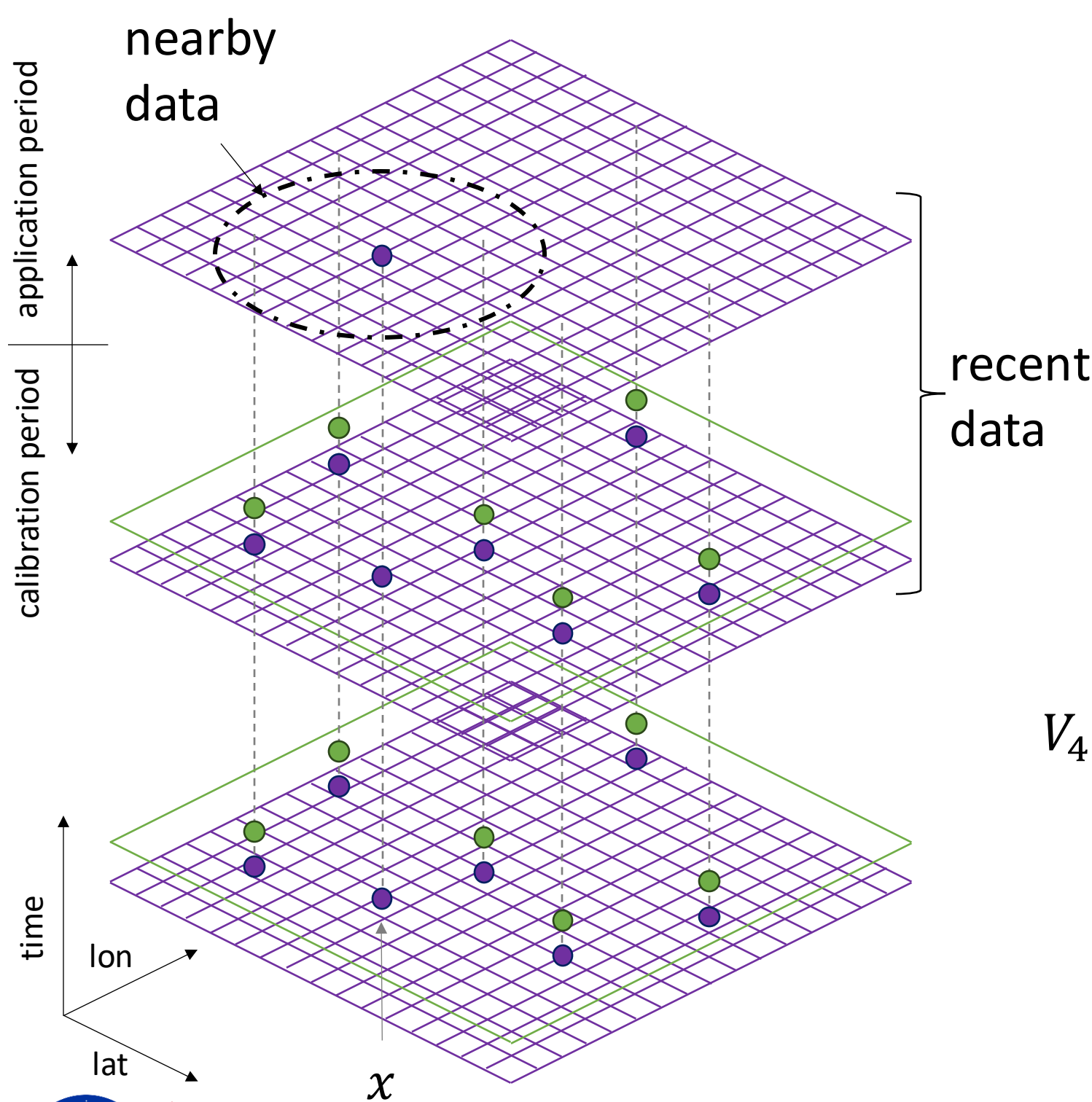


Phase	Estimate	Uncertainty
1	forecast model (GEOS-CF)	cell-to-cell variability of model
2	satellite (TROPOMI) informs sub-model-grid variability	satellite-to-model and surface-to-column ratios vary over time
3	phase 2 corrected to match surface monitor data	uncertain regression parameters between phase 2 output and surface monitor data

$$V_3(x, t) = V_{F3}(x, t, \tau) \leftarrow \text{uncertainty due to forecasting by } \tau \text{ ahead}$$
$$+ \theta_1^2 [V_M(x, t) + V_D(x, t) + 2V_{MD}(x, t)] \leftarrow \text{rescaled from phase 2}$$
$$+ \text{var}[\theta_1] F_2(x, t)^2$$
$$+ 2\text{cov}[\theta_0, \theta_1] F_2(x, t)$$
$$+ \text{var}[\theta_0]$$
$$+ \sigma_{\text{residual}}^2$$

variance and co-variance of regression parameters as well as regression residual are known

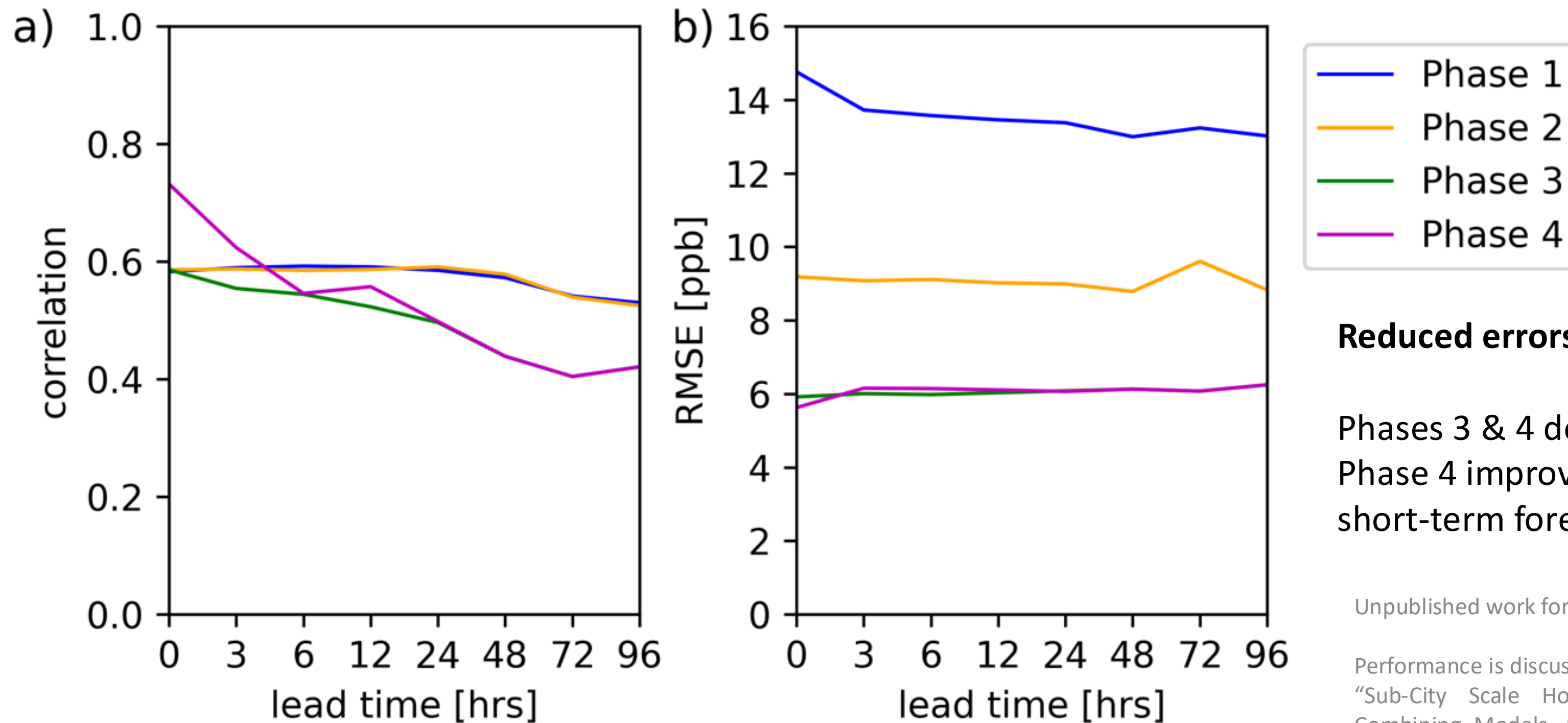
Phase 4: Residual Kriging



Phase	Estimate	Uncertainty
1	forecast model (GEOS-CF)	cell-to-cell variability of model
2	satellite (TROPOMI) informs sub-model-grid variability	satellite-to-model and surface-to-column ratios vary over time
3	phase 2 corrected to match surface monitor data	uncertain regression parameters between phase 2 output and surface monitor data
4	update phase 3 based on recent surface monitor data	uncertainty reduction via updating with nearby & recent data (kriging)

$$V_4(x, t) = V_3(x, t) - \sum_{x' \in X_n(x), t' \in T_n(t)} K(x, x', t, t') \text{cov}[G(x', t'), F_3(x', t')]$$

Data Fusion Results



Reduced errors as phase increases

Phases 3 & 4 degrade correlation,
Phase 4 improves correlations for
short-term forecasts

Unpublished work for a case study example

Performance is discussed in detail in Malings et al. (2021),
“Sub-City Scale Hourly Air Quality Forecasting by
Combining Models, Satellite Observations, and Ground
Measurements” *Earth & Space Science*.

[DOI: 10.1029/2021EA001743](https://doi.org/10.1029/2021EA001743)

Uncertainty Quantification Results

Case Study Details

San Francisco

September 2019

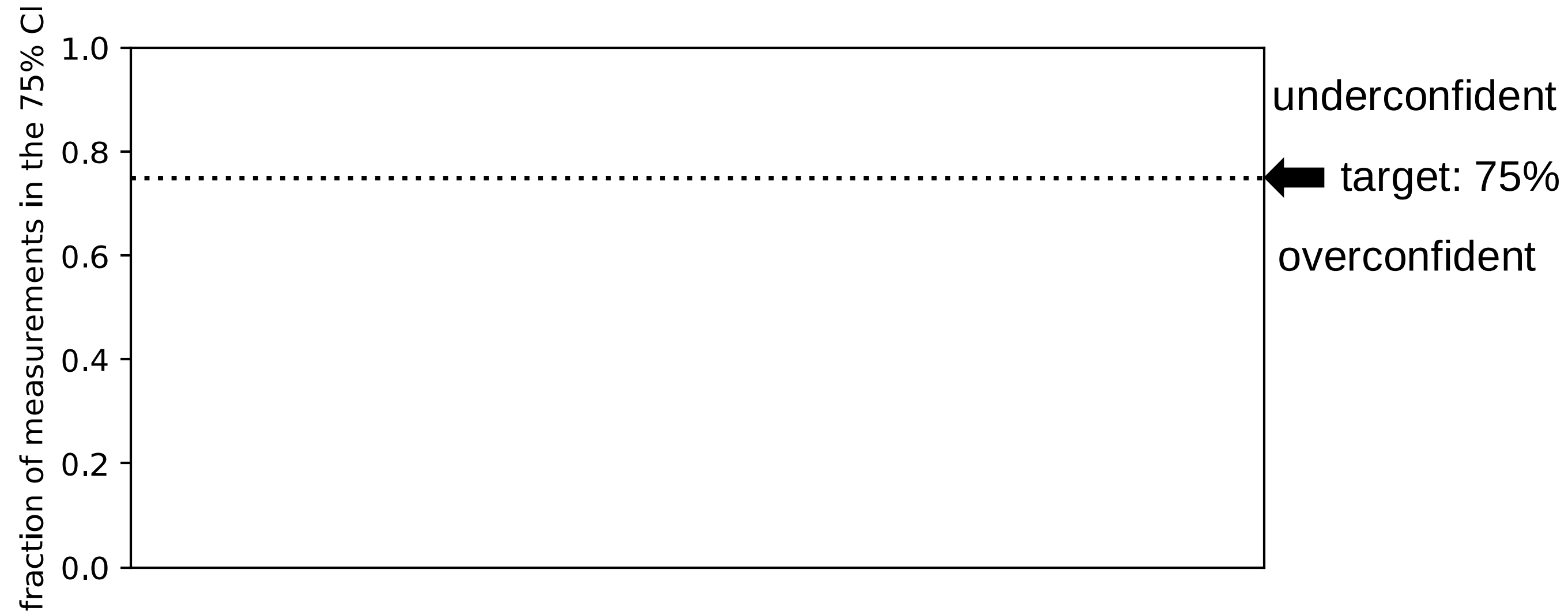
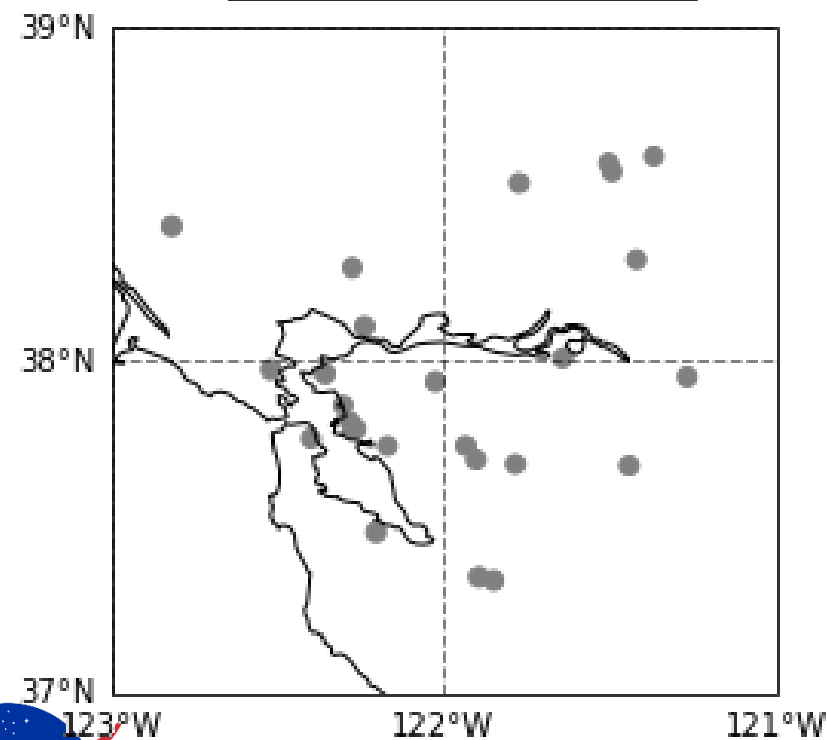
Surface NO₂

Lognormal distribution

Cross-validation test

25 ground monitors

Ground Sites



Uncertainty Quantification Results

Case Study Details

San Francisco

September 2019

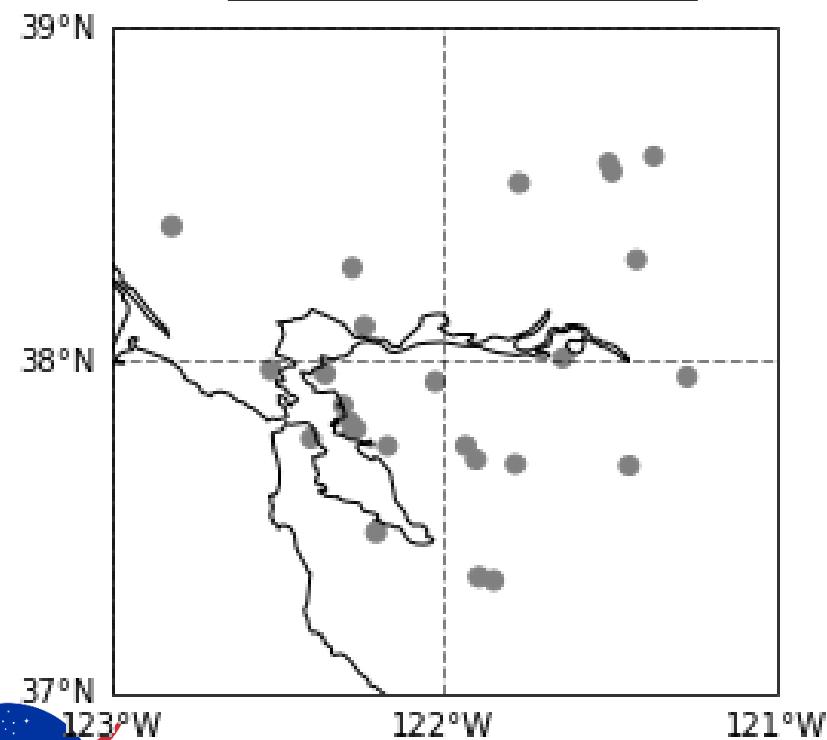
Surface NO₂

Lognormal distribution

Cross-validation test

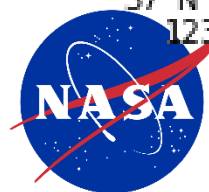
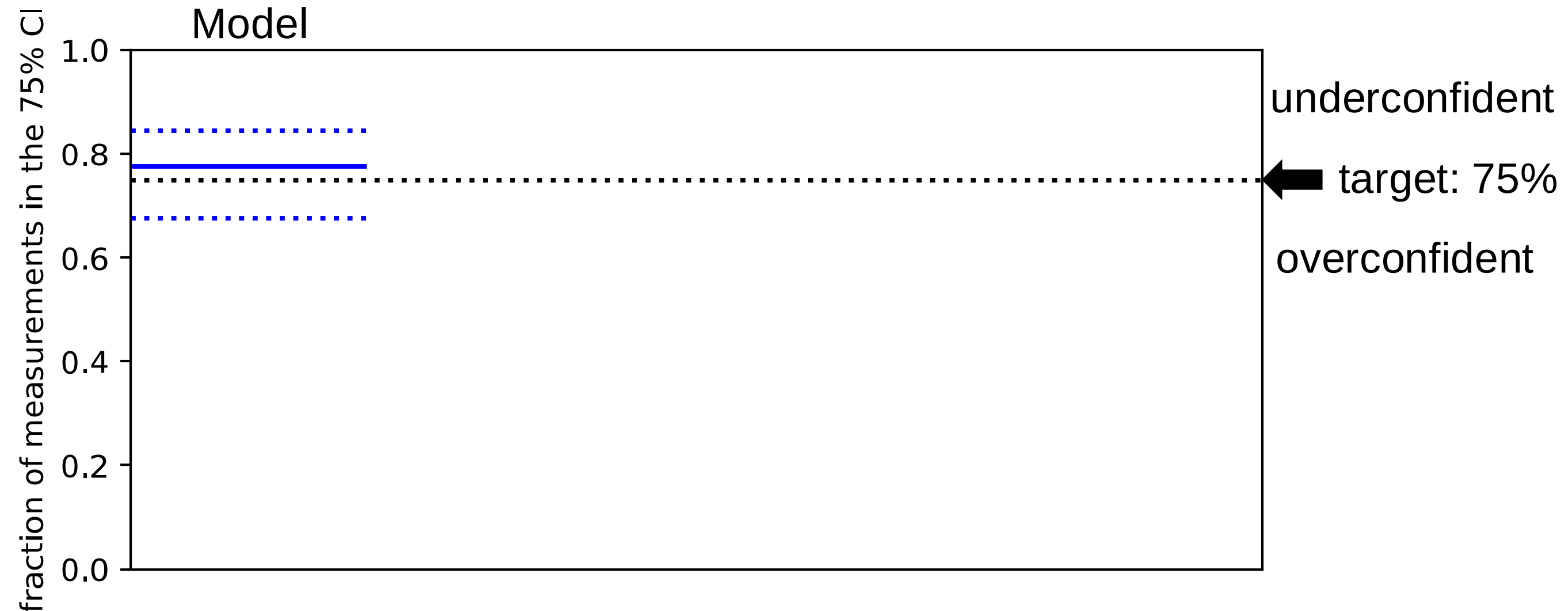
25 ground monitors

Ground Sites



Phase 1

Model



Partner



Presentation for GEEDEVs Nairobi Community
carl.a.malings@nasa.gov

17 Oct. 2024

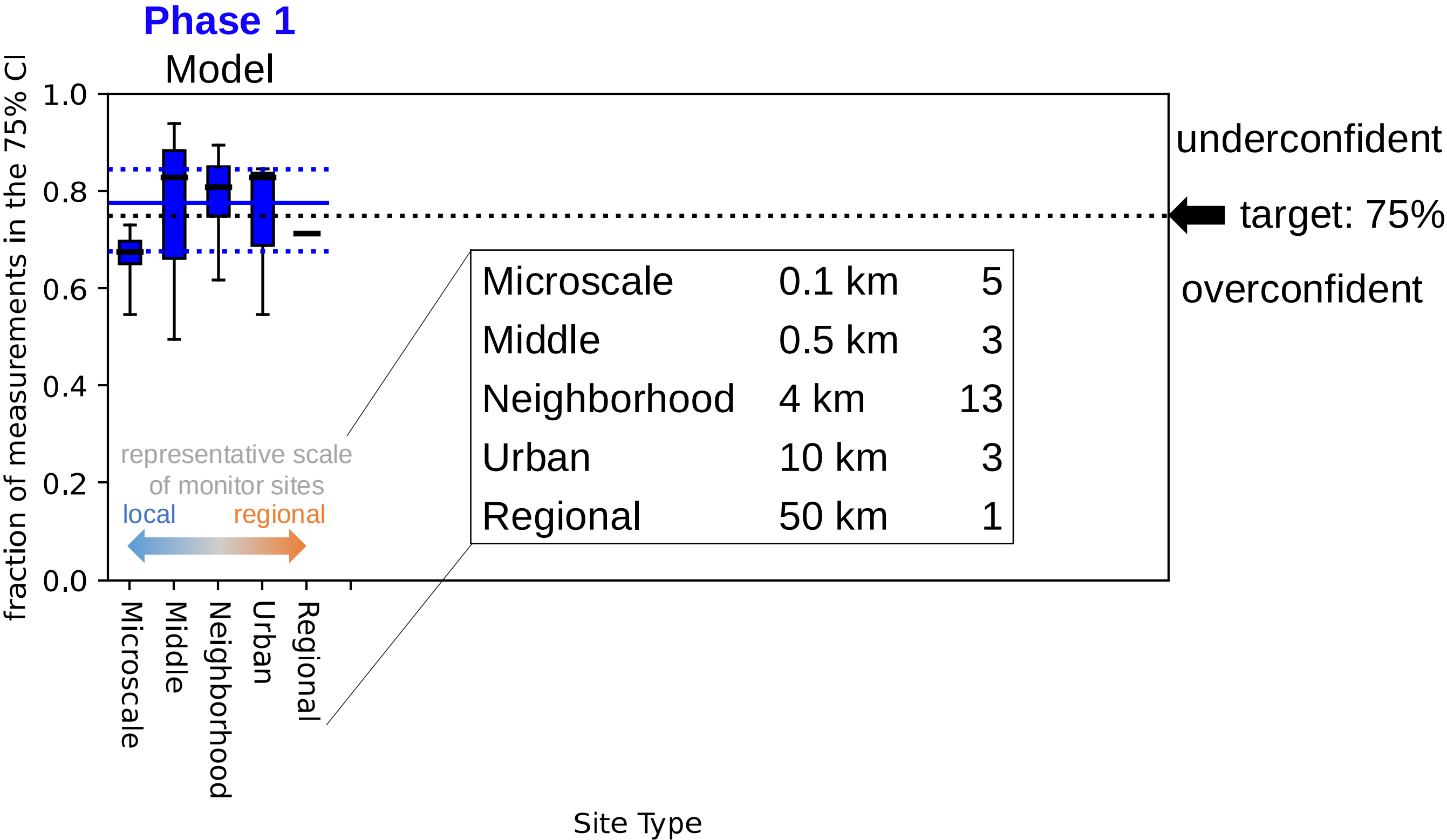
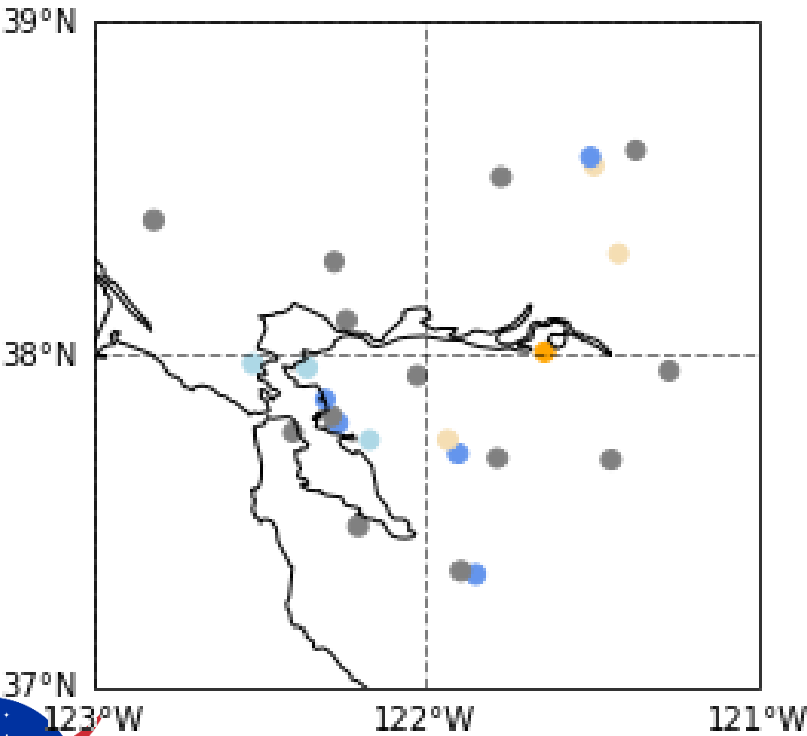
63

Uncertainty Quantification Results

Case Study Details

San Francisco
September 2019
Surface NO₂
Lognormal distribution
Cross-validation test
25 ground monitors

Ground Sites



Uncertainty Quantification Results

Case Study Details

San Francisco

September 2019

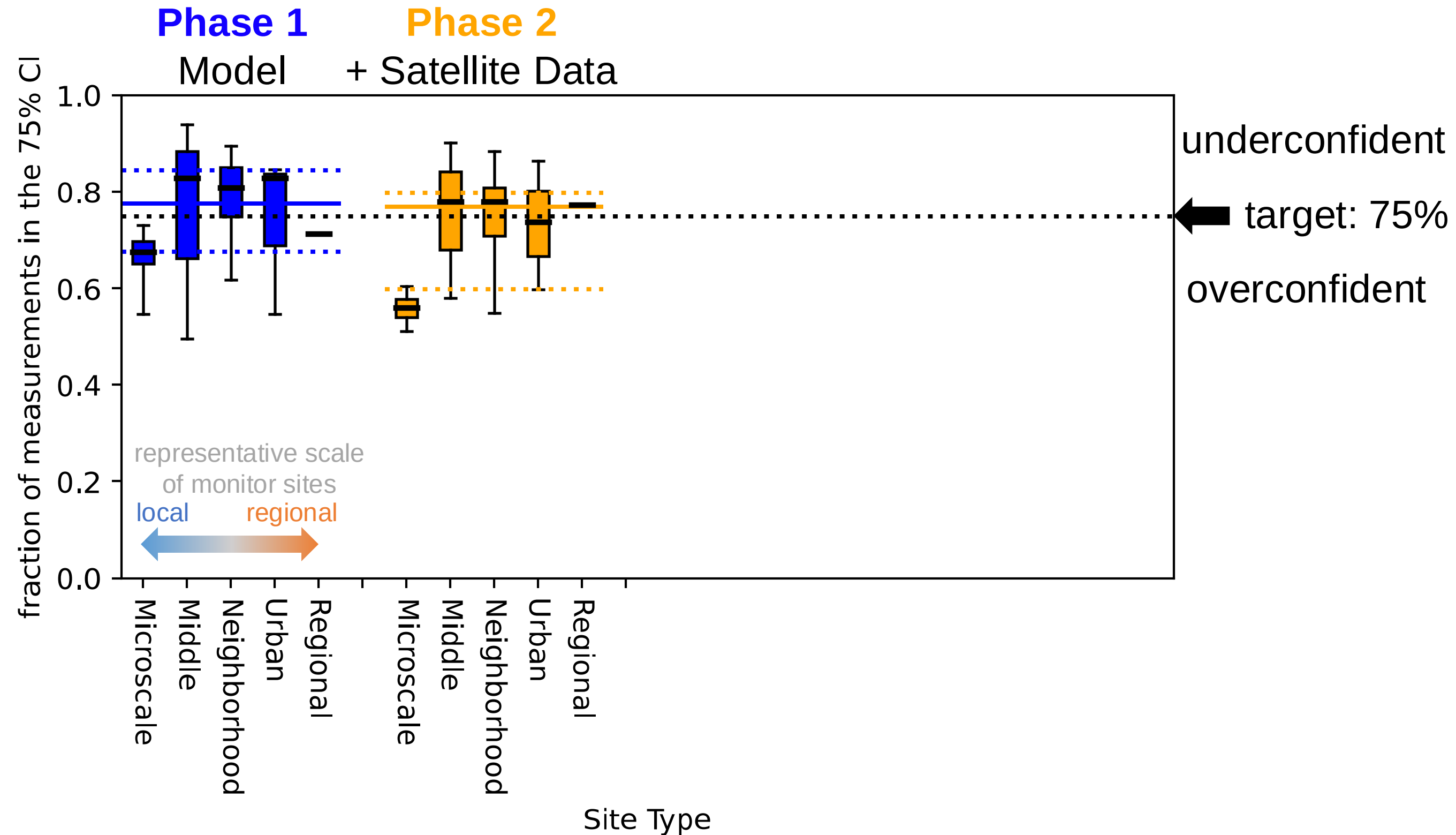
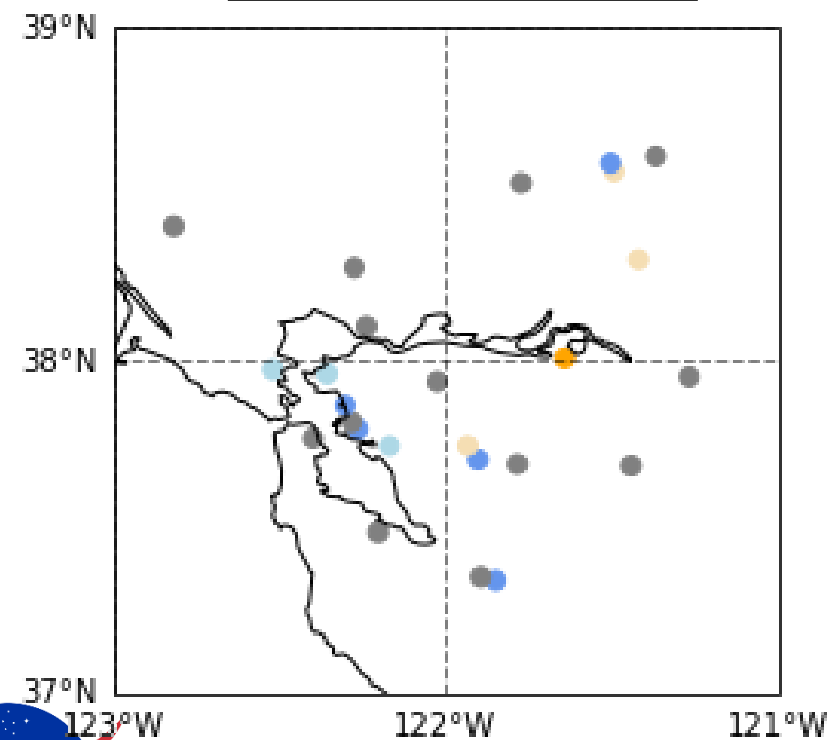
Surface NO₂

Lognormal distribution

Cross-validation test

25 ground monitors

Ground Sites



Uncertainty Quantification Results

Case Study Details

San Francisco

September 2019

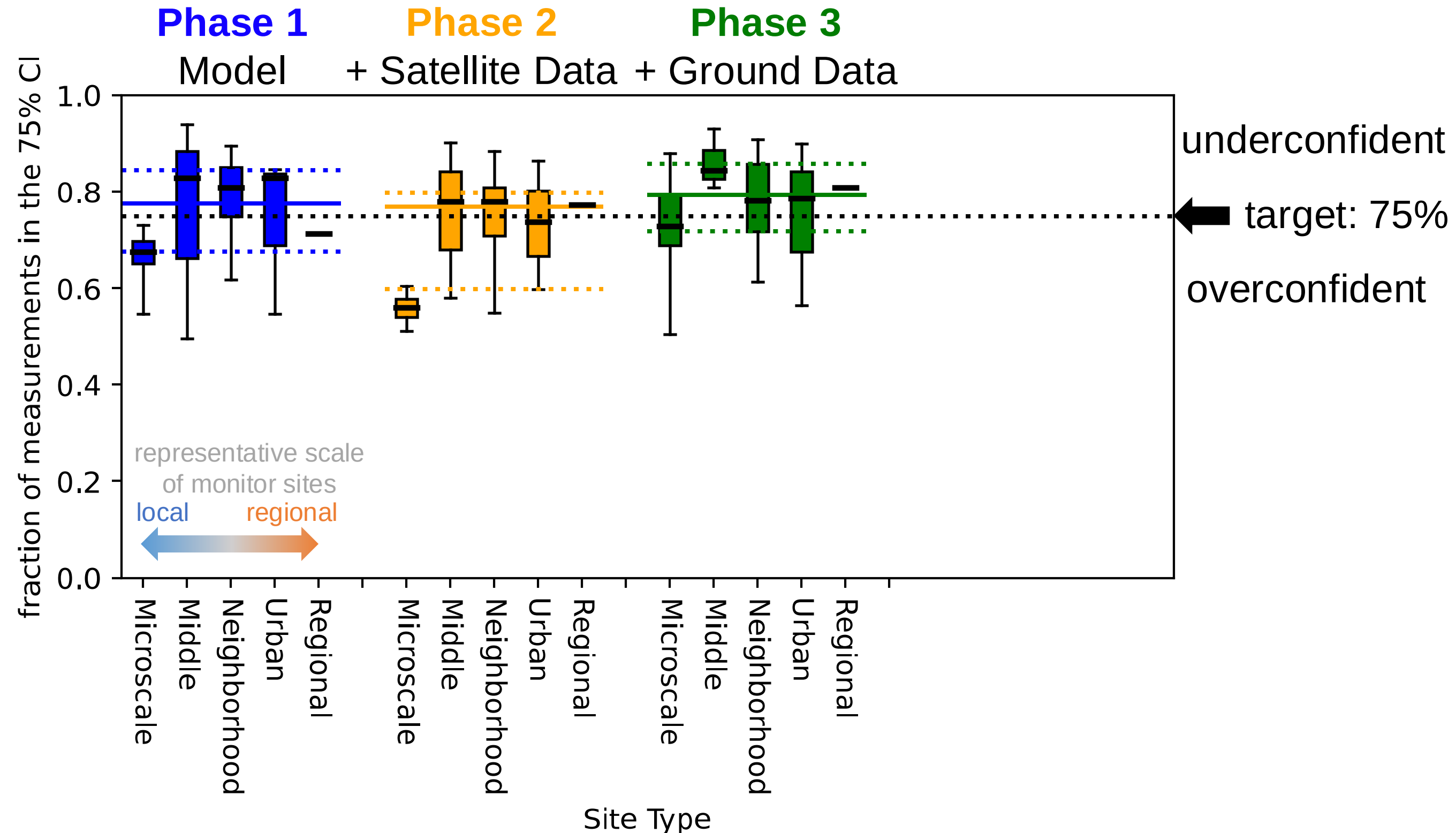
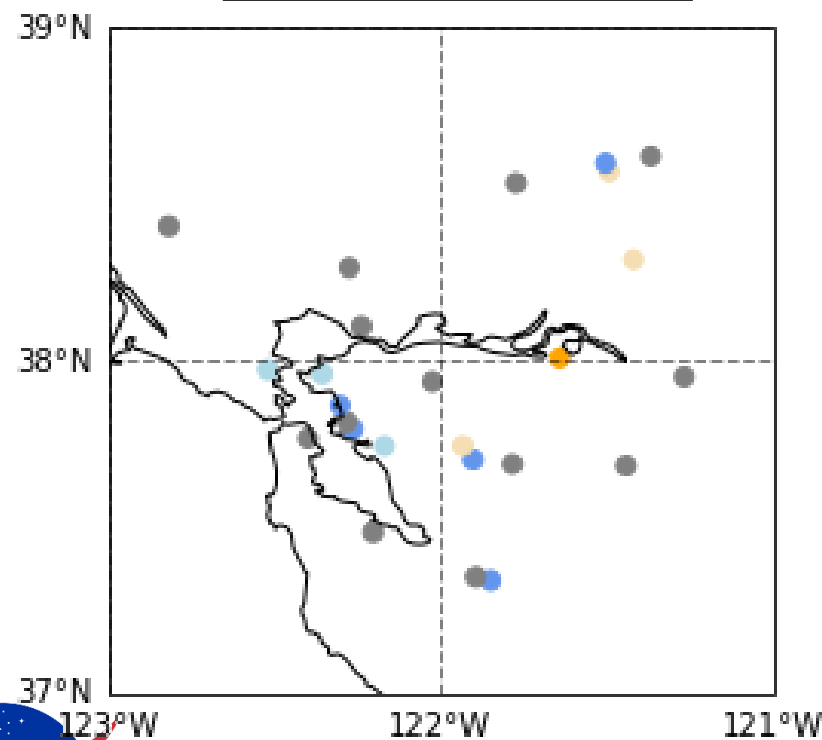
Surface NO₂

Lognormal distribution

Cross-validation test

25 ground monitors

Ground Sites



Uncertainty Quantification Results

Case Study Details

San Francisco

September 2019

Surface NO₂

Lognormal distribution

Cross-validation test

25 ground monitors

Ground Sites

