



Google and NASA Air Quality Partnership – A Collaboration Using GEOS-CF Data and Google Earth Engine

Callum Wayman^{1,3}, K. Emma Knowland^{2,3}, Christoph Keller^{2,3}, Joseph Ardizzone^{1,3}, Steven Pawson³

¹Science Systems and Applications, Inc.

²Morgan State University

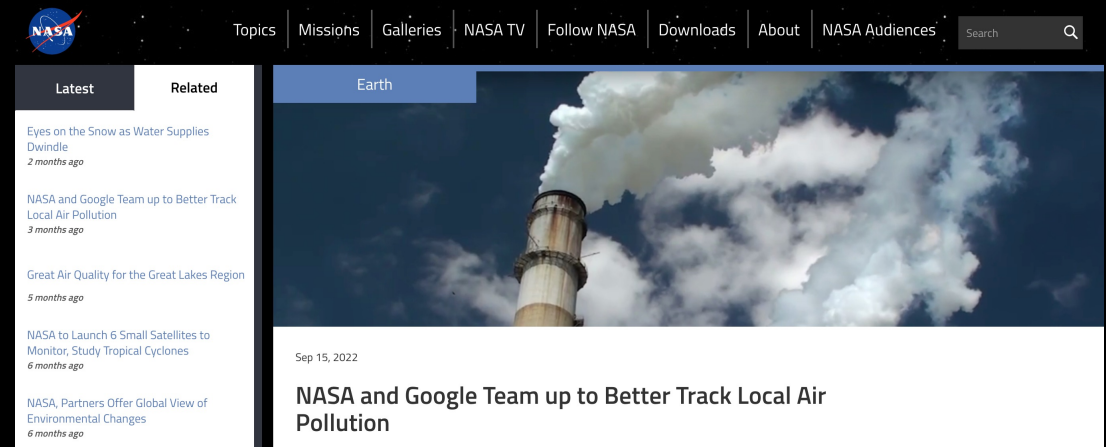
³Global Modeling and Assimilation Office, NASA Goddard Space Flight Center

Google NASA Air Quality Partnership

NASA and Google have expanded their partnership to create data and tools that help with pollution mitigation and decision making on a local government scale

The goal is to use the technologies available at NASA and Google to create city-scale data estimates and forecasts of air pollutants such as NO_2 derived from the GEOS Composition Forecast (GEOS-CF) model

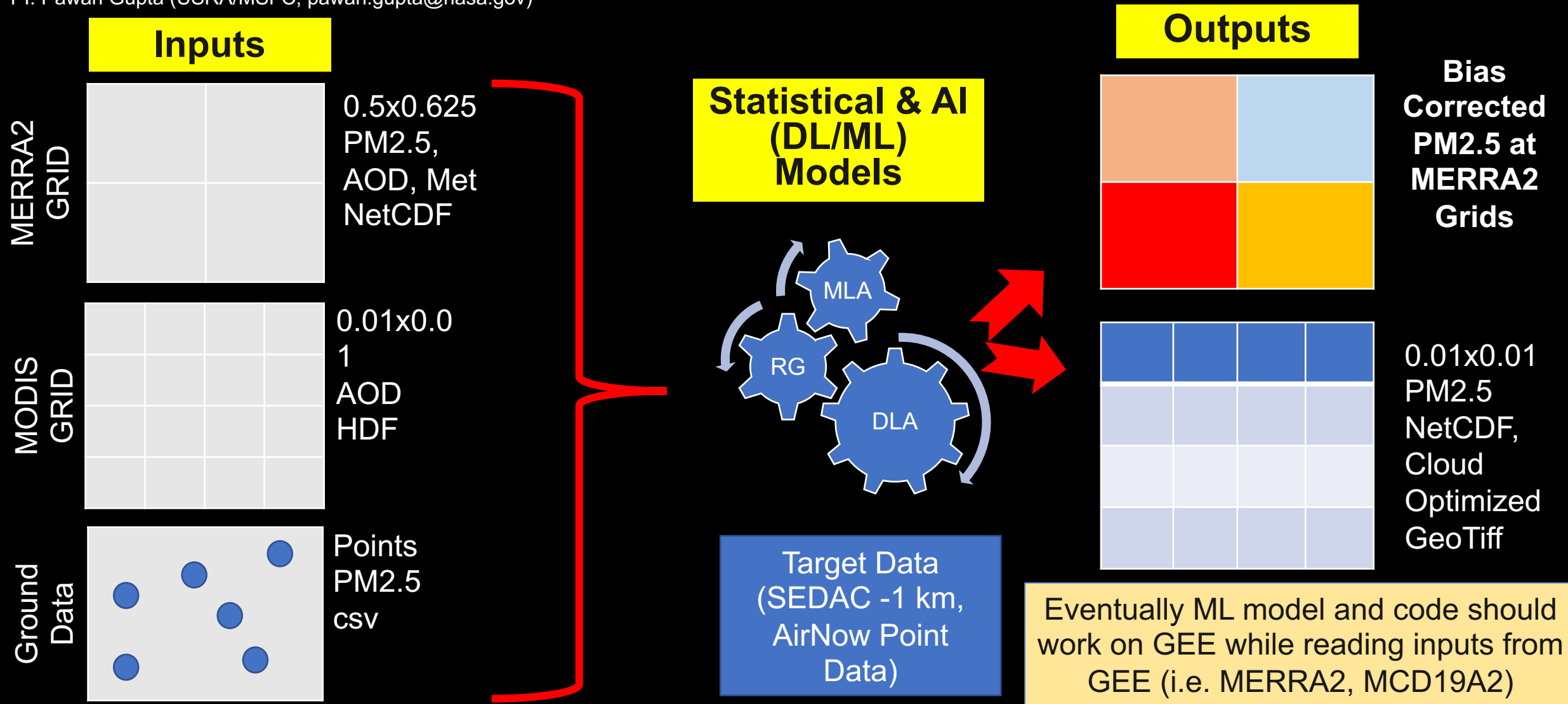
Efforts are also being led by Pawan Gupta to create a downscaled MERRA-2 PM_{2.5} product



Featured on the NASA homepage!

Bias Correction and Downscaling – A Machine Learning Approach

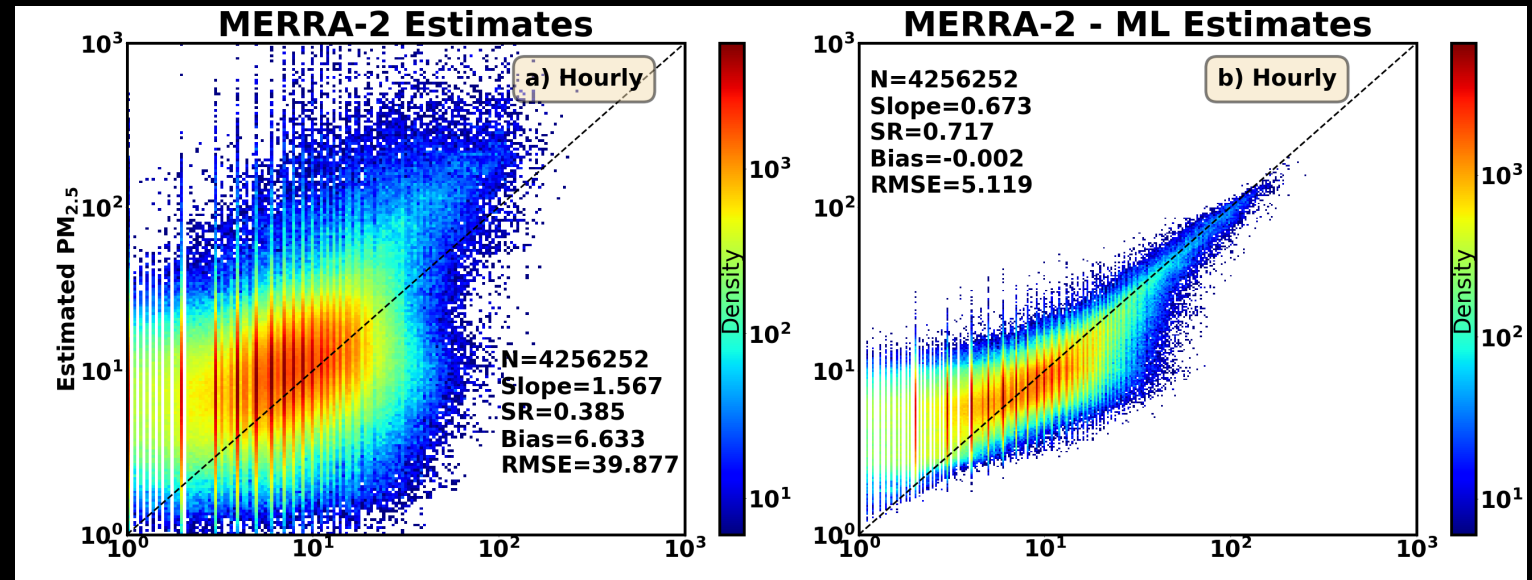
PI: Pawan Gupta (USRA/MSFC, pawan.gupta@nasa.gov)



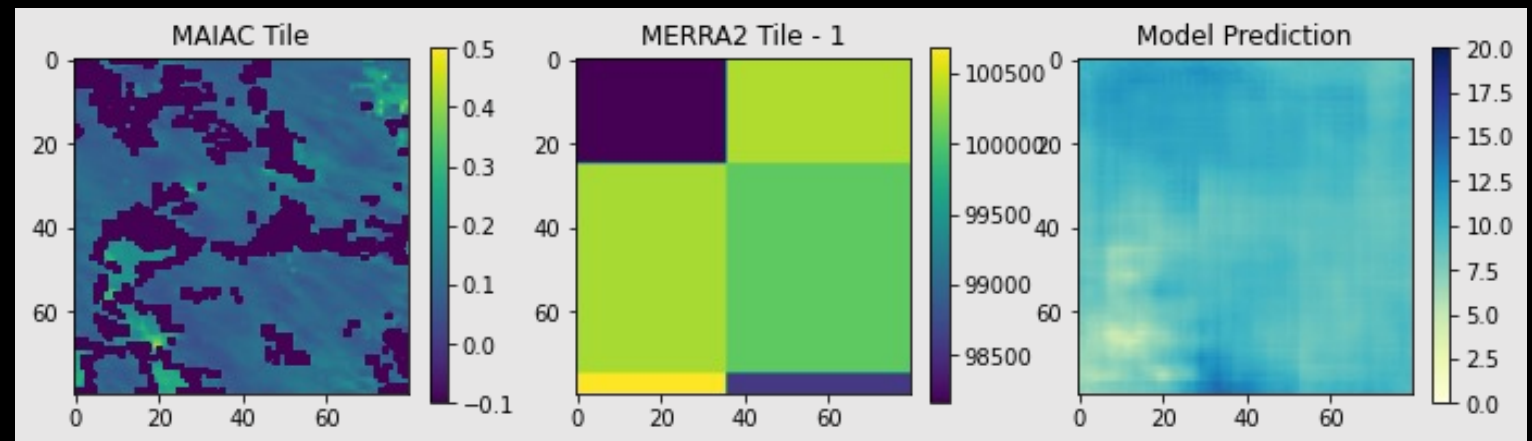
Preliminary Results

- Machine Learning Approach reduces the bias in PM_{2.5} w.r.t. EPA's measurements (Sayeed et al., 2022, under revision) - <https://www.essoar.org/doi/abs/10.1002/essoar.10511086.1>

- The coarse resolution MERRA2 PM_{2.5} (middle panel) is downscaled to a higher resolution (last panel) using satellite aerosol data (first panel) – preliminary results

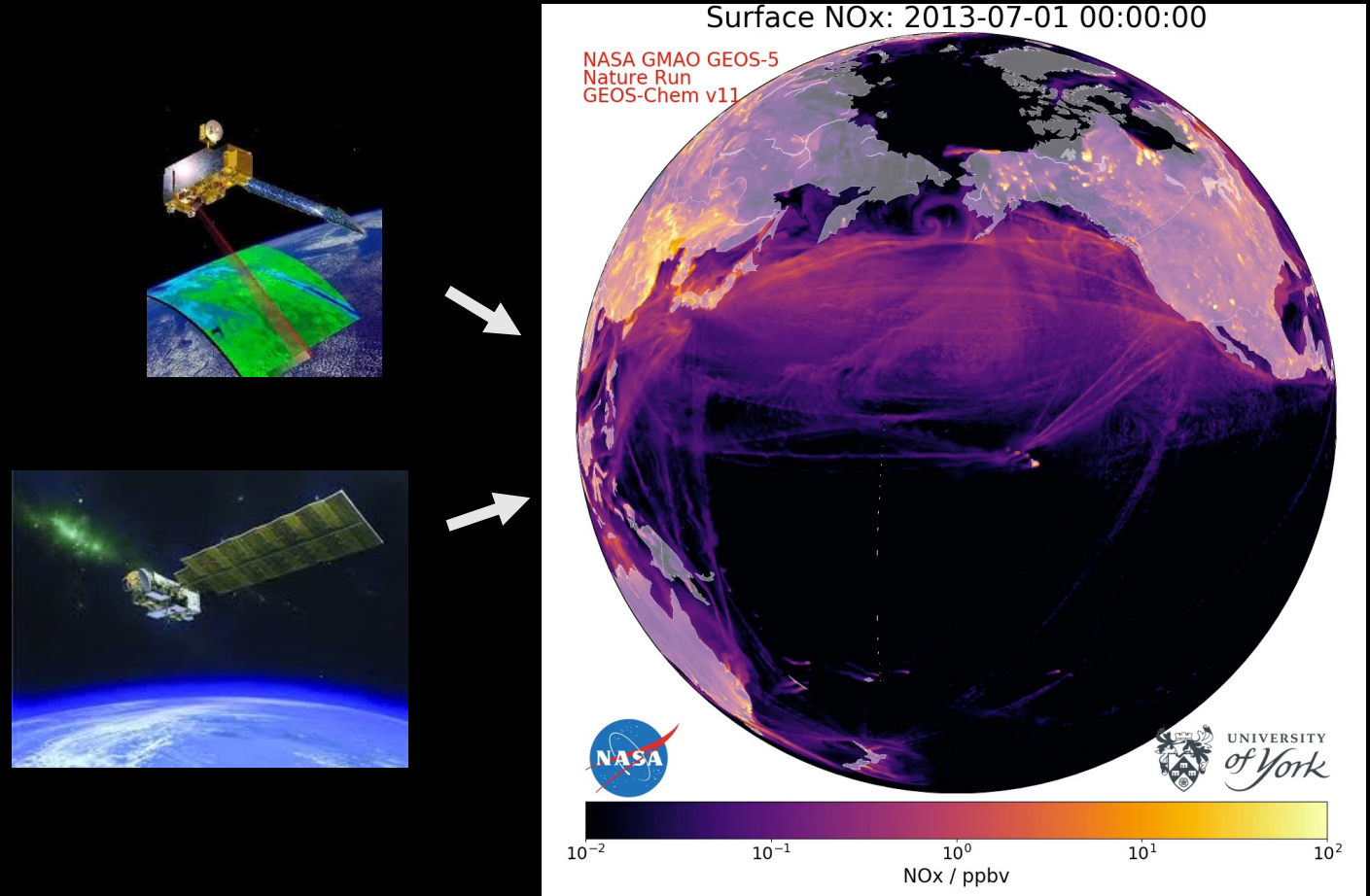


Lead: Alqmah Sayeed (USRA/MSFC)

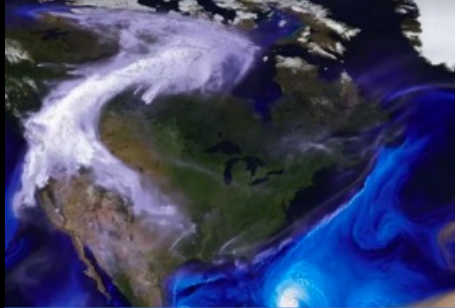


Lead: Michael Von Phole (USRA/Ames)

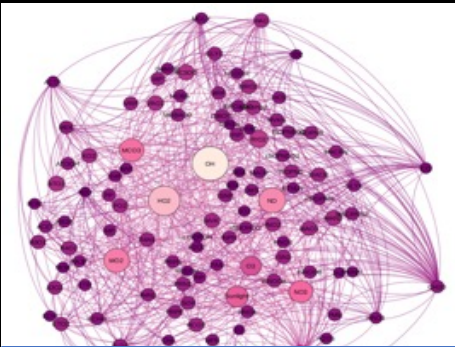
We can now simulate atmospheric composition with unprecedented level of detail



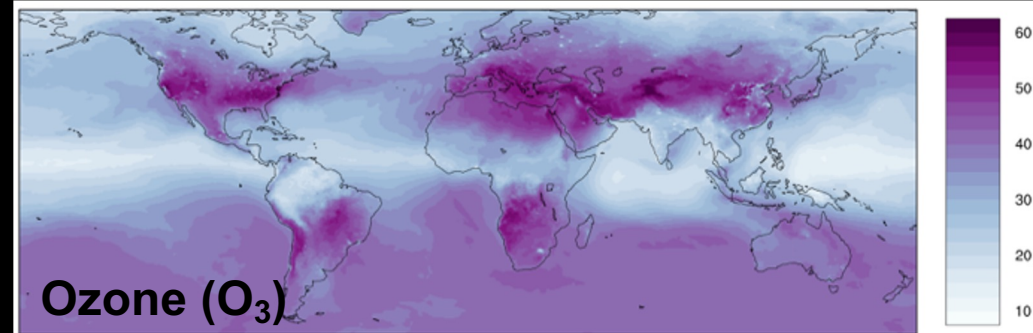
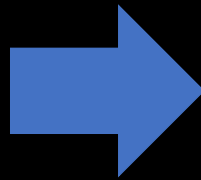
NASA GEOS Composition Forecast



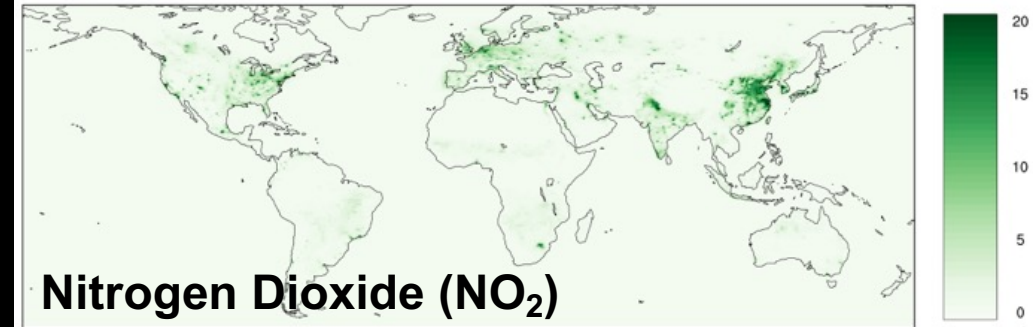
GEOS Weather



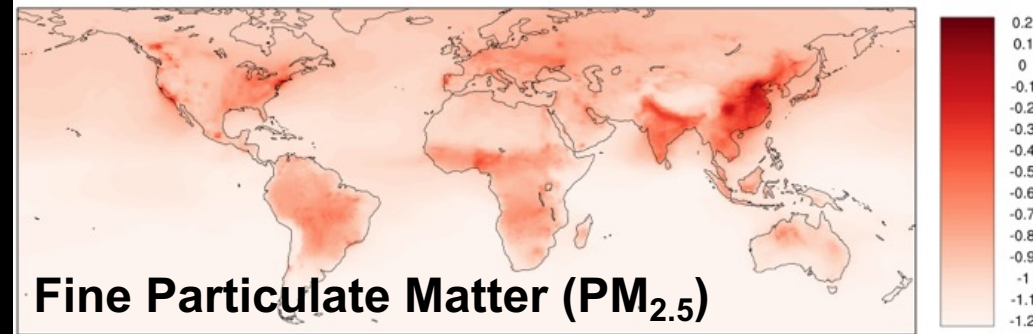
GEOS Chemistry



Ozone (O_3)

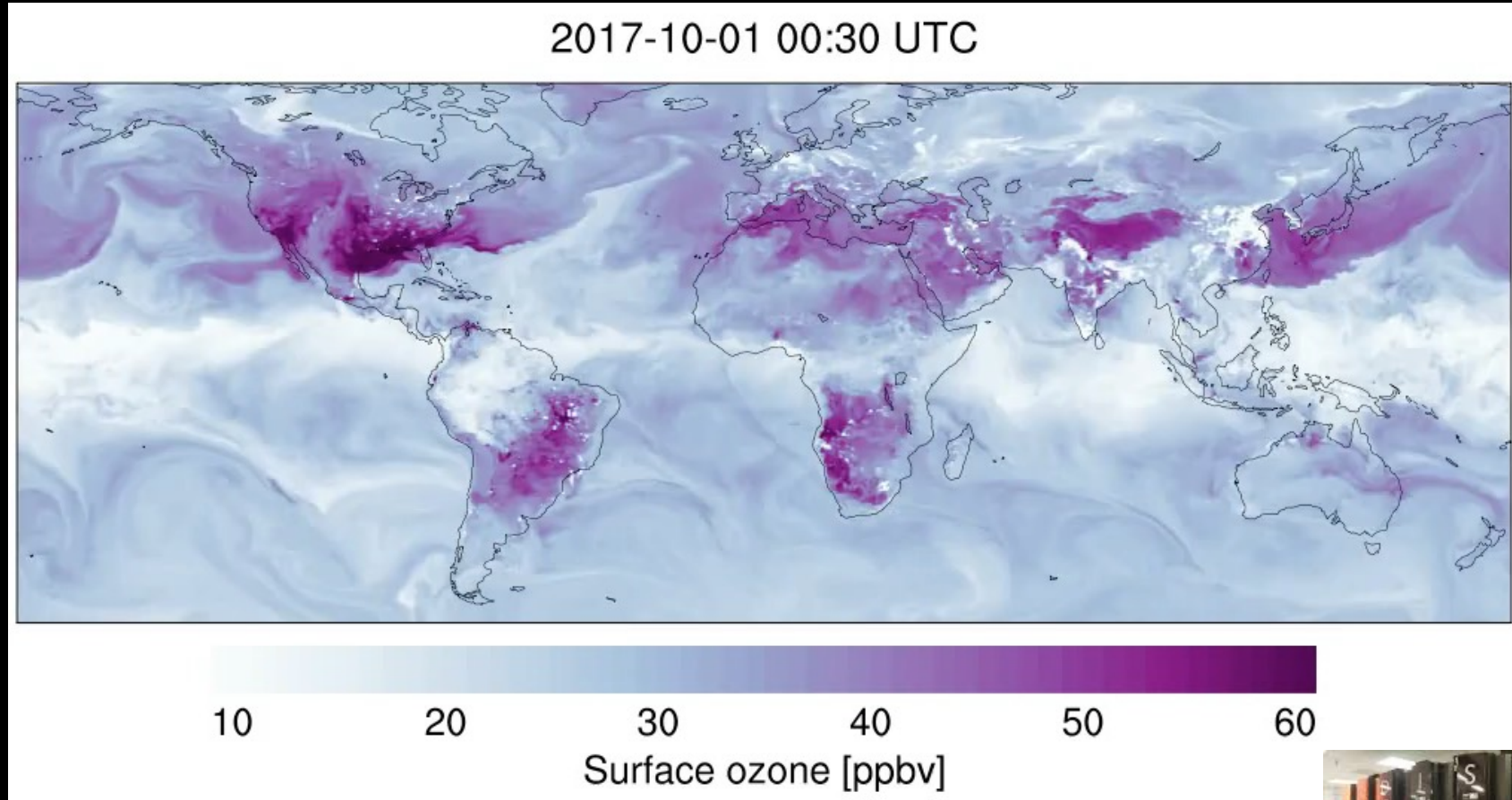


Nitrogen Dioxide (NO_2)



Fine Particulate Matter ($PM_{2.5}$)

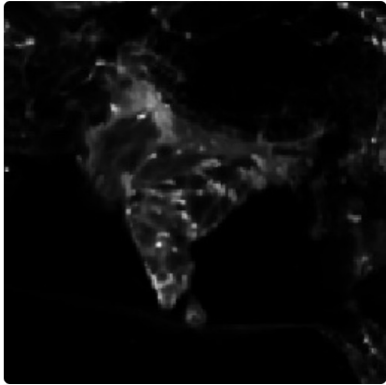
Producing the GEOS-CF forecasts is computationally expensive!



- 56 million grid cells (25x25 km², 72 levels), 250 chemical species
- ~3500 CPUs, 1 simulation days / hour, 1TB of output per day



GEOS-CF rpl_tavg1hr_v1: Goddard Earth Observing System Composition Forecast



DESCRIPTION BANDS TERMS OF USE CITATIONS DOIS

This dataset contains meteorological replay (rpl) of time-average one hour data (tavg1hr). It is built by merging the original GEOS-CF collections `chm_tavg_1hr_g1440x721_v1`, `met_tavg_1hr_g1440x721_x1`, and `xgc_tavg_1hr_g1440x721_x1`. 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.

Evaluation of GEOS-CF against satellite, ozonesonde, and surface observations for years 2018–2019 shows realistic simulated concentrations of O3, NO2, and CO, with normalized mean biases of -0.1 to 0.3, normalized root mean square errors between 0.1–0.4, and correlations between 0.3–0.8. Comparisons against surface observations highlight the successful representation of air pollutants in many regions of the world and during all seasons, yet also highlight current limitations, such as a global high bias in SO2 and an overprediction of summertime O3 over the Southeast United States.

CLOSE

IMPORT

Dataset Availability

2018-01-01T00:00:00 -

Dataset Provider

[NASA / GMAO](#)

Collection Snippet 

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

[See example](#)

Tags

composition

forecast

geos

gmao

nasa

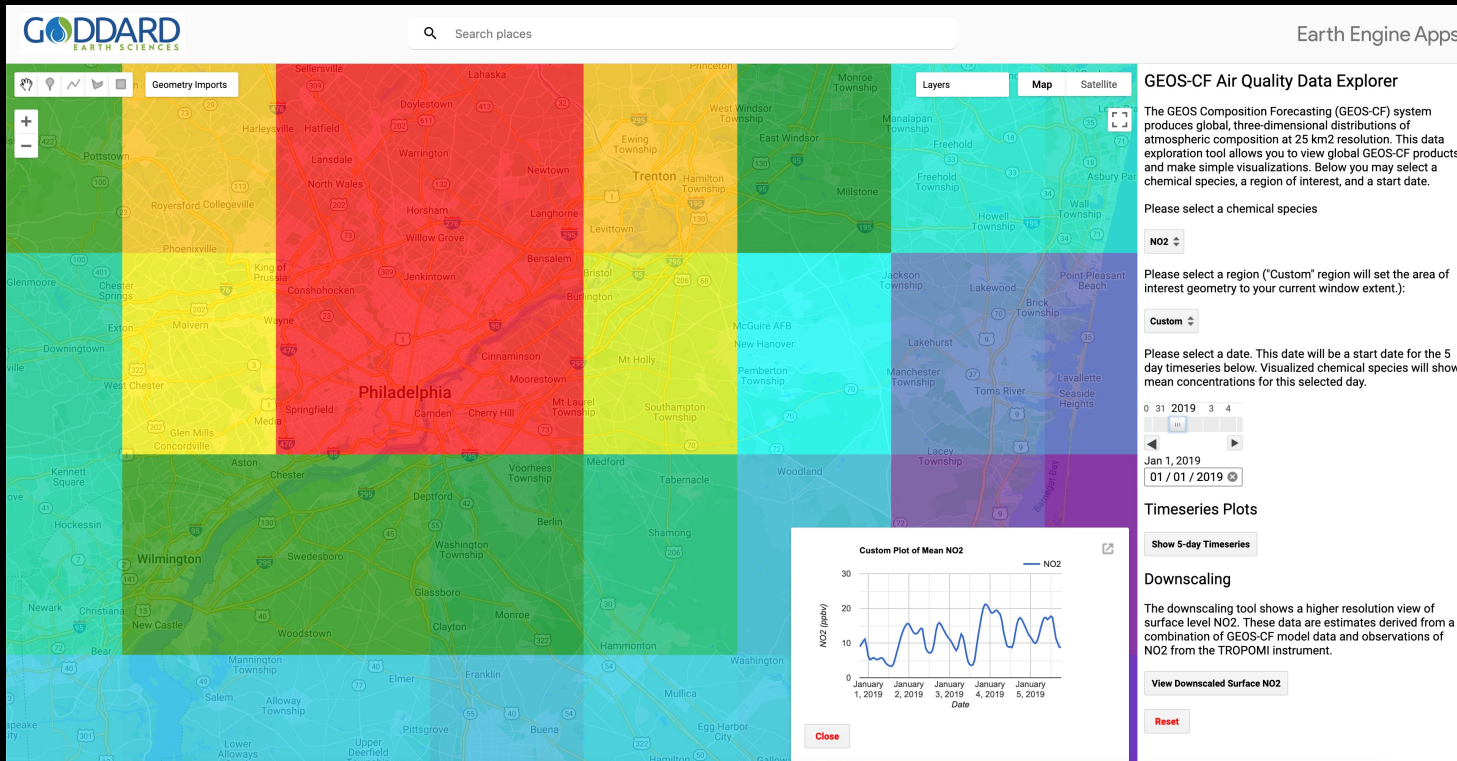
Accessing GEOS-CF through Google Earth Engine

Google Earth Engine (GEE) provides a rich open-access data repository to users

The data catalog contains metadata and documentation that allows anyone to manipulate and visualize data sets

The NASA and Google AQ partnership is providing a framework for GMAO data to be made available on GEE

Accessing GEOS-CF through Google Earth Engine



Currently the GEOS-CF 15-minute high-temporal frequency (“htf”) replay collection and the hourly time-averaged (tavg1hr) replay collection have been ingested in the GEE data repository

These collections consist of surface and column chemistry and meteorology fields

The htf and tavg collections are both publicly available through the Earth Engine data repository

Our partners at Google are currently ingesting forecast data sets which will be used for various machine learning applications

<https://callumwayman.users.earthengine.app/view/geoscfexplorer>

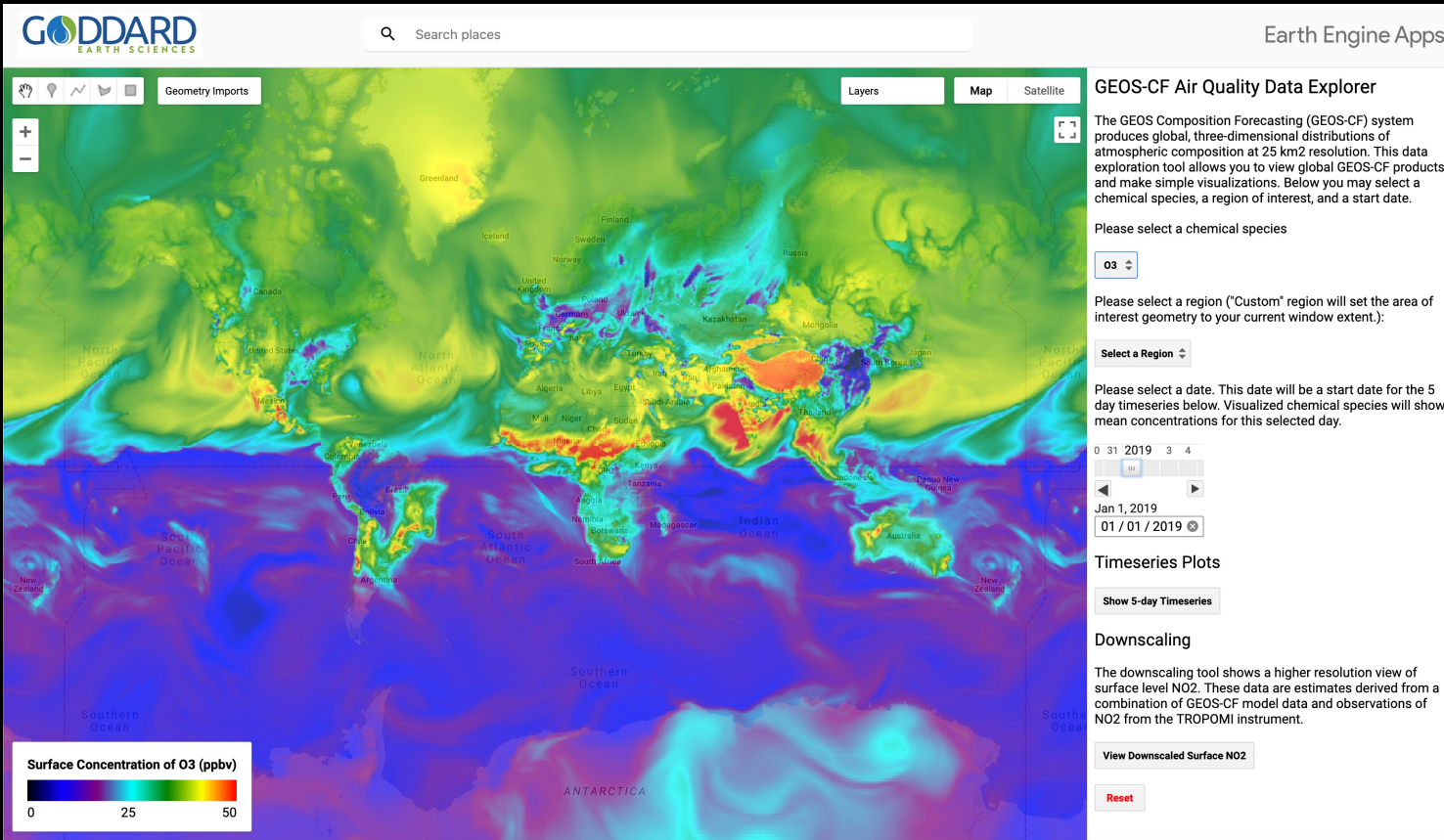
Dynamic Applications with GEE

GEE applications can be produced and easily exported

Dynamic interactive web-map tools will allow for greater flexibility in the ways in which the GMAO can provide data

Open-source data provides opportunities for anyone to produce tools using GMAO data

We are incorporating GMAO data and TROPOMI observations into these applications to provide localized surface level estimates of NO_2 with downscaling techniques



<https://callumwayman.users.earthengine.app/view/geoscfexplorer>

TROPOMI Observed Air Quality

The TROPOMI instrument is the single payload on the Sentinel-5P spacecraft

This instrument is dedicated to a mission of atmospheric monitoring

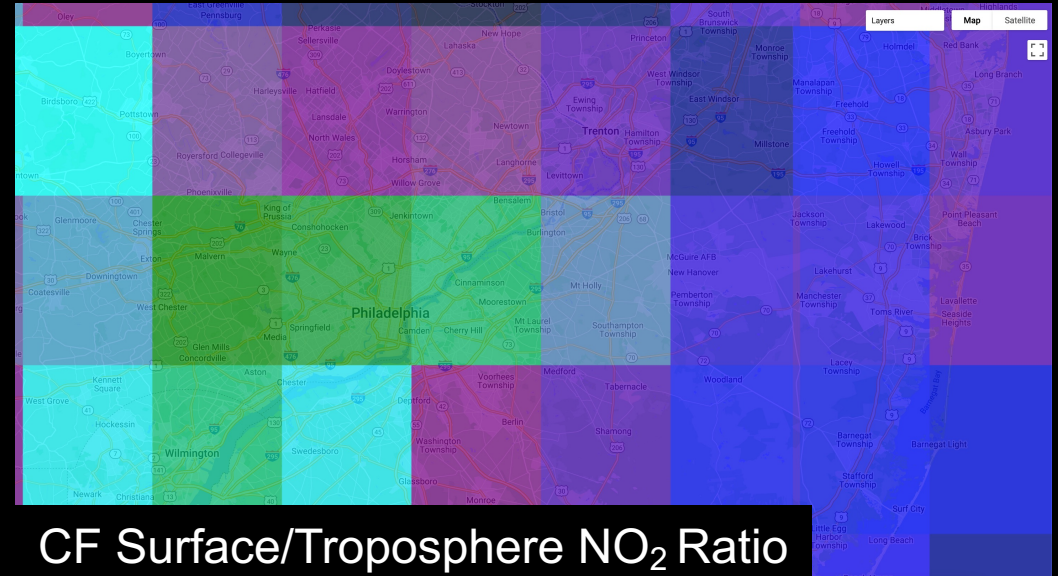
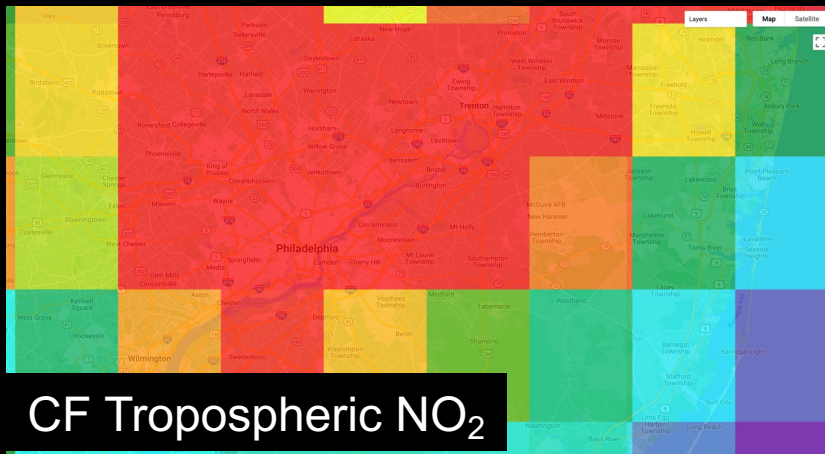
TROPOMI observations of tropospheric NO₂ have an approximate 7x3.5 km² pixel resolution

<https://sentinel.esa.int/web/sentinel/user-guides/sentinel-5p-tropomi>

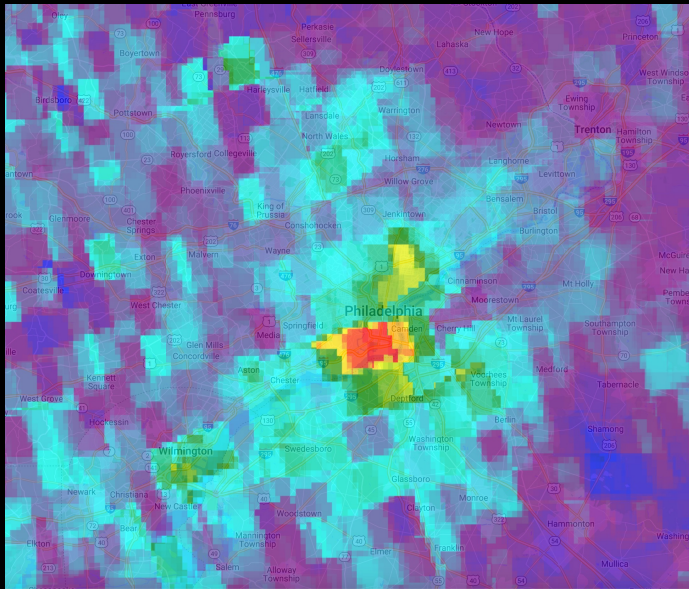


https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P/Introducing_Sentinel-5P

Using Observations with Model Data to Downscale GEOS-CF



Using Observations with Model Data to Downscale GEOS-CF



TROPOMI Tropospheric NO₂

*



CF Surface/Troposphere NO₂ Ratio

=

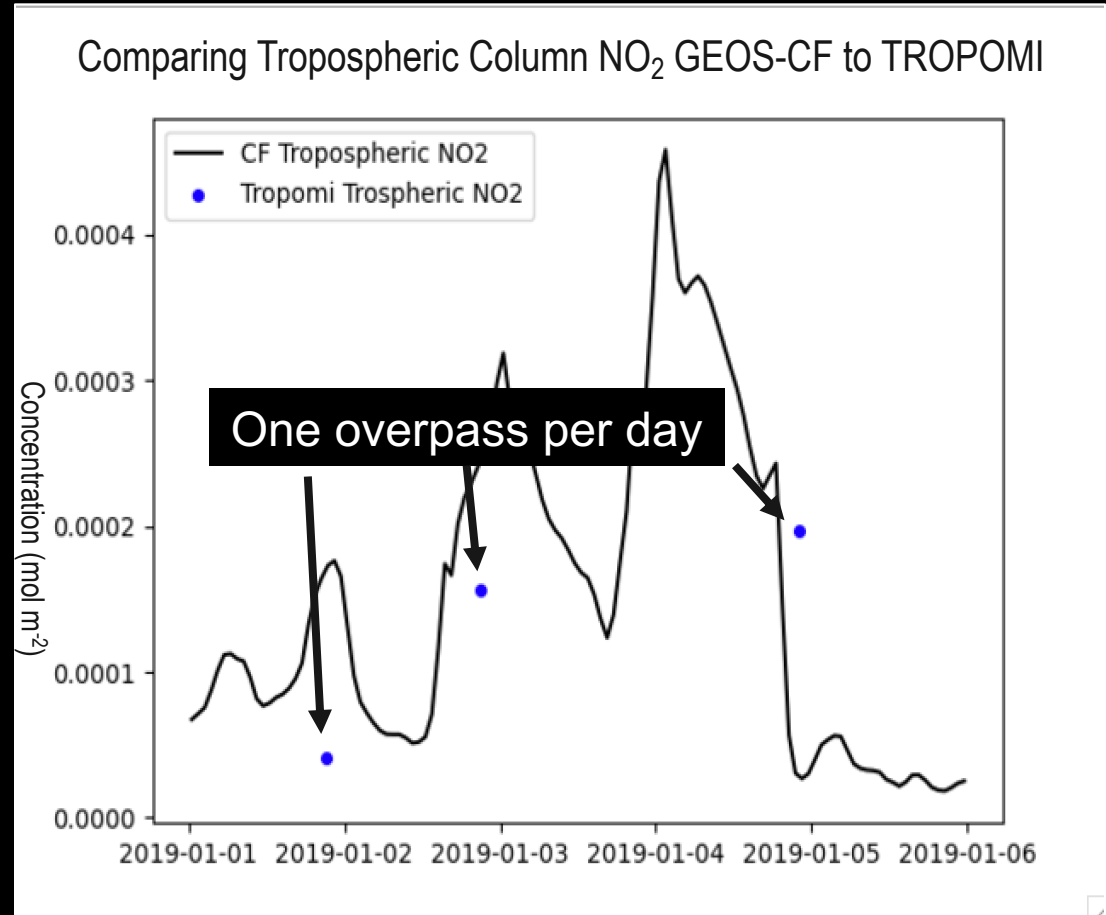


Surface Scaled TROPOMI NO₂

Using Earth Engine Built-In Machine Learning

Google Earth Engine has built-in machine learning algorithms such as the random forest classifier

Our next step is to incorporate other data sets into machine learning models to downscale GEOS-CF imagery on time steps without TROPOMI overpasses





Incorporating Proprietary Data

After incorporating GEOS-CF meteorology and TROPOMI data, Google proprietary data sets will be incorporated into the machine learning model

Google has developed some rich data sets which provide point source air quality concentrations using Google Street View vehicles in several cities around the world

In addition, the Open Buildings data set and the National Landcover Dataset may be used to better understand how buildings, pieces of infrastructure, and changing landuse are contributing to atmospheric NO₂ concentrations



Summary


The combined data infrastructures of Google and NASA provide an excellent foundation for new tools and applications

Simple Ratio techniques have been used to produce higher spatial resolution surface estimates of NO₂

Next steps:

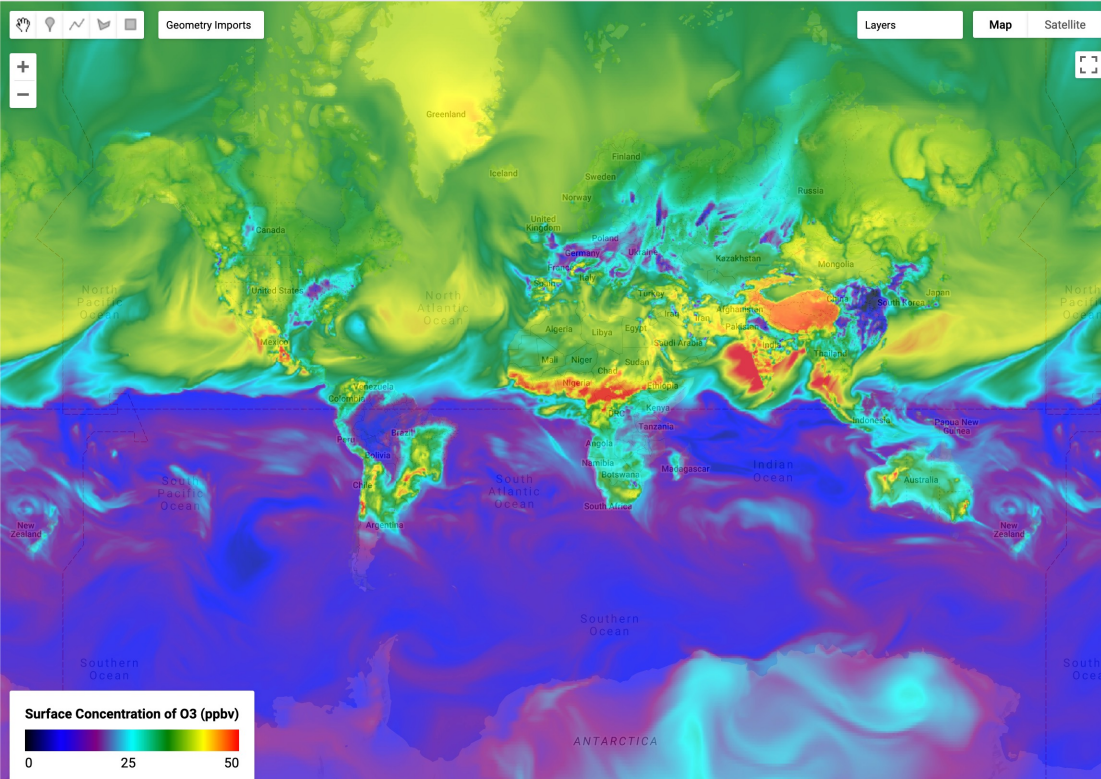
- Use machine learning within GEE to spatially downscale the GEOS-CF data to finer grids
- Access proprietary Google data at the street-level
 - This could improve ability to locate high concentrations of NO₂ within cities.

Questions?

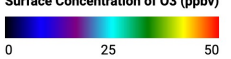


Earth Engine Apps

Geometry Imports
Layers
Map
Satellite



Surface Concentration of O3 (ppbv)



GEOS-CF Air Quality Data Explorer

The GEOS Composition Forecasting (GEOS-CF) system produces global, three-dimensional distributions of atmospheric composition at 25 km² resolution. This data exploration tool allows you to view global GEOS-CF products and make simple visualizations. Below you may select a chemical species, a region of interest, and a start date.

Please select a chemical species

Please select a region ("Custom" region will set the area of interest geometry to your current window extent):

Please select a date. This date will be a start date for the 5 day timeseries below. Visualized chemical species will show mean concentrations for this selected day.

0 31 2019 3 4

Jan 1, 2019

Timeseries Plots

Downscaling

The downscaling tool shows a higher resolution view of surface level NO₂. These data are estimates derived from a combination of GEOS-CF model data and observations of NO₂ from the TROPOMI instrument.

callum.r.wayman@nasa.gov