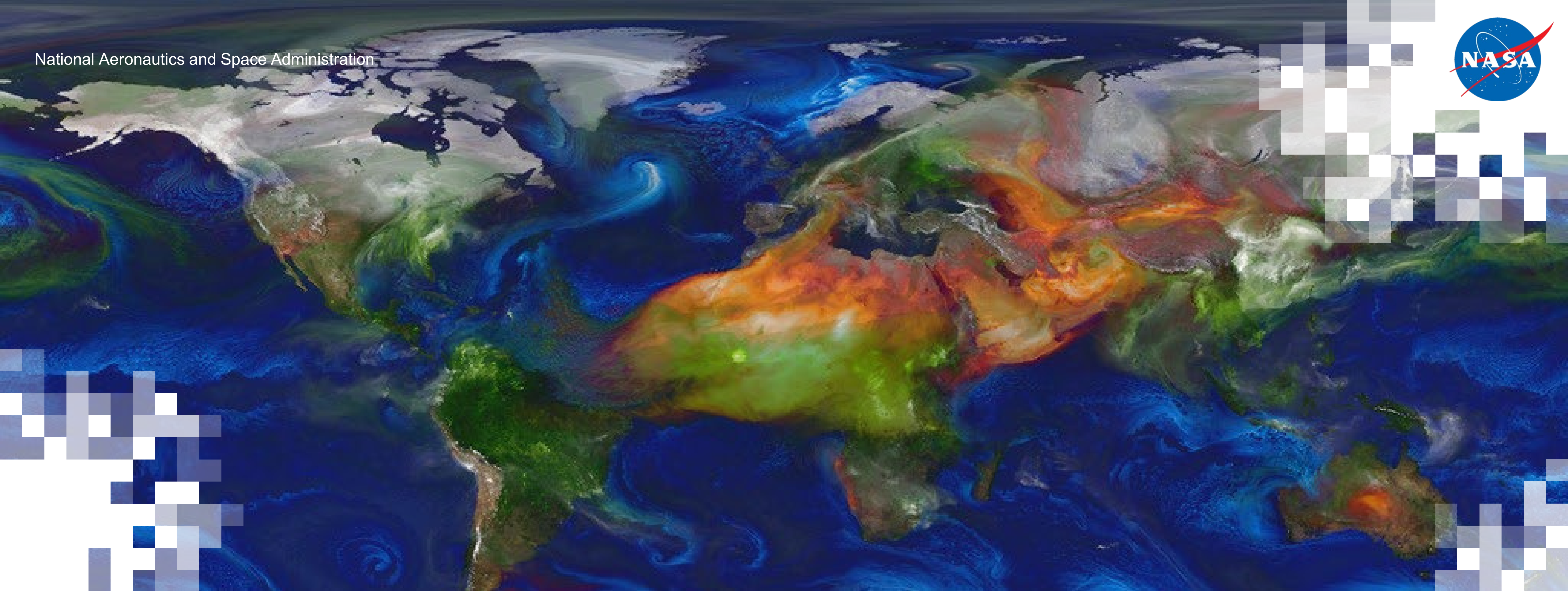


Global Air Quality Forecasting at NASA

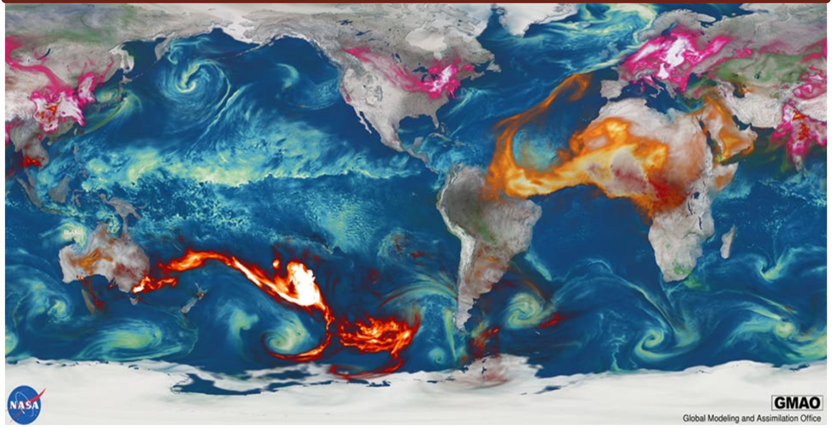
Melanie Follette-Cook & Pawan Gupta

September 28, 2021

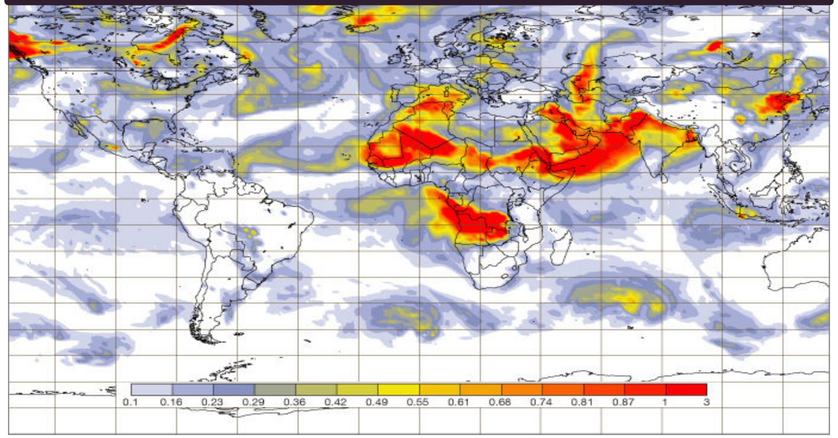
Webinar Agenda

Part 1:
Air Quality (AQ) Basics

Part 2:
NASA GEOS Model

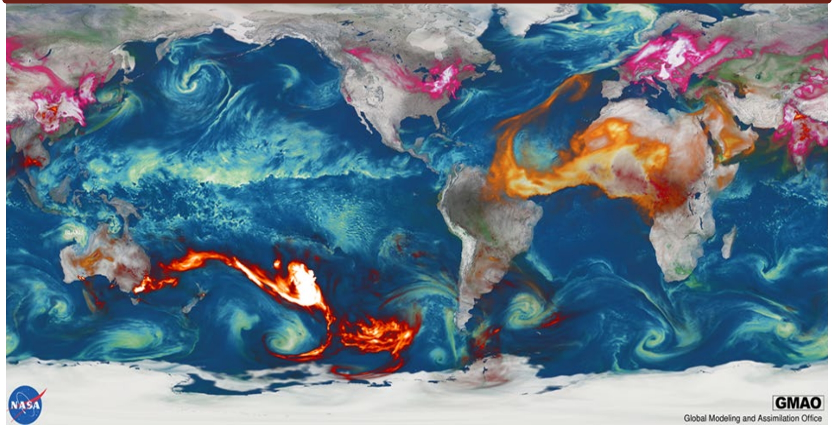


Part 3:
ECMWF CAMS Model



Webinar Agenda

Session 2:
NASA GEOS Model



Melanie Follette-Cook

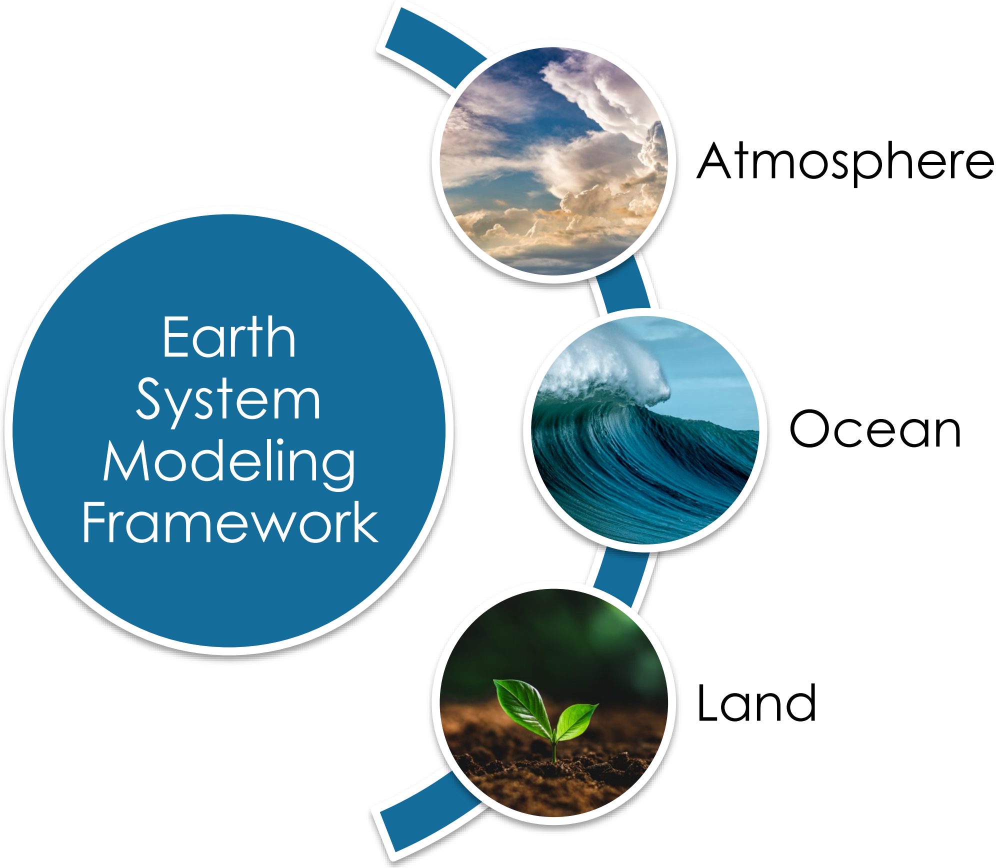
Pawan Gupta

Learning Objectives

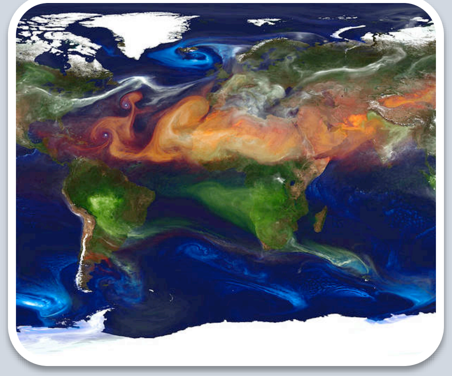
- Identify the different Air Quality (AQ) relevant model outputs available from Goddard Earth Observing System (GEOS) Earth System Model
- Understand the difference between analysis, reanalysis, and forecasting
- Understand the different ways satellite observations are used for forecasting, reanalysis, and evaluation
- Discover how to subset and visualize reanalysis and forecast outputs

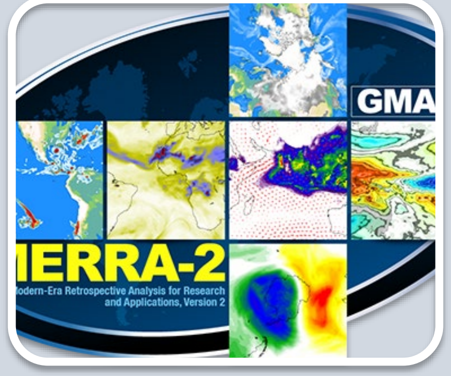
NASA GEOS Earth System Model

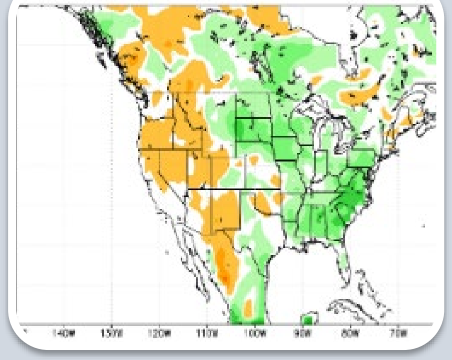
https://gmao.gsfc.nasa.gov/

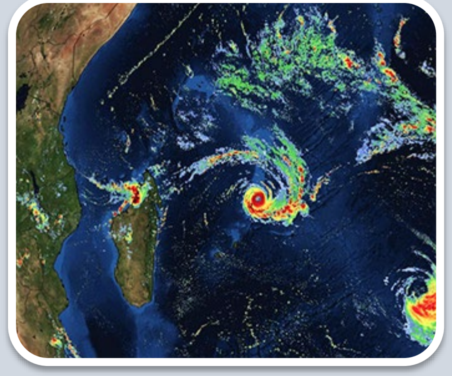


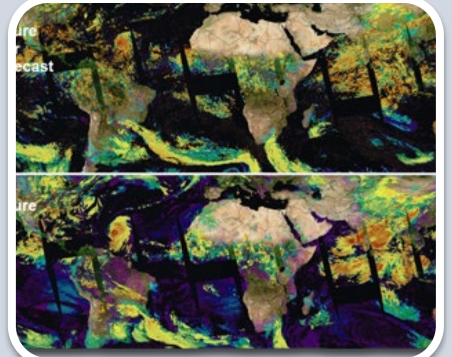
NASA GEOS Earth System Model







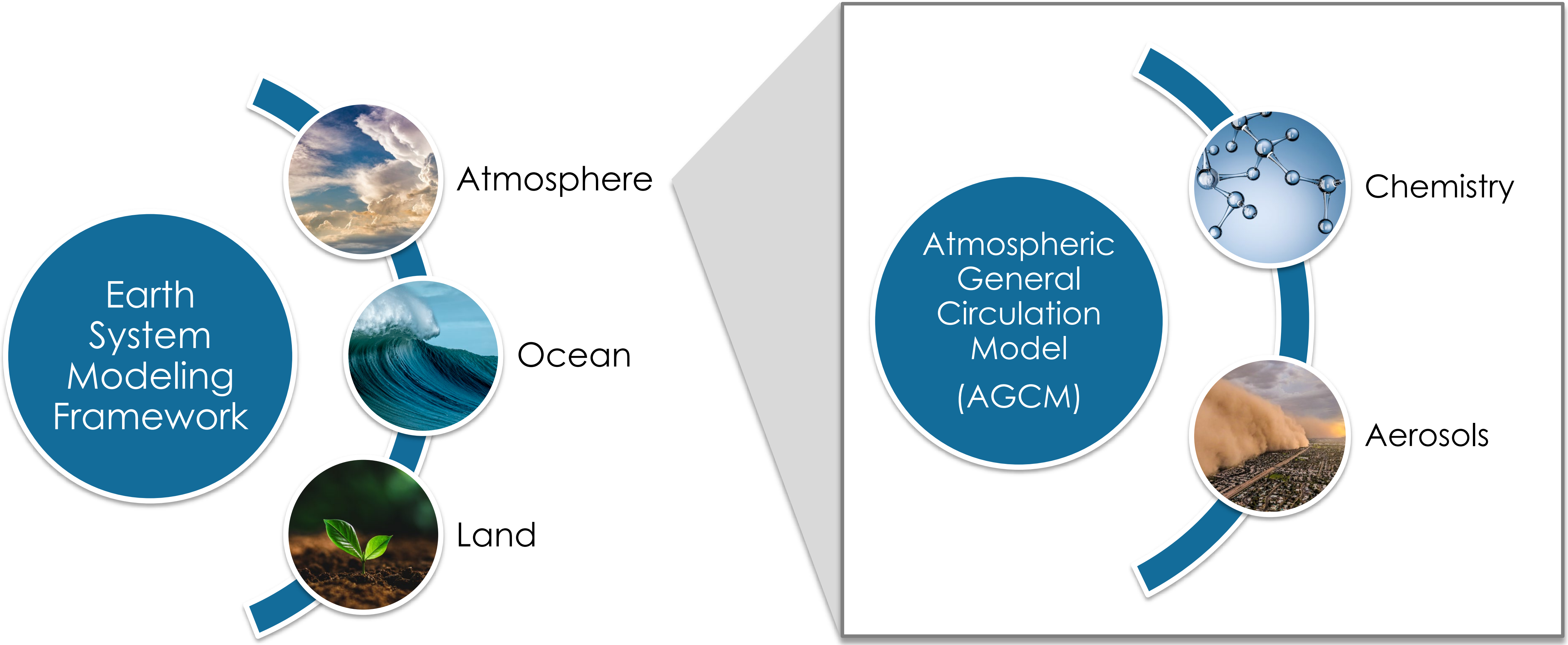




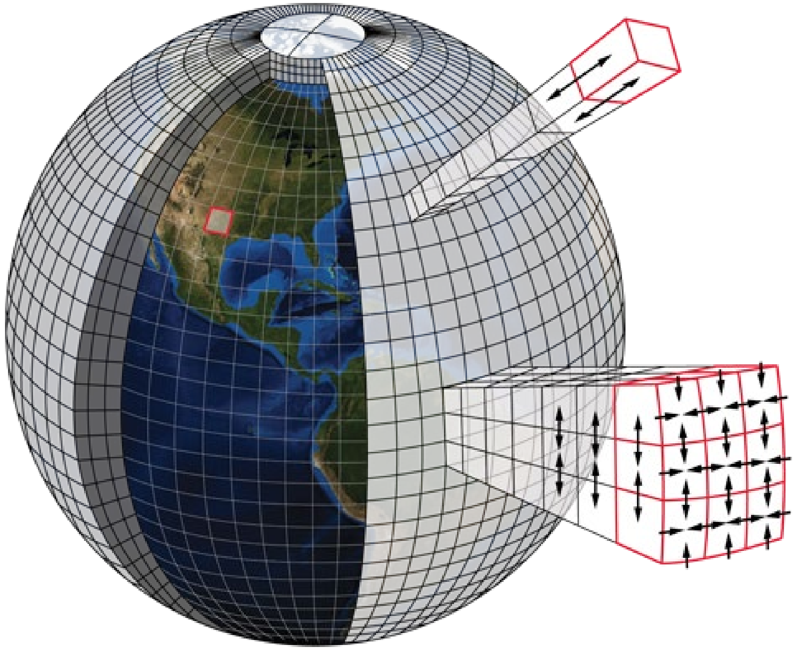
- NRT Weather and Chemical Forecasts
- Reanalysis
- Seasonal to Sub-Seasonal (S2S) and Decadal Prediction
- High Resolution Mesoscale Modeling
- Observing System Science

NASA GEOS Earth System Model

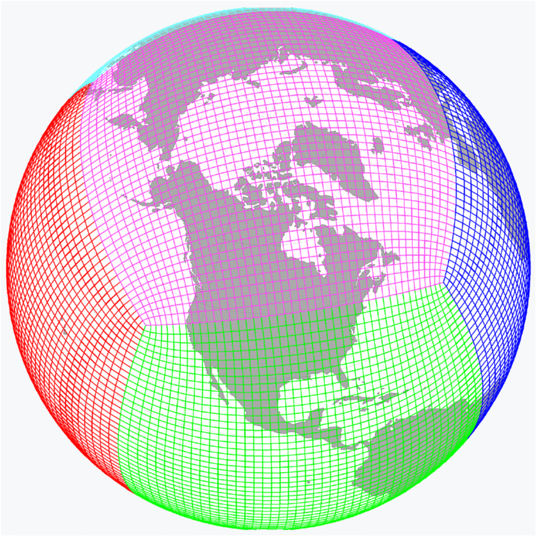
https://gmao.gsfc.nasa.gov/



Modeling the Atmosphere



Three-dimensional (3D) atmospheric chemistry models divide the atmosphere into a set of 3D grid cells.



The GEOS AGCM is run on a *cubed sphere* grid.

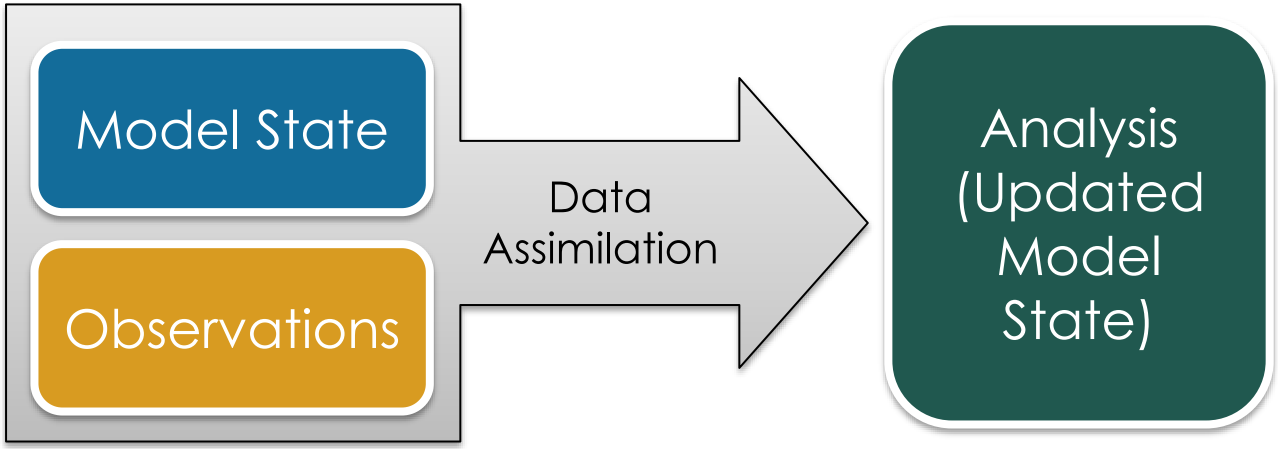
- Ensures uniform spatial grid
- Better for scalability

Models solve equations for physics, transport, and chemistry within each grid cell.

Image Source: https://wxguys.ssec.wisc.edu/2019/03/04/models/

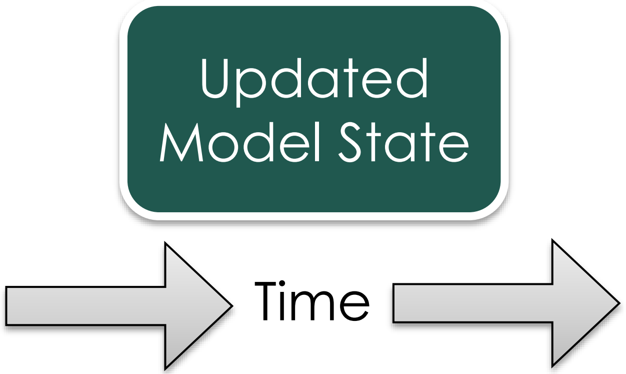
Forecast, Analysis, Reanalysis, Data Assimilation
What are the differences between these?

Data assimilation describes the process of assimilating, or incorporating, observations into a model state to produce the best estimate of the atmosphere, land, and ocean conditions.



An **analysis** is the blend of the model and observations.

A **reanalysis** blends observations with model simulations of the past using a single model version.



A **forecast** is a model simulation run forward in time to predict a future state.

GEOS Forecast and Reanalysis Products

GEOS Forward Processing (GEOS FP) NRT Analysis and Forecast

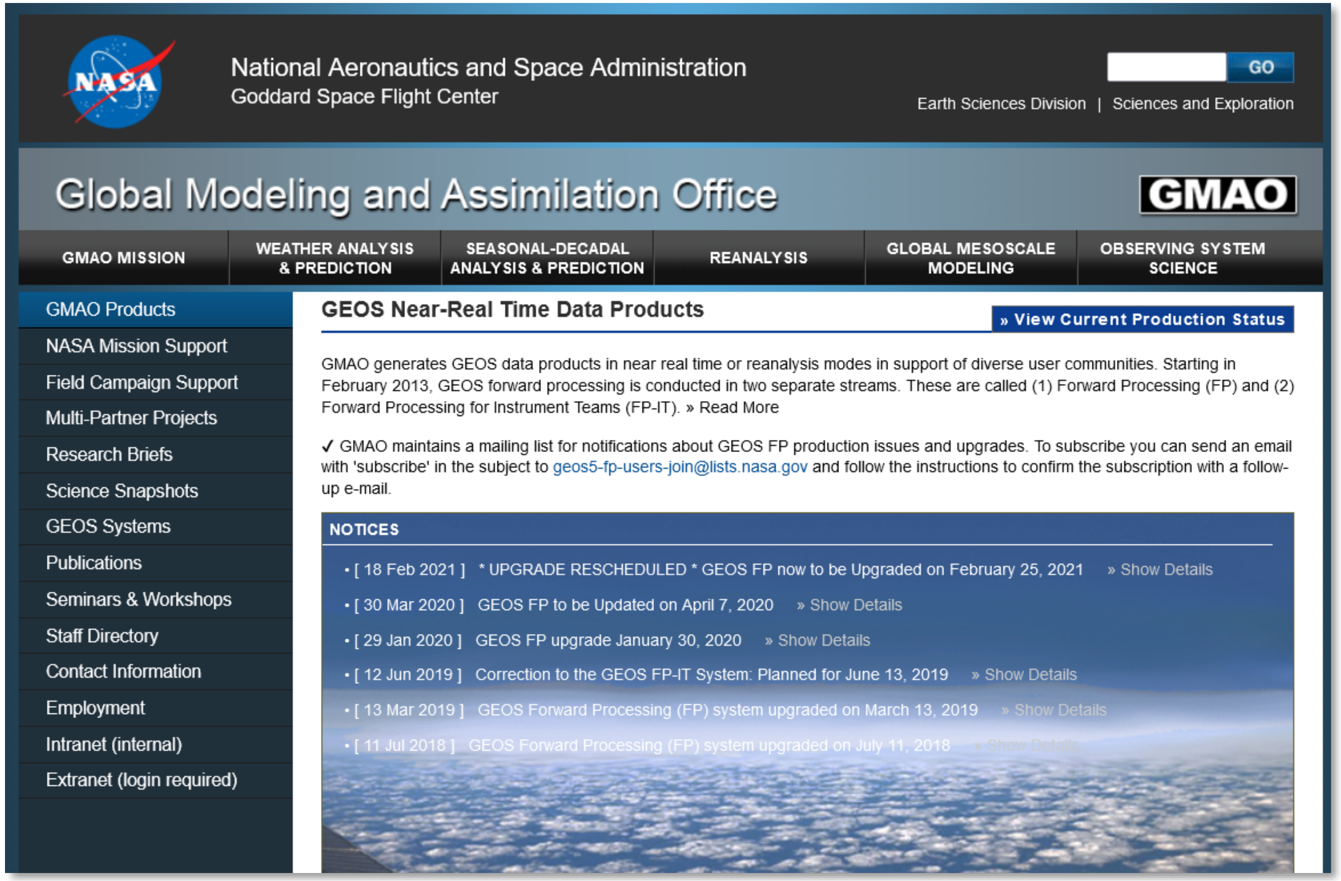
GEOS-Composition Forecast (GEOS-CF) NRT Forecast

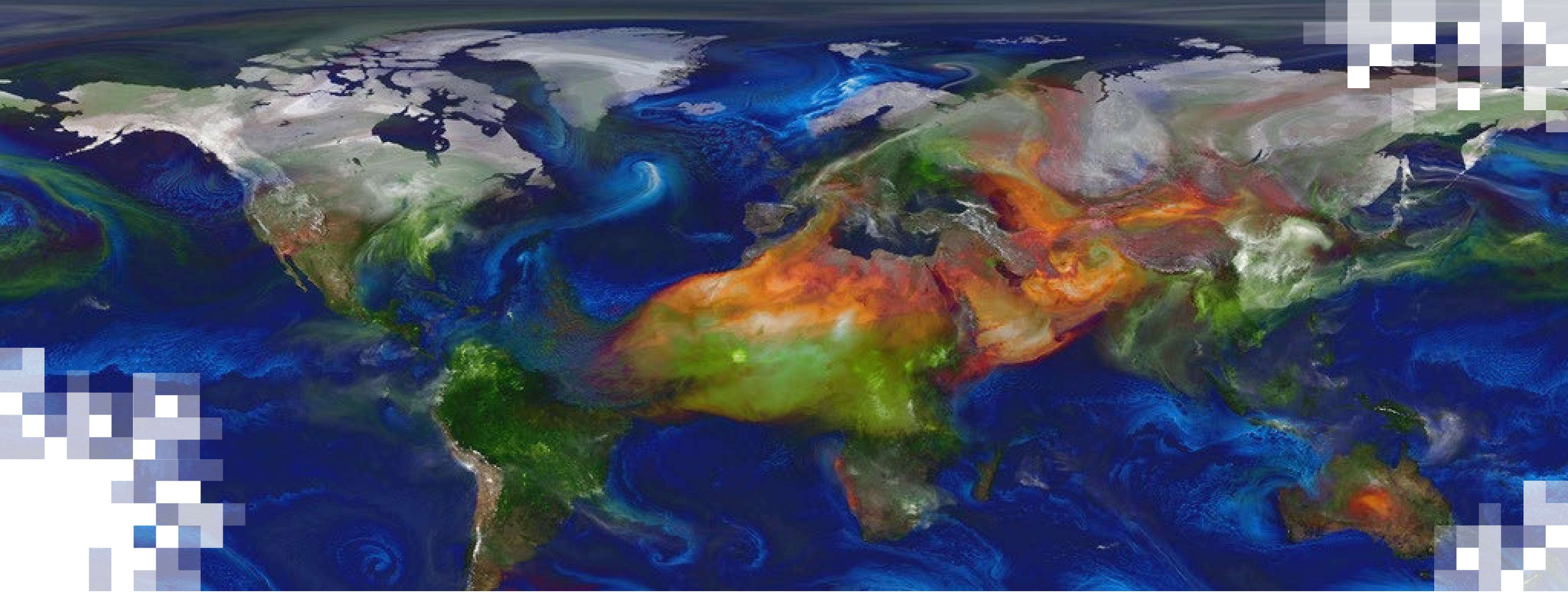
Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) Reanalysis

Global Modeling and Assimilation Office (GMAO)

https://gmao.gsfc.nasa.gov/GMAO_products/NRT_products.php

- The GEOS forecasts are a dynamic system.
- For the most current information about GEOS NRT products, visit the GMAO products website.
- The most current version of file specification documents are covered here: https://gmao.gsfc.nasa.gov/pubs/office_notes.php





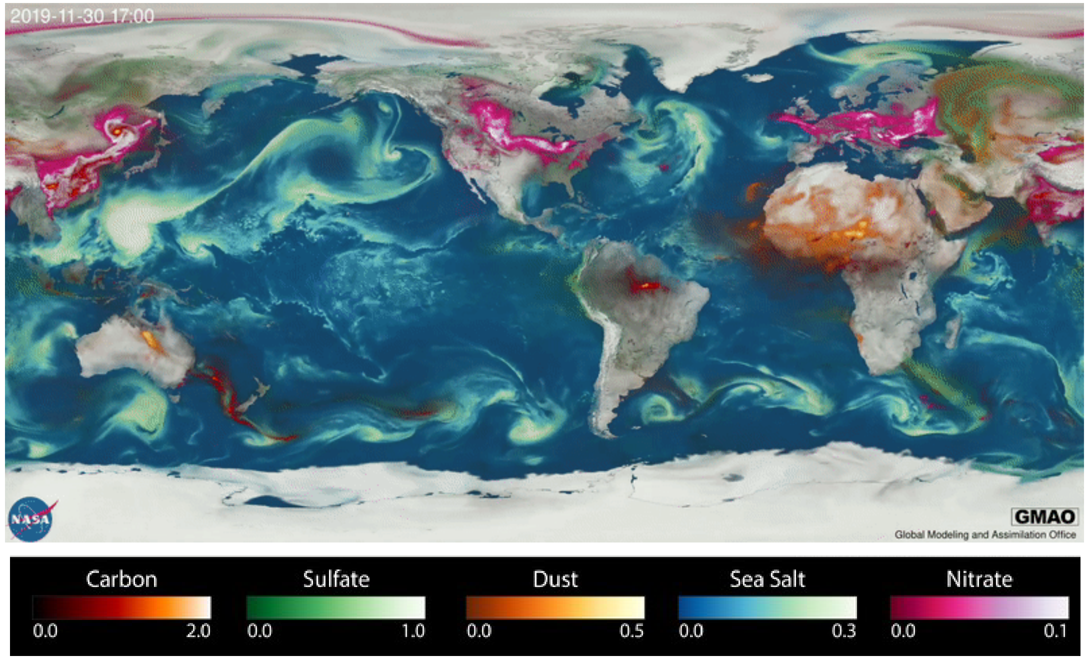
GEOS FP

GEOS FP

https://gmao.gsfc.nasa.gov/weather_prediction/

- GEOS FP analyses and forecasts support NASA field campaigns and provide a testbed for assimilation and forecast development
- Publicly available
- Includes weather, aerosols, and carbon monoxide (CO) on the same spatial scale
- State of the science forecast system – model physics or observing system updated every 6-12 months
  – Not suitable for trend analyses

AOD (550 nm)



https://svs.gsfc.nasa.gov/31100

- Meteorology used to drive chemistry models:
  – GEOSChem, Whole Atmosphere Community Climate Model (WACCM)
- When using FP meteorology fields to drive another model, must ensure your simulation does not span an update
  – GMAO NRT Product Page has updated details and dates

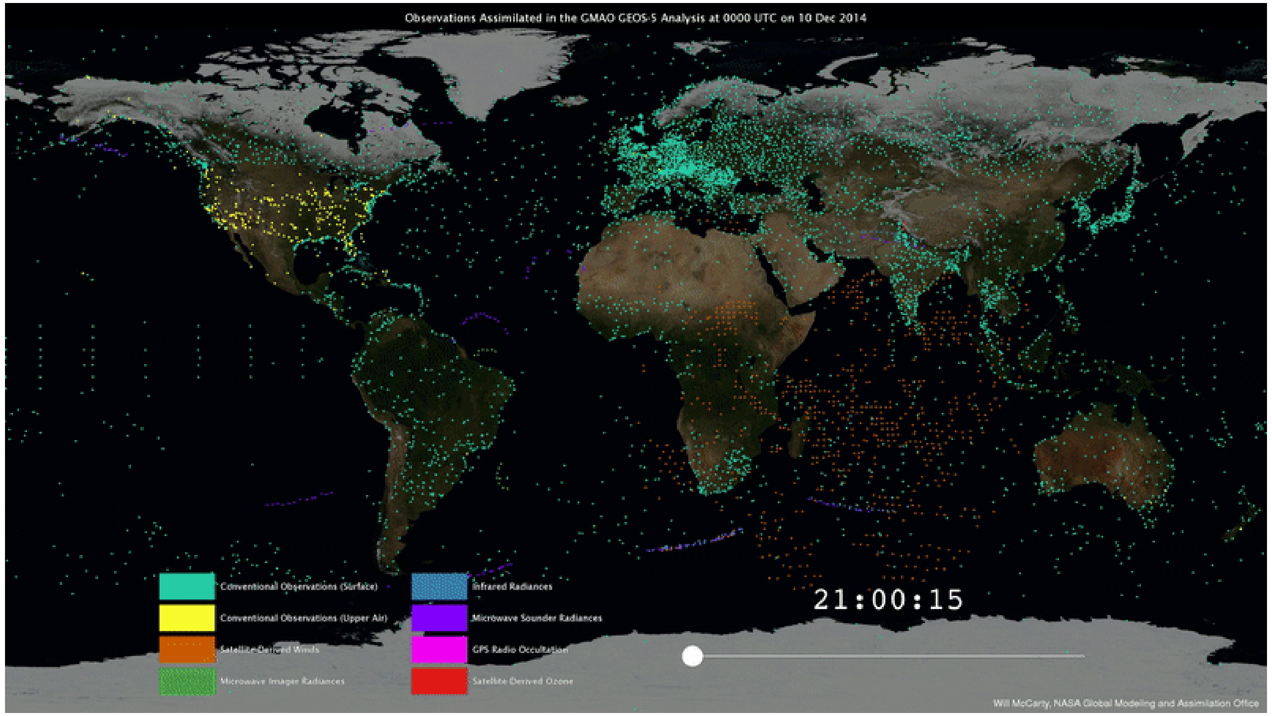
GEOS Output Quick Guide

	GEOS FP
Type	Analysis + Forecast
Domain	Global
Spatial Resolution	Simulation: ~12 km Output: ~25 km (0.25°x0.312°)
Temporal Resolution	2-D data: Hourly 3-D data: Every 3 h
Vertical Levels	72 (near surface-0.1 hPa)
Output Available	Analysis: 2014 – Present Forecast: ~21 days
Initialization	Daily 10-day forecast at 00Z Daily 5-day forecast at 12Z
Data Assimilation	Yes
File Specification Doc	https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi1203.pdf *

* Find most current File Specification at https://gmao.gsfc.nasa.gov/pubs/office_notes.php

Data Assimilation in GEOS FP

- **Data assimilation** describes the process of assimilating, or incorporating, observations into a model state to produce the best estimate of the atmosphere, land, and ocean conditions.
- GEOS uses a Hybrid 4D-Ensemble Variational (Hyb-4DEnVar) approach.
- Analyses are created every 6 hours using over 5 million observations.
- GEOS assimilates AOD from MODIS (Terra and Aqua).

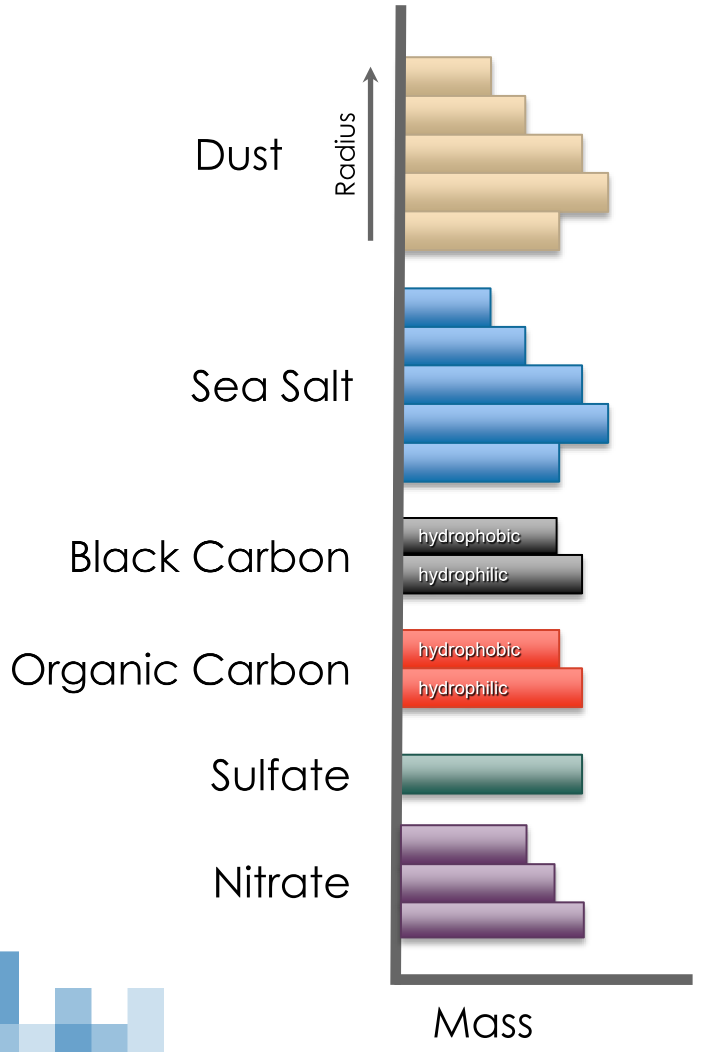


https://svs.gsfc.nasa.gov/30590

- Important: AOD is the total extinction of all aerosol species, so when AOD is increased or decreased as a result of the assimilation, assumptions are made about how that is distributed among the species.

GOCART in GEOS

- Goddard Chemistry, Aerosol, Radiation and Transport Model (GOCART, Chin et al. 2002, Colarco et al. 2010)
- Sources and sinks for 6 non-interactive species
- Radiatively active



Wind and topographic sources, 5 mass bins

Wind-driven source, 5 mass bins

Anthropogenic and wildfire sources, mass hydrophobic & hydrophilic

Anthropogenic, biogenic, and fire sources, mass hydrophobic and hydrophilic

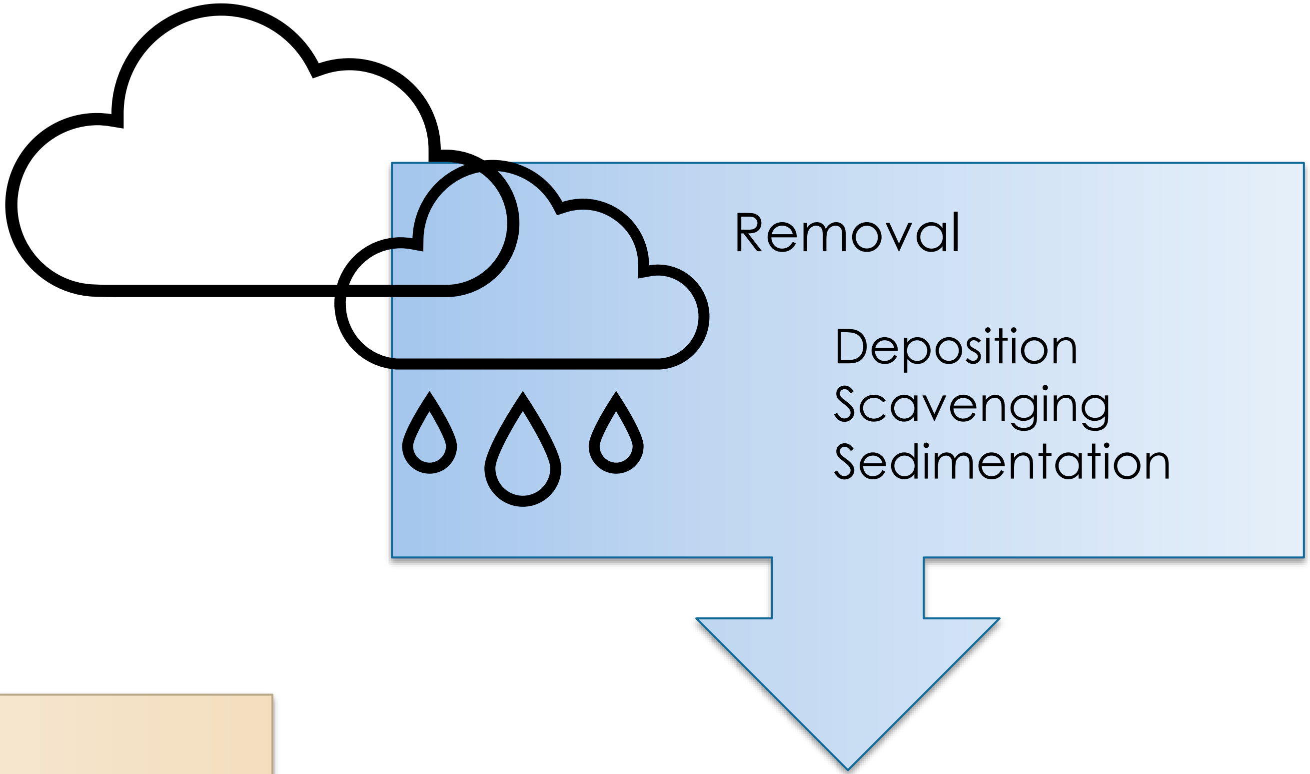
Anthropogenic, wildfire, and volcanic

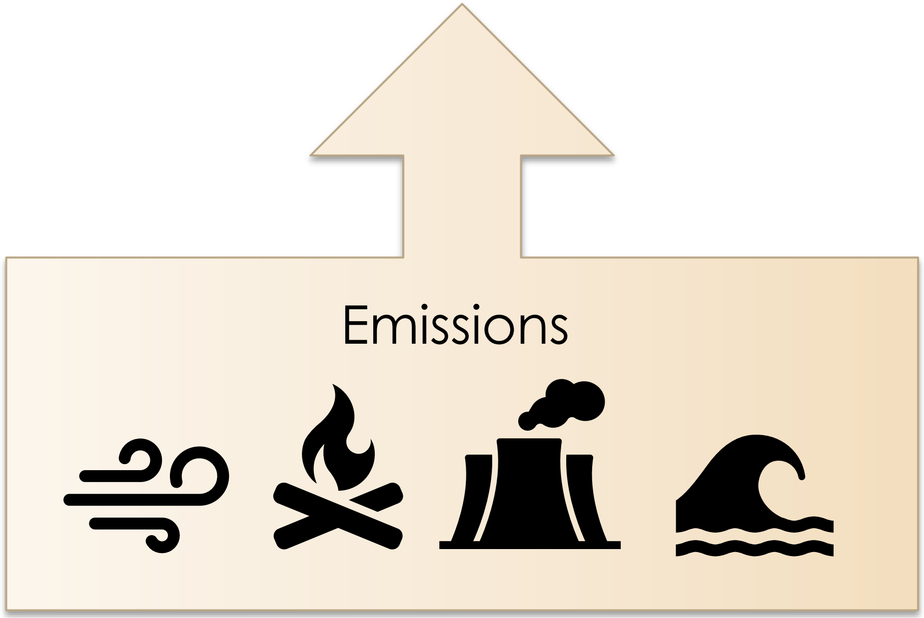
Anthropogenic and wildfire sources

Future version of GOCART in GEOS FP will include Brown Carbon (BC)

Aerosol Processes

GOCART simulates the emission, chemistry, and removal of aerosols and CO within each grid cell.





Aerosol Emissions:

- Dust: Ginoux et al., 2001
- Fire emissions: Quick Fire Emissions Database (QFED), Darmenov et al., 2015
- SO_2: Hemispheric Transport of Air Pollution (HTAP), enhanced by OMI SO_2 observations, Janssens-Maenhout et al. (2015), Liu et al. (2018)
- Anthropogenic emissions: HTAP
- Volcanic emissions: Carn (2019)

For more information about QFED fire emissions, see Sessions 3 and 4 of our webinar [Satellite Observations and Tools for Fire Risk, Detection, and Analysis]

Example GEOS FP File Name

GEOS.fp.fcst.tavg3_2d_aer_Nx.20210901_00+20210902_1330.V01.nc4

GEOS FP File Collections

- GEOS FP output is organized into file *collections* that contain related variables.
- These have the form:

Frequency_Dimensions_Group_HV

Frequency

Frequency or averaging interval

- const = time-independent
- inst = instantaneous
- tavg = time-average

Dimensions

Dimensions of variables

- 2d = only 2d fields
- 3d = can have 2d and 3d

Group

Three letter abbreviation for the type of variables

- Also used in the short name
- Ex. aer = Aerosol fields
- See documentation for full list

HV

Horizontal and vertical grid

- H = typically N, for nominal grid
- V = x, horizontal only
- V = p, pressure level
- V = v, model level
- V = e, model layer edges

GEOS FP Output File Names

- Each GEOS FP file has the form:

GEOS.fp.mode.collection.time.file_ver.nc4

Mode

- asm = assimilation
- fcst = forecast

Collection

See previous slide

File_ver

File version (usually V01)

Time

Date and time of data file

- For assimilation:
  yyyymmdd_hhmm
- For forecast:
  yyyymmdd_hhmm+yyyymmdd_hhmm

Assimilation cycle | Forecast time

All GEOS FP output files are in NetCDF-4 format.

Example GEOS FP File Name

GEOS.fp.fcst.tavg3_2d_aer_Nx.20210901_00+20210902_1330.V01.nc4

- fp – forward processing
- fcst – forecast product
- tavg3_2d_aer_Nx : 2D time-averaged aerosol diagnostics
- 20210901_00+20210902_1330 : Forecast initialized at 2021-09-01 00 Z. The valid time for the data in this file is 2021-09-02 1330 Z, which represents the center point of a 3-hour time-averaging period between 1200 and 1500 Z.

AQ-Relevant Collections and Variables

https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi1203.pdf *

Collection Name	Description
tavg3_2d_aer_Nx	2D time-averaged aerosol diagnostics *Optical properties (Extinction AOT, Scattering AOT, Angstrom parameter)* *Surface concentration (kg/m^3), Column Density (kg/m^2)*
tavg3_2d_adg_Nx	2D time-averaged aerosol diagnostics (extended) *Emissions and removal processes (deposition, sedimentation, and scavenging)*
tavg3_2d_chm_Nx	2D time-averaged chemistry diagnostics *Surface CO, column CO, emissions, chemical loss, chemical production*
inst3_3d_aer_Nv	3D instantaneous aerosol diagnostics *Mass mixing ratios (kg/kg) of aerosol species in each size bin*
inst3_3d_chm_Nv	3D instantaneous chemistry diagnostics *CO molar mixing ratio (mol/mol)*

To calculate $PM_{2.5}$, use the formula:

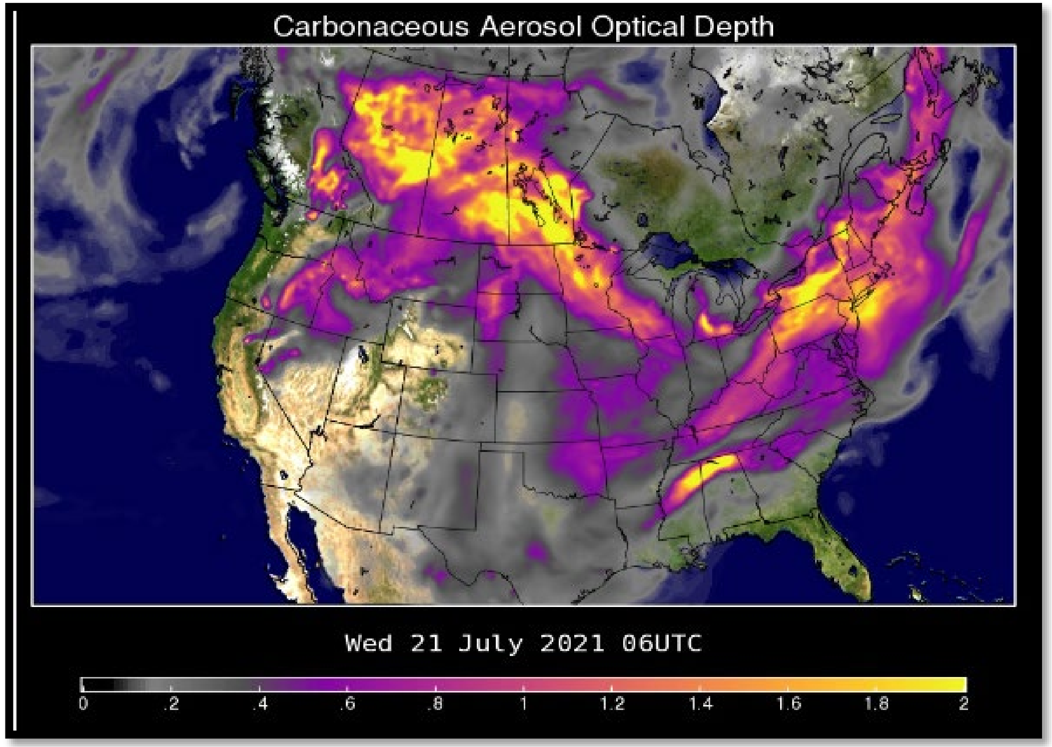
$PM_{2.5}$=[DUSMASS25]+[SSSMASS25]+[BCSMASS]+[OCSMASS]+1.375×[SO4SMASS]+1.29×[NISMASS25]

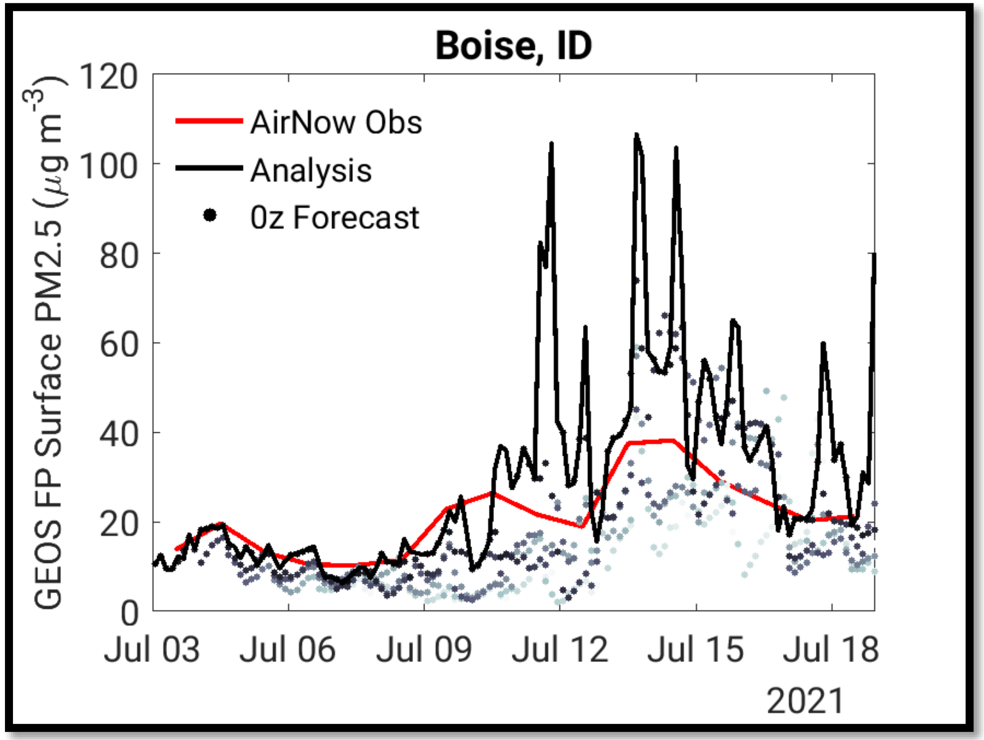
These variables are contained in the *tavg3_2d_aer_Nx* collection.

In the near future, GEOS FP output will include a $PM_{2.5}$ variable.

* Find most current File Specification at https://gmao.gsfc.nasa.gov/pubs/office_notes.php

Case Study: 2021 Biomass Burning Season





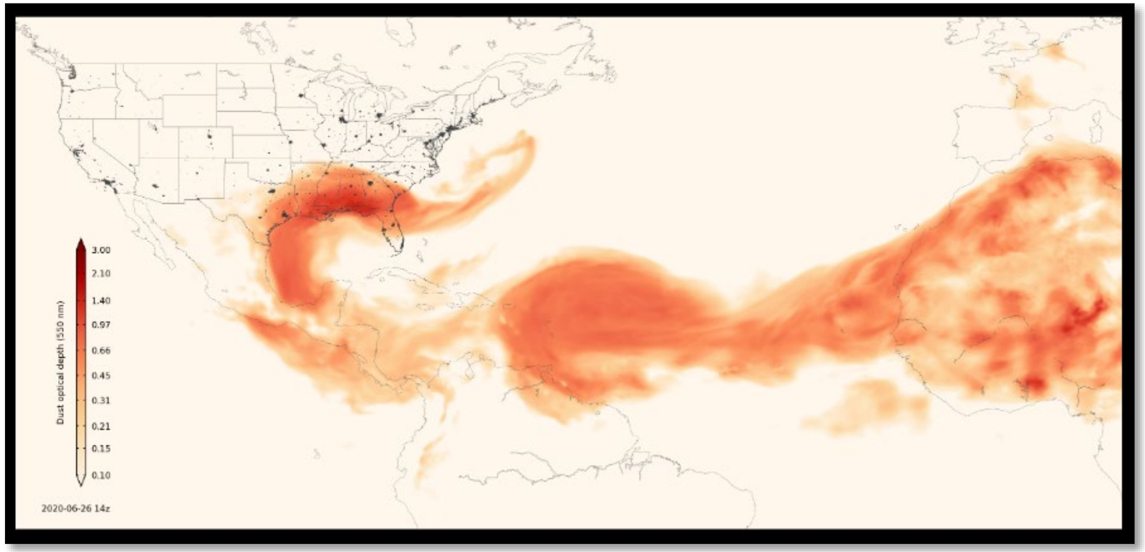
https://gmao.gsfc.nasa.gov/research/science_snapshots/2021/na_biomass_burning_2021.php

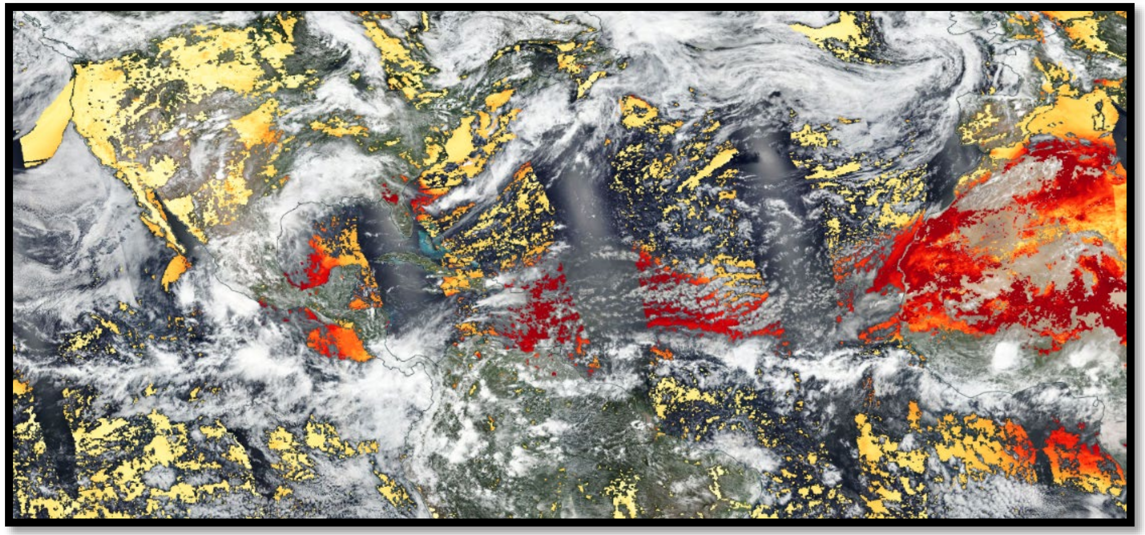
At the beginning of July 2021, GEOS FP was able to accurately simulate current conditions and forecast the amount of $PM_{2.5}$ as a result from the smoke from the wildfires in the region. However, beginning on July 10th, GEOS FP overestimated $PM_{2.5}$ at the surface near Boise, Idaho.

Case Study: June 2020 Dust Storm

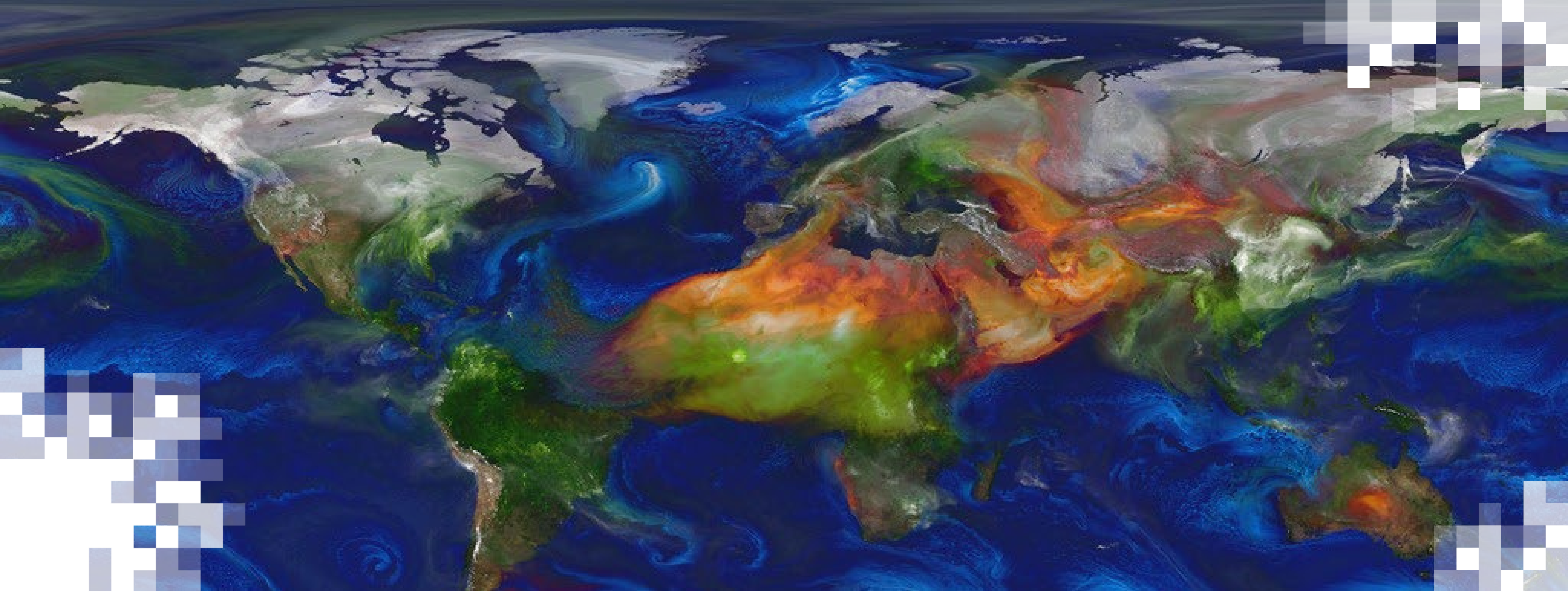
- During June 2020, dust was transported across the Atlantic Ocean.
  - Highest Caribbean aerosol loading in the last 20-year period
  - Surface concentrations almost three times higher than the 24-hour EPA standard

- GEOS FP analysis (top) shows the Saharan dust on June 26, 2020 at 14z
- VIIRS AOD (Deep Blue, Land and Ocean) and true color imagery shows the satellite observed plume on the same day (bottom, Worldview)





https://gmao.gsfc.nasa.gov/research/science_snapshots/2020/Saharan_dust_2020.php

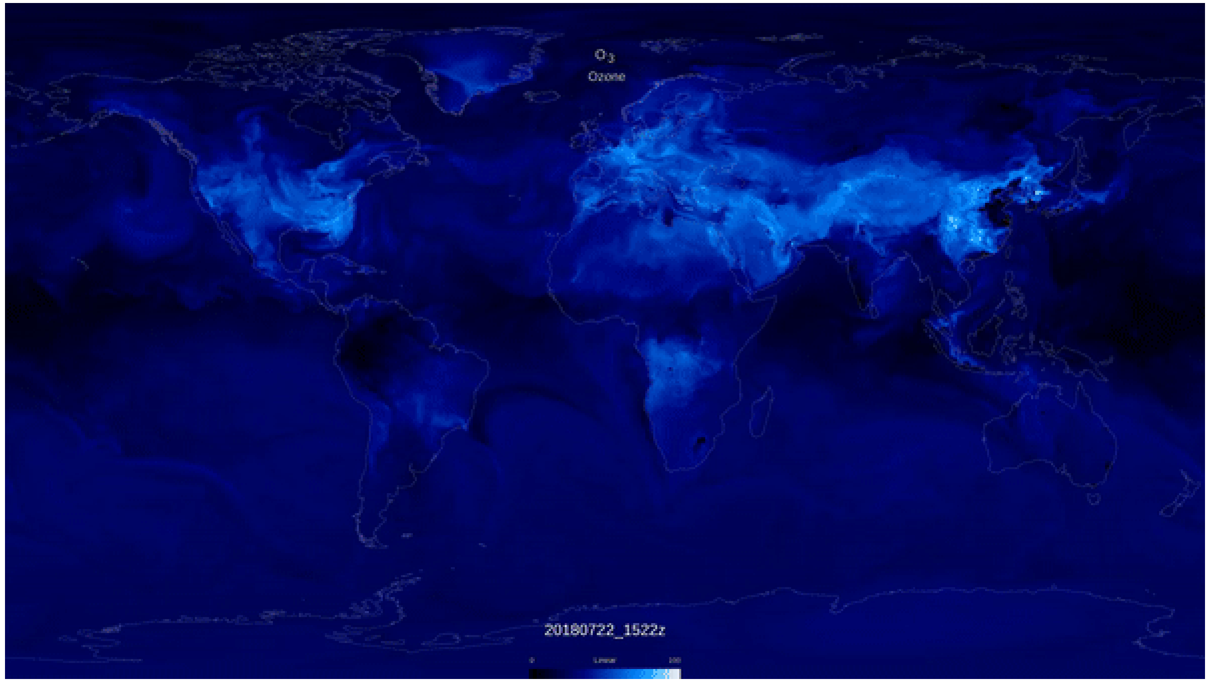


GEOS-CF

NASA Composition Forecasts (GEOS-CF)

https://gmao.gsfc.nasa.gov/weather_prediction/GEOS-CF/

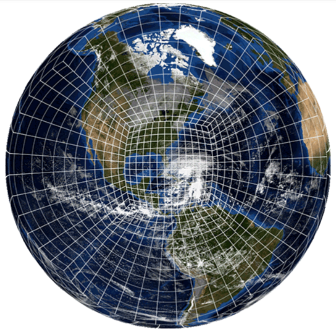
- The GEOS-Composition Forecast (CF) system forecasts trace gas and aerosol fields using constrained meteorology from GEOS and the GEOS-Chem chemical mechanism.
- Publicly available
- GEOS-Chem is a community-developed global 3-D model of atmospheric chemistry.
  - 250 chemical species
  - 725 chemical reactions
- Questions about GEOS-CF can be sent to geos-cf@lists.nasa.gov



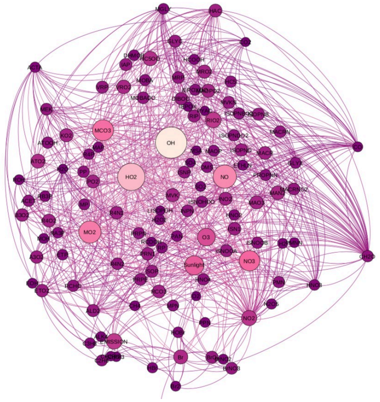
https://svs.gsfc.nasa.gov/4754

NASA Composition Forecasts (GEOS-CF)

GEOS Meteorology



GEOS-Chem Chemistry



GEOS-CF

- GEOS-Chem: Global chemistry transport model driven by GEOS meteorology
- 1-day simulation of the previous day using the analysis from FP-IT
  - Uses a **replay** technique to force the meteorology towards the FP-IT analysis
  - FP-IT is a 'frozen' version of FP used for satellite retrievals, similar to the version used to make MERRA-2.
- 5-day forecast
- Two aerosol schemes:
  - GOCART – Radiatively coupled to AGCM
  - GEOS-Chem – No feedbacks to model physics
- Full description in Keller et al., 2021

Emissions in GEOS-CF

Table 1
Emissions Used by GEOS-CF

Description	Reference	Comments
Offline inventories		
Anthropogenic NO, CO, black carbon (BC), organic carbon (OC), Ammonia (NH_3)	HTAP v2.2 (Janssens-Maenhout et al., 2015)	Global except Africa
Anthropogenic SO_2	OMI-HTAP (Liu et al., 2018)	Global except Africa
Anthropogenic VOCs	RETRO (Schultz et al., 2008)	Global except Africa
Anthropogenic NO, CO, SO_2, BC, OC, NH_3, VOCs	DICE-Africa (Marais and Wiedinmyer, 2016)	Africa
Arctic seabird NH_3	Croft et al. (2016)	
Volcanic SO_2	Carn (2019)	5% of the sulfur emitted as SO_4
Aircraft NO_x (=NO + NO_2), CO, SO_2, VOCs, BC, OC	AEIC (Stettler et al., 2011)	
Emissions calculated online based on real-time environment		
Biogenic VOCs	MEGAN v2.1 (Guenther et al., 2012)	
Biomass burning (wildfires) NO_x, CO, SO_2, VOCs, BC, OC	QFED v2.5 (Darmenov and da Silva, 2015)	35% emitted between 3.5 and 5.5 km altitude (Fischer et al., 2014).
Lightning NO_x	Murray et al., 2012	
Soil NO_x	Hudman et al., 2012	
Soil dust	Zender et al., 2003	
Sea salt aerosols	Gong, 2003; Jaeglé et al., 2011	
Oceanic DMS, CH_2O, C_3H_6O	Johnson, 2010; Nightingale et al., 2000	
Oceanic iodine	Carpenter et al., 2013	

Keller et al., 2021

GEOS Output Quick Guide

	GEOS FP	GEOS-CF
Type	Analysis + Forecast	Replay + Forecast
Domain	Global	Global
Spatial Resolution	Simulation: ~12 km Output: ~25 km (0.25°x0.312°)	~25 km (0.25°x0.312°)
Temporal Resolution	2-D data: Hourly 3-D data: Every 3 h	15 min, Hourly
Vertical Levels	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)
Initialization	Daily 10-day forecast at 00Z Daily 5-day forecast at 12Z	Daily 5-day forecast at 12Z
Data Assimilation	Yes	No
File Specification Doc	https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi1203.pdf *	https://gmao.gsfc.nasa.gov/pubs/docs/Knowland1204.pdf *

* Find most current File Specification at https://gmao.gsfc.nasa.gov/pubs/office_notes.php

GEOS-CF File Collections

- GEOS-CF output is organized into file *collections* that contain related variables.
- These have the form:

Group_Time_#Frequency_H_V

Group

Three letter abbreviation for the type of variables

- aqc= AQ relevant
- chm = chemistry
- htf = High-temporal freq
- met = meteorology
- xgc = extra chem fields

Time

- inst = instantaneous
- tavg = time-average

Frequency

Frequency or averaging time interval

- mn = minute
- hr = hour

H

Horizontal grid, hIxJ

- h = horizontal domain
- g (global) or r (regional)
- IxJ = horizontal resolution (# lon points, # lat points)

V

Vertical resolution, vL

- v = x, 2d fields
- v = p, pressure levels
- v = v, model levels
- L = # vertical levels

GEOS-CF File Names

- Each GEOS-CF file has the form:

GEOS-CF.version.mode.collection.timestamp.nc4

Version

File version (usually V01)

Mode

- rpl = replay
- fcst = forecast

Collection

See previous slide

Timestamp

Date and time of data file

- For assimilation: yyyymmdd_hhz
- For forecast: yyyymmdd_hhz+yyyymmdd_hhmmz

Assimilation cycle (yyyymmdd_hhz) — Forecast time (yyyymmdd_hhmmz)

All GEOS-CF output files are in NetCDF-4 format.

Example GEOS-CF File Name

GEOS-CF.v01.fcst.chm_tavg_1hr_g1440x721_v1. 20190309_12z+20190314_0730z.nc4

- GEOS-CF.v01.fcst - GEOS-CF forecast filename
- chem_tavg_1hr_g1440x721_v1 – Chemical species collection ("chm"), 1-hour time-averaged ("tavg_1hr") at the global ~0.25° horizontal resolution ("g1440x721") for single model layer data ("v1")
- 20190309_12z+20190314_0730z – Forecast initialized at 20190309 at12 Z. The valid time for the data in this file is 20190314 at 0730 Z, which represents the center point of a one-hour time-averaging period between 0700 and 0800 Z

AQ-Relevant Collections and Variables

https://gmao.gsfc.nasa.gov/pubs/docs/Knowland1204.pdf *

Collection Name	Description
htf_inst_15mn_g1440x721_x1	High Temporal Frequency Chemistry and Meteorology *Surface CO, NO_2, O_3, SO_2, $PM_{2.5}$ (GCC & GOCART), and meteorology (RH, T, P, etc)*
aqc_tavg_1hr_g1440x721_v1	Air Quality Concentrations *One-hour time-averaged surface CO, NO_2, O_3, $PM_{2.5}$(GCC), SO_2*
chm_tavg_1hr_g1440x721_v1	Chemistry Fields *One-hour time-averaged surface mixing ratios of many chemical species and speciated $PM_{2.5}$ (GCC)*
xgc_tavg_1hr_g1440x721_x1	Extra GEOS-Chem Fields *One-hour time-averaged AOD, column quantities, and removal processes (deposition)*
chm_inst_1hr_g1440x721_p23	Chemistry Fields *3D (23 pressure levels,1000 to 10 hPa) instantaneous CO, NO_2, O_3, $PM_{2.5}$(GCC, speciated), SO_2*

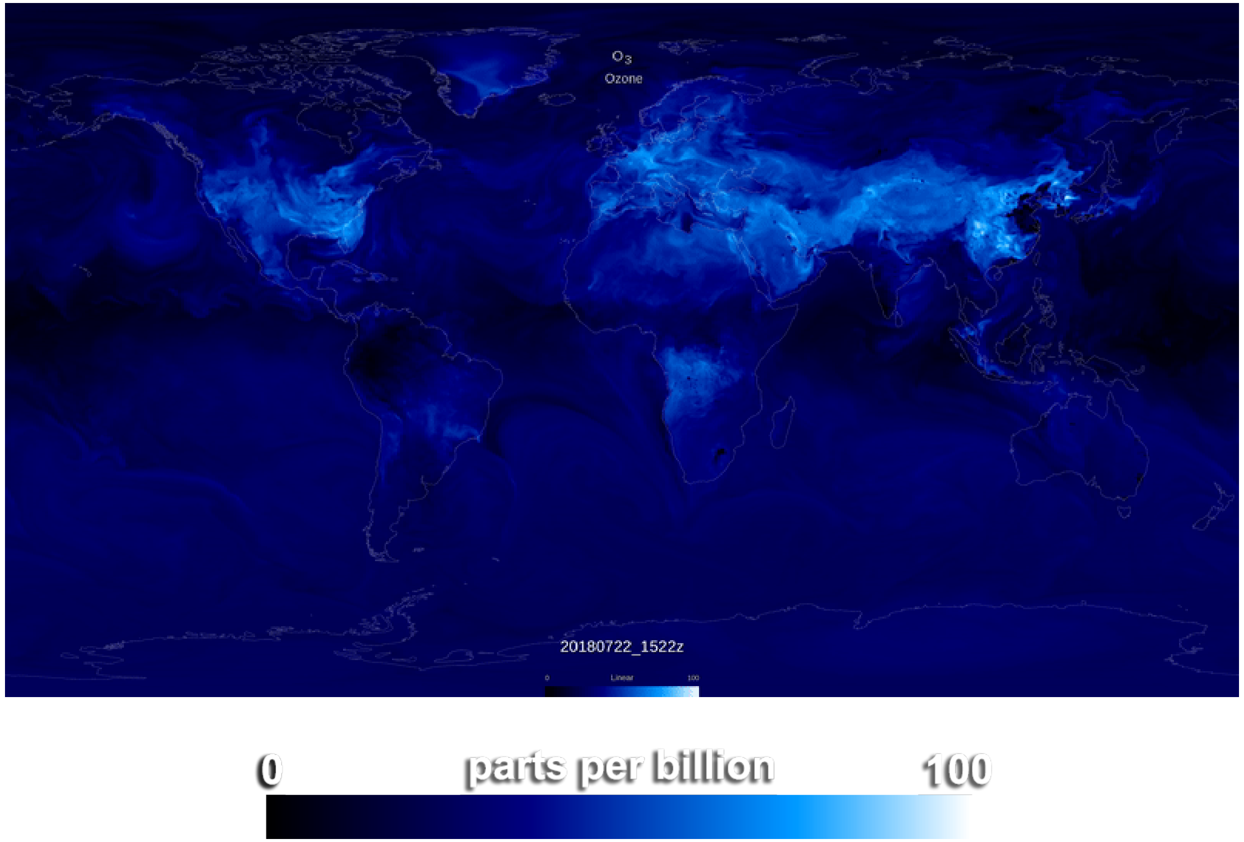
There are two $PM_{2.5}$ variables:

- PM25_RH35_GCC: $PM_{2.5}$ at 35% RH from GEOS-Chem (GCC)
- PM25_RH35_GOCART: $PM_{2.5}$ at 35% RH from GOCART (only available in htf)

* Find most current File Specification at https://gmao.gsfc.nasa.gov/pubs/office_notes.php

Case Study: High Resolution Simulation of Ozone

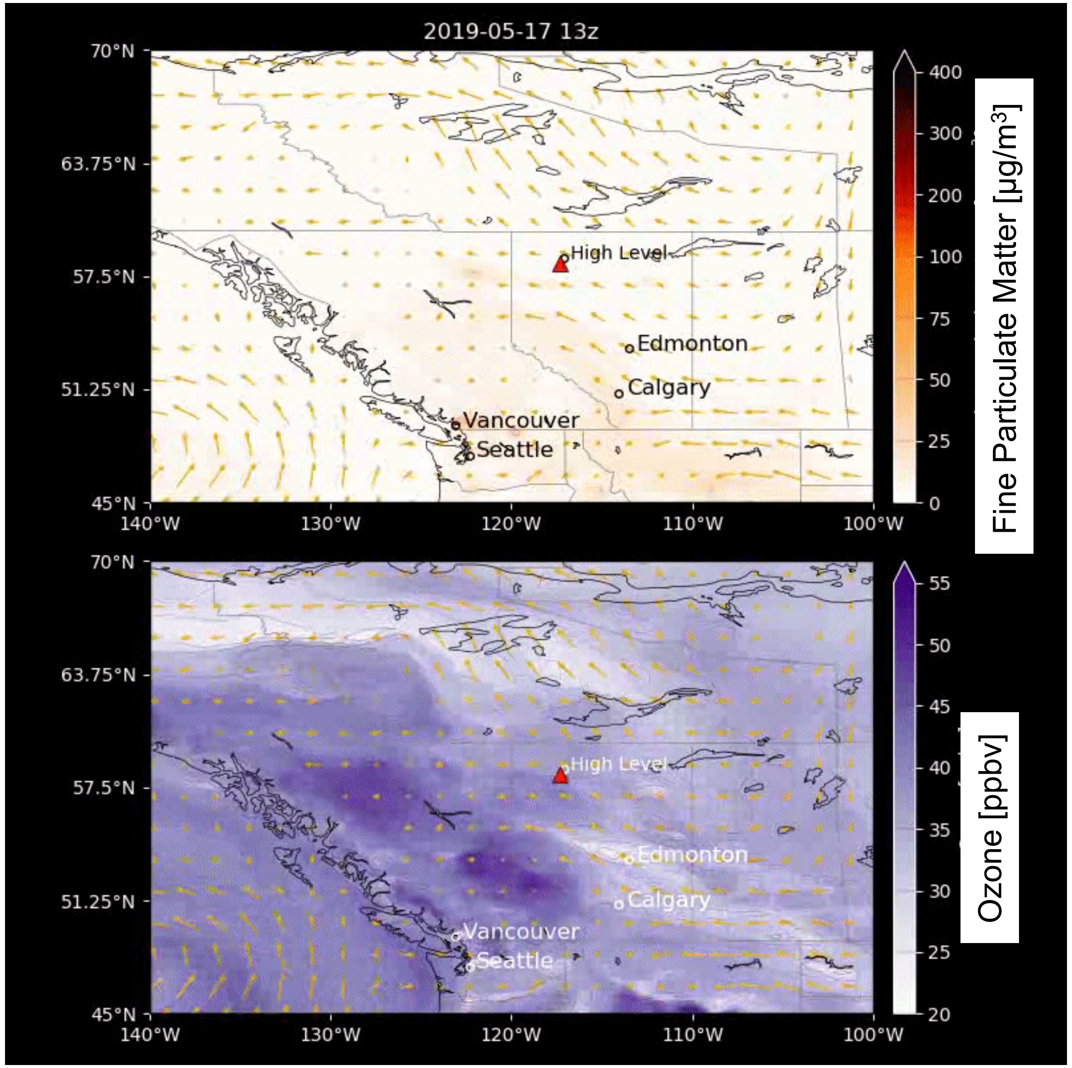
- Surface O_3 from GEOS-CF during the summer of 2018
- O_3 is a pollutant produced and destroyed through interactions of various chemical species such as nitrogen oxides (NO_2, NO) and volatile organic compounds (VOCs).
- Forecasted concentrations of pollutants like O_3, NO_2, and $PM_{2.5}$ can be combined to calculate air quality indices.



https://svs.gsfc.nasa.gov/4764

Case Study: O_3 and $PM_{2.5}$ Forecasts During 2019 Canadian Wildfires

- Smoke from fires led to poor air quality in the Northwest US.
- Within the smoke plumes, high concentrations of NO consume O_3, leading to lower O_3 levels (lighter colors) near the source.
- As the plume mixes with surrounding air, O_3 is produced, leading to increased concentrations near the plume edge.
- O_3 produced in wildfire plumes can be comparable to urban pollution levels.



Fire Detection

https://gmao.gsfc.nasa.gov/research/science_snapshots/2019/air_pollution_GEOS-CF.php

GEOS-CF Evaluation

- Keller et al. (2021) includes a detailed evaluation of GEOS-CF
- Variety of observations used for evaluation
  - Sondes
  - Ground based in-situ
  - Ground based remote sensing
  - Satellite
- Evaluation of both replay and forecast skill included
- Aerosol evaluation of GEOS-Chem aerosols

Table 2
Overview of Observation Data Sets Used for GEOS-CF Model Validation

Description	Species	# Sites	Reference
Ozonesonde	O_3	24	Tilmes et al., 2012; Thompson et al., 2017
NASA OMI NO_2 v4.0	Tropospheric column NO_2	global	Lamsal et al., 2021
MOPITT v8	Total column CO	global	Deeter et al., 2019
MODIS	AOD at 550 nm	Global	Remer et al., 2005; Levy et al., 2010, 2015
AERONET	AOD at 550 nm	195	Giles et al., 2019
GAW WDCGG	CO	54	https://gaw.kishou.go.jp/
GAW WDCRG	O_3, NO_2, SO_2	48 (O_3), 6 (NO_2), 9 (SO_2)	https://www.gaw-wdcrg.org/
OpenAQ	O_3, NO_2, SO_2, $PM_{2.5}$	3151 (O_3), 2789 (NO_2), 1221 (SO_2), 2667 ($PM_{2.5}$)	https://openaq.org

Table 2 from Keller et al., 2021

GEOS-CF Replay Evaluation - AOD

- Comparisons with MODIS Aqua and AERONET AOD
- GEOS-CF shows minimal bias in background regions.
- Overall high bias in AOD, but spatial and seasonal patterns captured
- Overprediction of AOD likely a result of outdated SO_2 emissions inventory, and underestimation of removal processes for nitrate

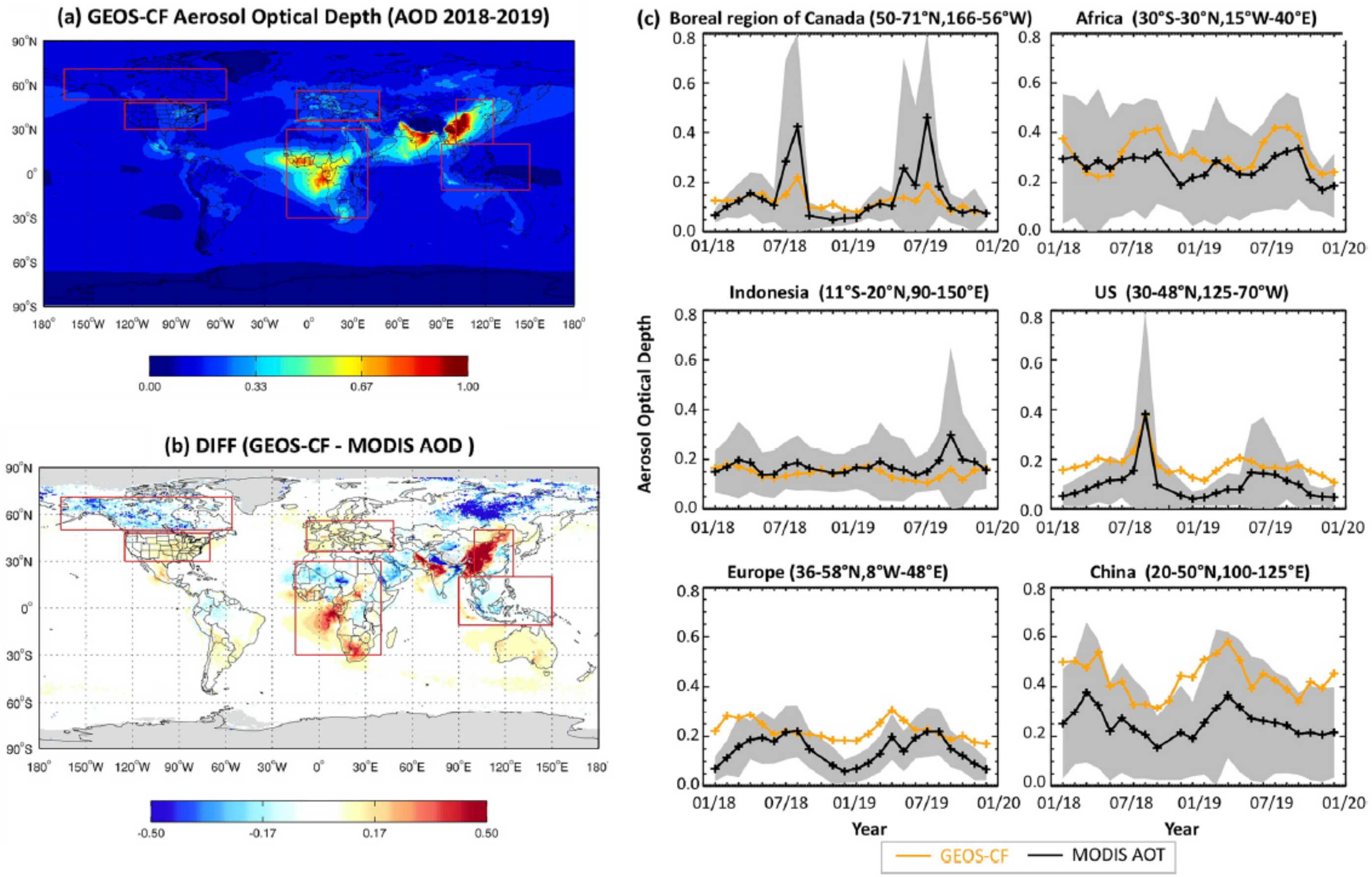


Figure 15 from Keller et al., 2021

GEOS-CF Replay Evaluation – $PM_{2.5}$

- Comparison with ~2600 surface $PM_{2.5}$ observations from OpenAQ
- Similar results as comparison with AOD
- Even though GEOS-CF has high resolution for a global forecast model, there are still biases when comparing with station-level observations, especially in urban areas

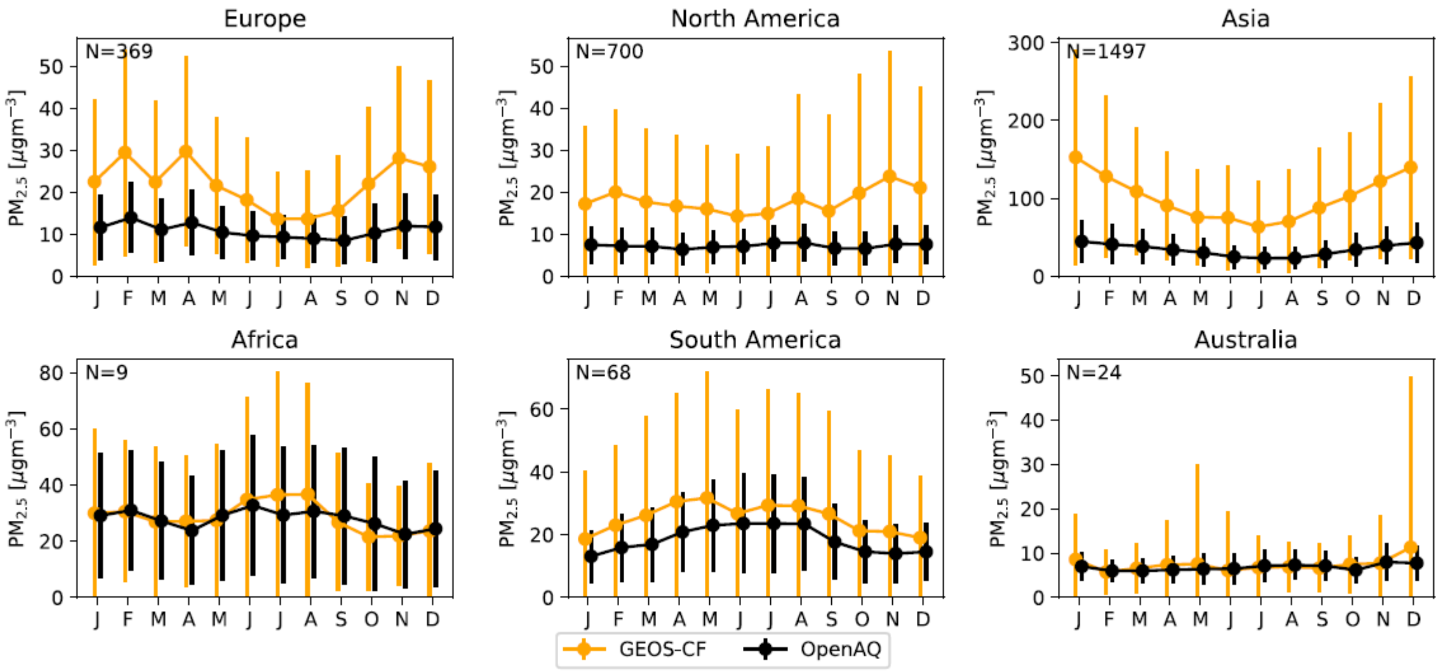


Figure 17 from Keller et al., 2021

GEOS-CF Forecast Evaluation

- 5-day forecasts have comparable normalized mean bias (NMB) and normalized root mean square error (NRMSE) to the 1-day replay.

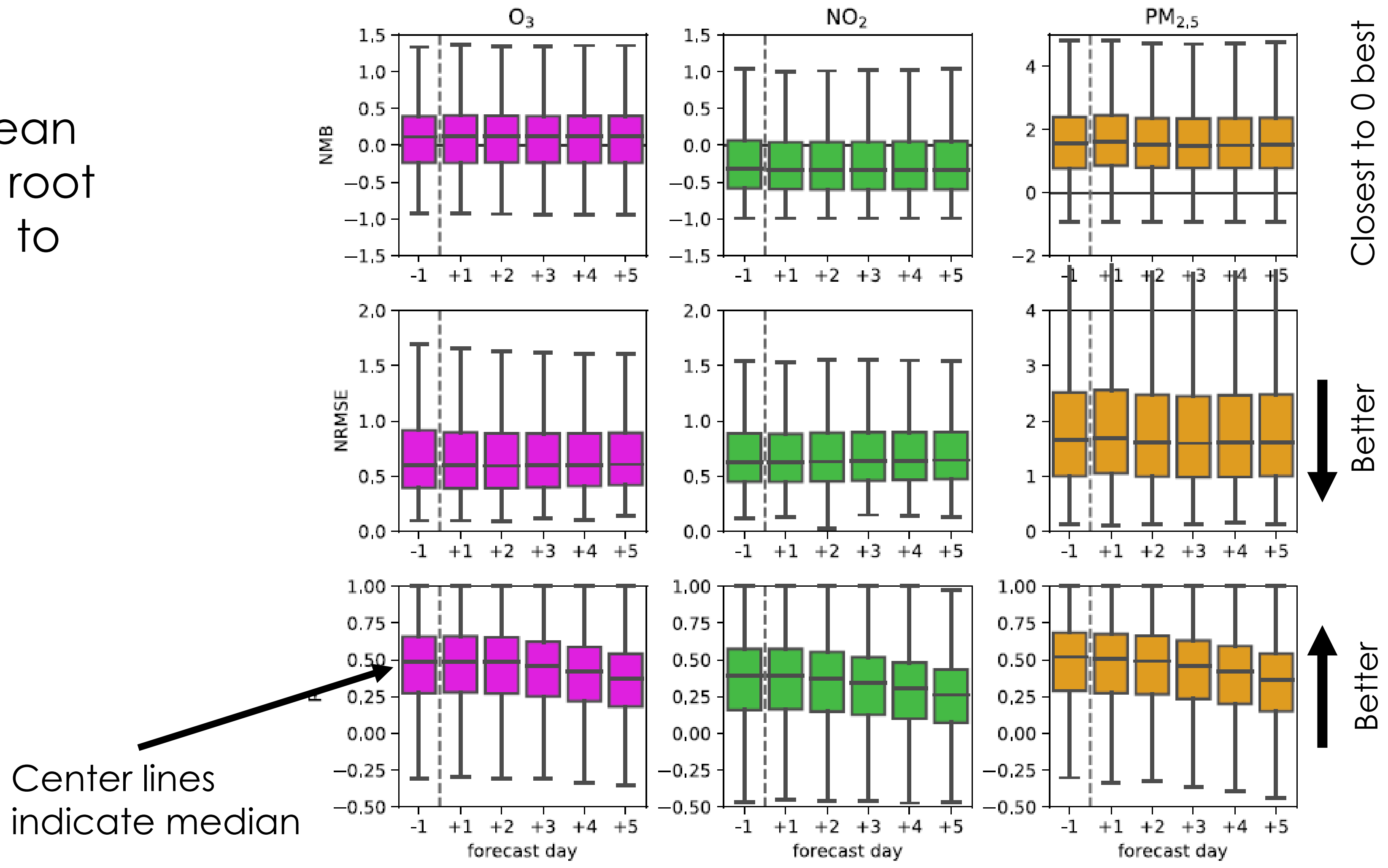
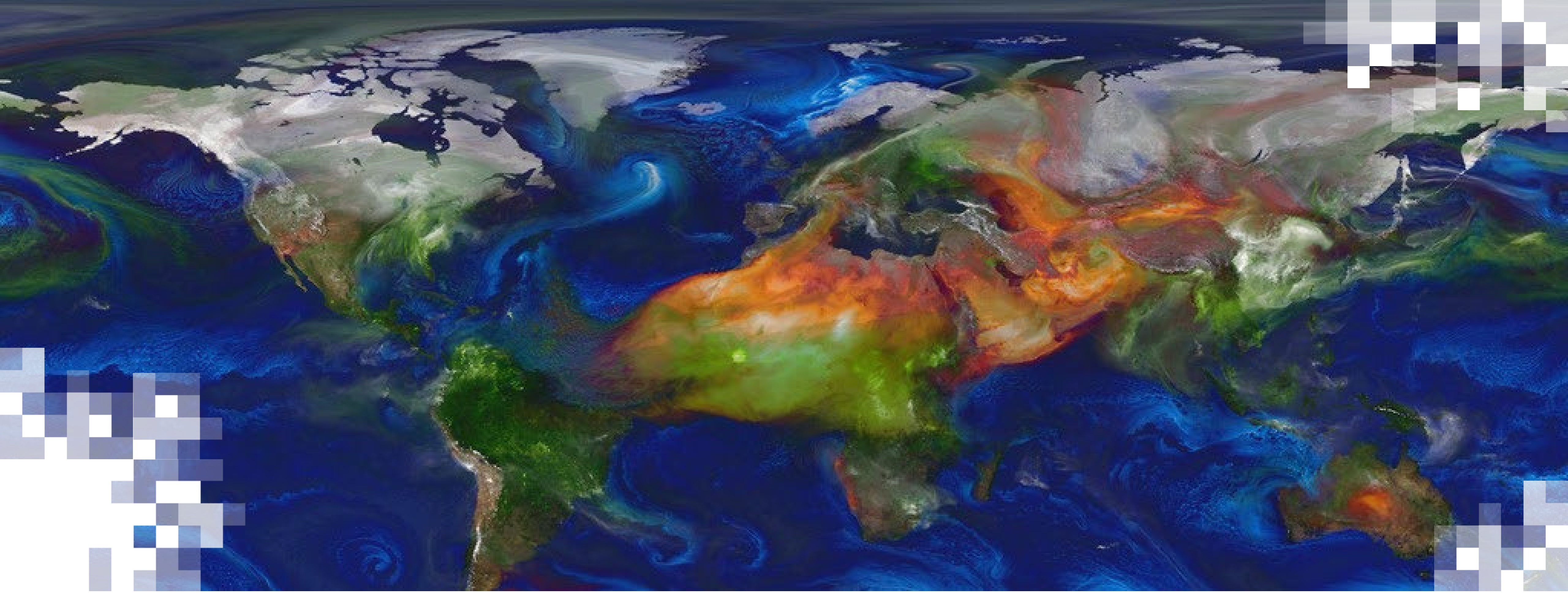


Figure 18 from Keller et al., 2021



MERRA-2

What is reanalysis, and why do we do it?

What:

- A consistent reprocessing of Earth system observations using a modern, unchanging data assimilation system
- Relies on models to interpret, relate, and combine different observations from multiple sources
- Successful reanalysis ***requires*** a good forecast model combined with bias-corrected and quality-controlled observations

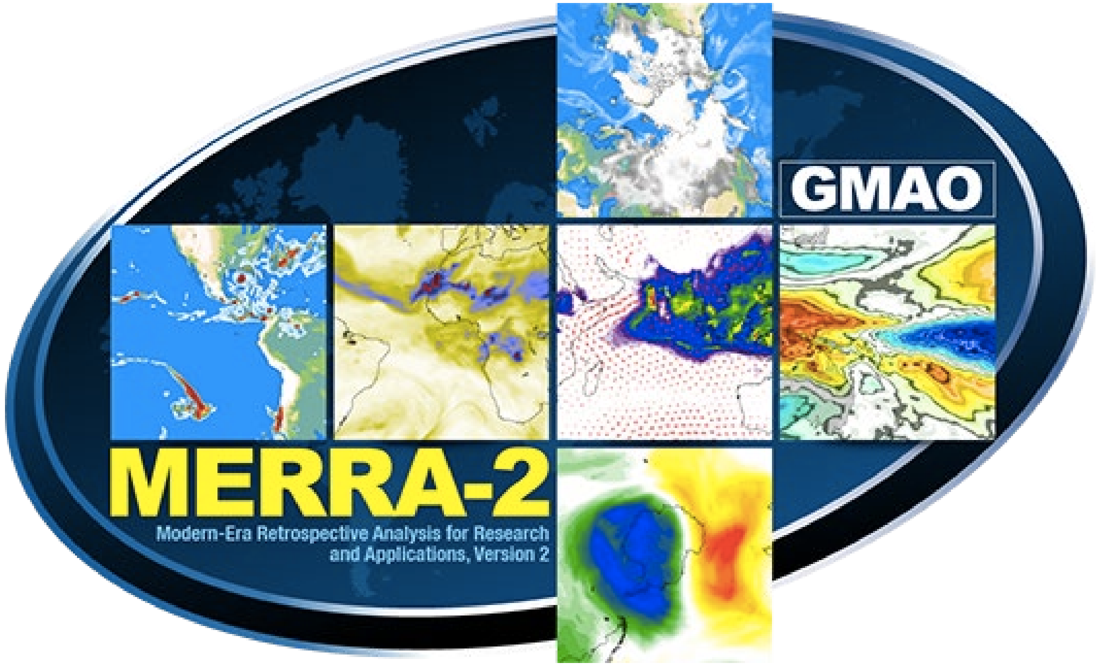
Why:

- Produces multi-decadal, gridded datasets that estimate a large variety of Earth system variables, including ones that are not directly observed
- Has become fundamental to research and education in the Earth sciences

MERRA-2 Reanalysis

https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/

- The **M**odern-**E**ra **R**etrospective analysis for **R**esearch and **A**pplications version **2** (MERRA-2) provides data beginning in 1980 and runs a few weeks behind real-time.
- Long-term, model-based analyses of multiple datasets using a fixed assimilation system
- Includes meteorology, stratospheric ozone, and aerosols at the spatial resolution of a 0.5° × 0.66° (~50 km) grid.



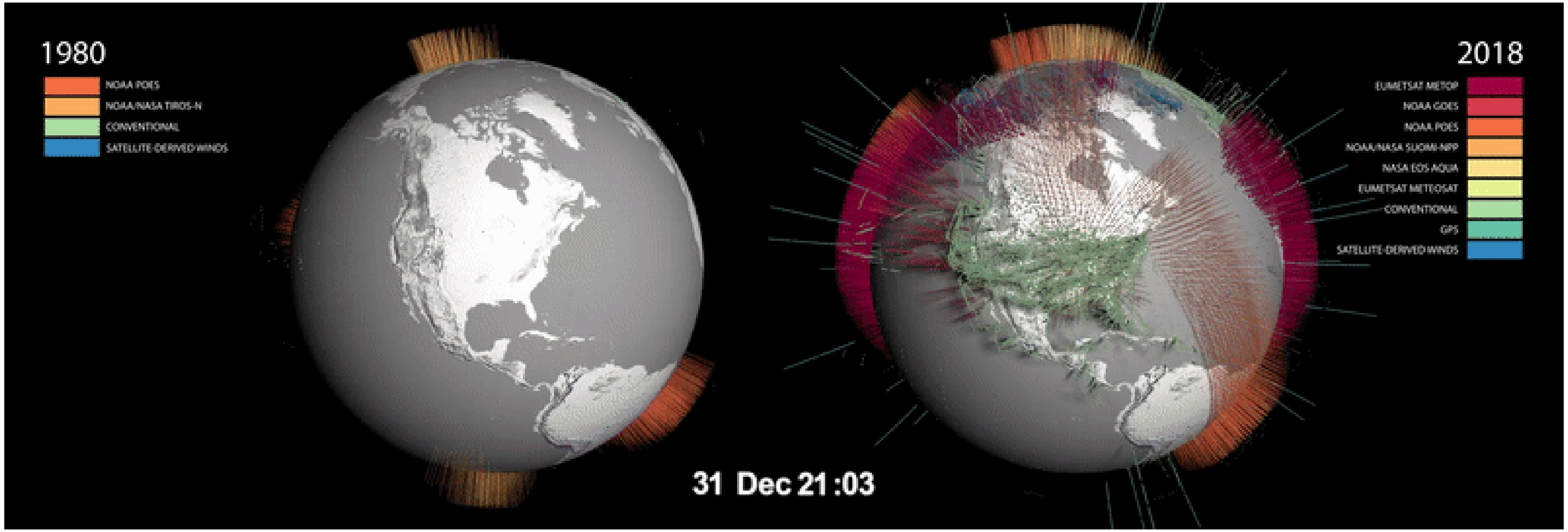
Source: https://gmao.gsfc.nasa.gov/reanalysis/

GEOS Output Quick Guide

	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 *

* Find most current File Specification at https://gmao.gsfc.nasa.gov/pubs/office_notes.php

Observing System in MERRA-2



In 1980, there were few satellites providing observations. These satellites, with global surface and upper-air observations were the first observations used for the beginning of MERRA-2 in 1980. Every 6 hours, a median number of 175,000 observations were assimilated.

Today, our observing system has advanced significantly, and MERRA-2 assimilates a median number of 5 million observations every 6 hours.

https://svs.gsfc.nasa.gov/4654

MERRA-2 Aerosol Observations

- Aerosol assimilation is described in detail in Randles et al. 2017 and https://gmao.gsfc.nasa.gov/pubs/docs/Randles887.pdf.
- In MERRA-2, AOD at 550 nm is assimilated.
- Some notes:
  - No information on vertical structure or composition
  - Daylight observations only
  - Subject to meteorological conditions (e.g., clouds) and viewing geometry (e.g., sun glint)
  - When there are no observations, MERRA-2 draws towards the GEOS/GOCART simulation.

Sensor	Temporal coverage	Description
AVHRR NNR	1980–August 2002	PATMOS-x radiances over ocean only (PM orbit)
AERONET	Station dependent (1999–October 2014)	AOD from land station network
MISR	February 2000–June 2014	AOD over bright land surfaces only (albedo > 0.15)
MODIS *Terra* NNR	March 2000 onward (NRT)	Collection 5 "Dark Target" land and ocean radiances (AM orbit)
MODIS *Aqua* NNR	August 2002 onward (NRT)	Collection 5 "Dark Target" land and ocean radiances (PM orbit)

Table 2 from Randles et al. 2017

MERRA-2 Aerosol Observations

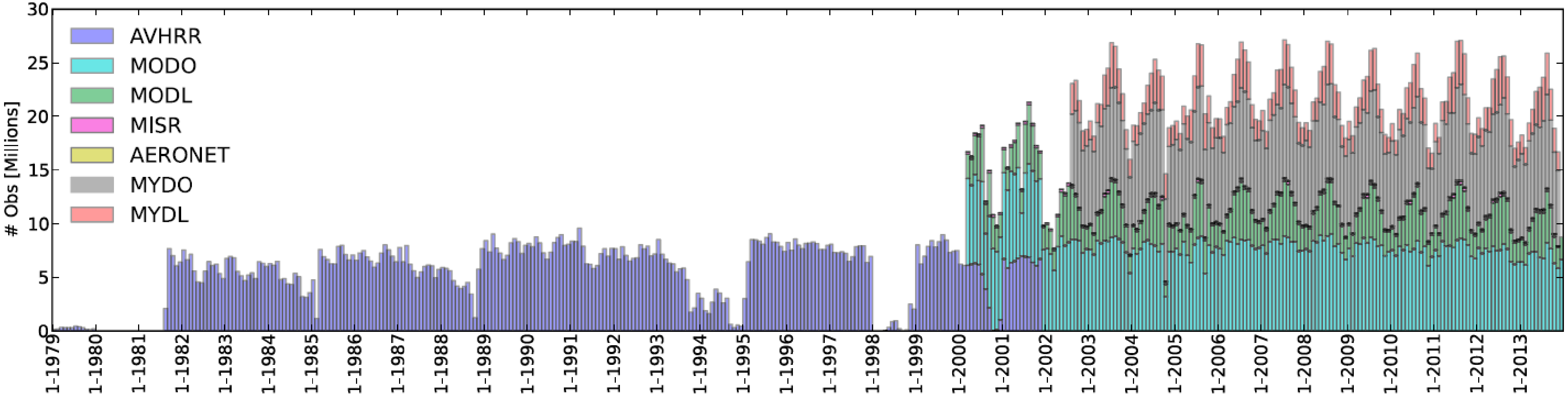
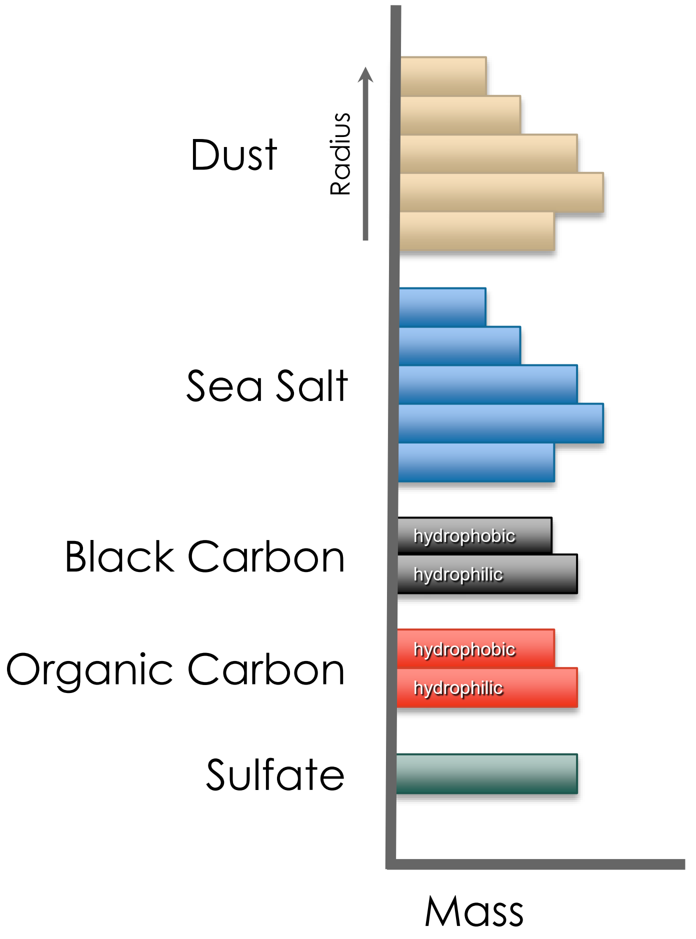


Figure 3 from Randles et al. 2017

- When using MERRA-2 products, one must take care to consider the changing observing system over time.

GOCART in MERRA-2

- Goddard Chemistry, Aerosol, Radiation and Transport Model (GOCART, Chin et al. 2002, Colarco et al. 2010)
- Sources and sinks for 5 non-interactive species
- Radiatively active



Wind and topographic sources, 5 mass bins

Wind-driven source, 5 mass bins

Anthropogenic and wildfire sources, mass hydrophobic & hydrophilic

Anthropogenic, biogenic, and fire sources, mass hydrophobic and hydrophilic

Anthropogenic and wildfire sources of SO_2, oxidation to SO_4 mass

There are no nitrate aerosols in MERRA-2.

Emissions in MERRA-2

Aerosol type	Source	Description
Dust	Wind-driven	Static topographic depression source map (0.3125° × 0.25°; Ginoux et al. 2001)
Sea salt	Wind-driven	See section 2b
Volcanic SO_2	AeroCom Phase II (HCA0 v2; Diehl et al. 2012)	Daily degassing and eruptive volcanos (1980–2010)
Biogenic terpene	Guenther et al. (1995)	Monthly mean climatology (2° × 2.5°)
Dimethyl sulfide (DMS) and methanesulfonic acid (MSA)	Lana et al. (2011)	Monthly mean climatology (1° × 1°)
Biomass burning	Scaled RETROv2 (Duncan et al. 2003)	Monthly mean varying (1980–96; 0.3125° × 0.25°)
SO_2, SO_4, POM, and BC	Scaled GFEDv3.1 (Randerson et al. 2006)	Monthly mean varying (1997–2009; 0.3125° × 0.25°)
	QFED 2.4-r6	Daily (2010 onward; 0.3125° × 0.25°)
Anthropogenic SO_2	EDGARv4.2 (energy + non-energy) (European Comission 2011)	Annually varying (1980–2008; 0.1° × 0.1°)
Anthropogenic SO_4, BC, and POM	AeroCom Phase II (HCA0 v1; Diehl et al. 2012)	Annually varying (1980–2006; 1° × 1°)
International ships (SO_2)	EDGARv4.1 (European Commission 2010)	Annually varying (1980–2005; 1° × 1°)
International ships (SO_4, POM, BC)	AeroCom Phase II (HCA0 v1; Diehl et al. 2012) and Eyring et al. (2005)	Annually varying (1980–2007; 1° × 1°)
Aircraft (SO_2)	AeroCom Phase II (HCA0 v1; Diehl et al. 2012)	Monthly varying (1980–2006; 1° × 1.25° × 72 levels)

Randles et al., 2017

MERRA-2 File Collections

- MERRA-2 outputs are organized into file *collections* that contain related variables.
- These have the form:

Frequency_Dimensions_Group_HV

Frequency

Frequency or averaging interval

- const = time-independent
- inst = instantaneous
- tavg = time-average
- Stat = statistics

Can be 1, 3, 6-hourly, daily (D), monthly (M), or a monthly-diurnal mean (U)

Dimensions

Dimensions of variables

- 2d = only 2d fields
- 3d = can have 2d and 3d

Group

Three letter abbreviation for the type of variables

- These are also used in the short name
- Ex. aer = Aerosol fields
- See documentation for full list

HV

Horizontal and vertical grid

- H = typically N, for nominal grid
- V = x, horizontal only
- V = p, pressure level
- V = v, model level
- V = e, model layer edges

MERRA-2 File Names

- Each MERRA-2 file has the form:

MERRA2_SVv.collection.timestamp.nc4

Stream and Version

File version (usually 100, 200, 300, or 400)

Collection

See previous slide

All MERRA-2 output files are in NetCDF-4 format.

Timestamp

Date and time of data file

- For instantaneous or time-averaged files: yyyymmdd
- For monthly files: yyyymm

For collections with instantaneous or time-averaging frequency < 1 day, the daily file will contain all of the timesteps

Example MERRA-2 File Name

MERRA2_400.tavgM_2d_aer_Nx.202106.nc4

- MERRA2_400 – MERRA-2 file from fourth assimilation stream
- tavgM_2d_aer_Nx – 2D monthly time-averaged (“tavgM_2d”) aerosol species collection (“aer”) on the horizontal grid (“Nx”)
- 202106 – This file contains monthly averages for June 2021

AQ-Relevant Collections and Variables

https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf *

Collection Name	Description
tavg1_2d_aer_Nx	Aerosol Diagnostics *Optical properties (Extinction AOT, Scattering AOT, Angstrom parameter)* *Surface concentration (kg/m^3), Column Density (kg/m^2)*
tavg1_2d_adg_Nx	Aerosol Diagnostics (extended) *Emissions and removal processes (deposition, sedimentation, and scavenging)*
tavg1_2d_chm_Nx	2D time-averaged chemistry diagnostics *Surface CO, column CO, emissions, chemical loss, chemical production, total column O_3*
inst3_3d_aer_Nv	3D instantaneous aerosol diagnostics *Mass mixing ratios (kg/kg) of aerosol species in each size bin*
inst3_3d_chm_Nv	3D instantaneous chemistry diagnostics *CO molar mixing ratio (mol/mol), O_3 (not for use in scientific analysis)*

To calculate $PM_{2.5}$, use the formula:

$PM_{2.5}$=[DUSMASS25]+[SSSMASS25]+[BCSMASS]+1.4×[OCSMASS]+1.375×[SO4SMASS]

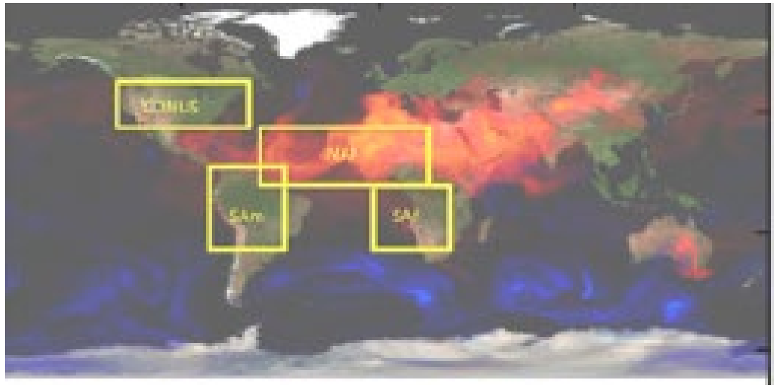
These variables are contained in the *tavg1_2d_aer_Nx* collection

* Find most current File Specification at https://gmao.gsfc.nasa.gov/pubs/office_notes.php

MERRA-2 Evaluation

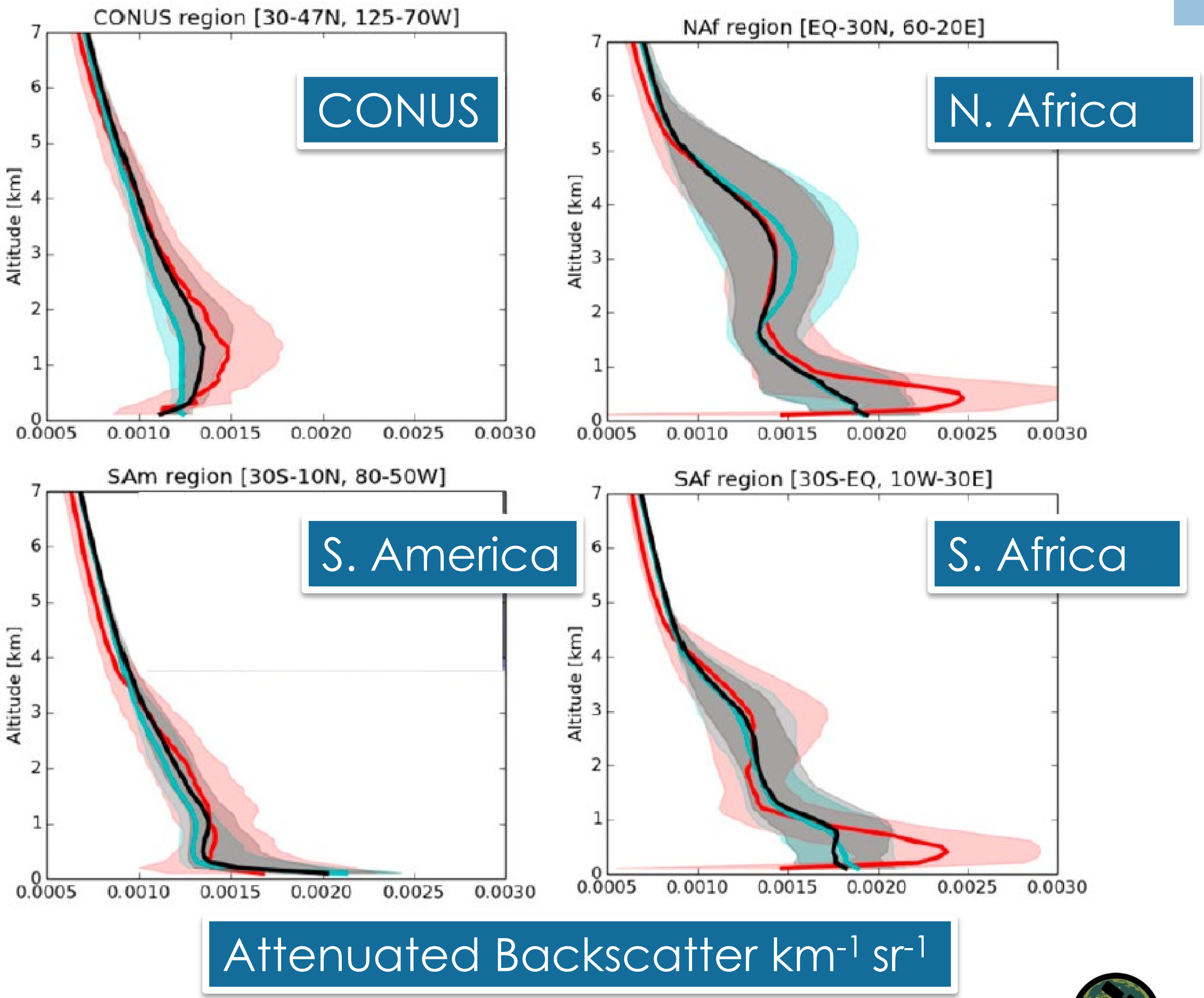
- In order to evaluate the performance of a reanalysis, it is important to compare the output with *independent* sources of data (i.e., those not used for assimilation).
- A detailed evaluation of MERRA-2 aerosols can be found in Buchard et al. (2017).
  - Optical properties, vertical distribution, and surface $PM_{2.5}$

MERRA-2 Evaluation: Vertical Structure



CALIOP observations
Model without AOD assimilation
MERRA-2

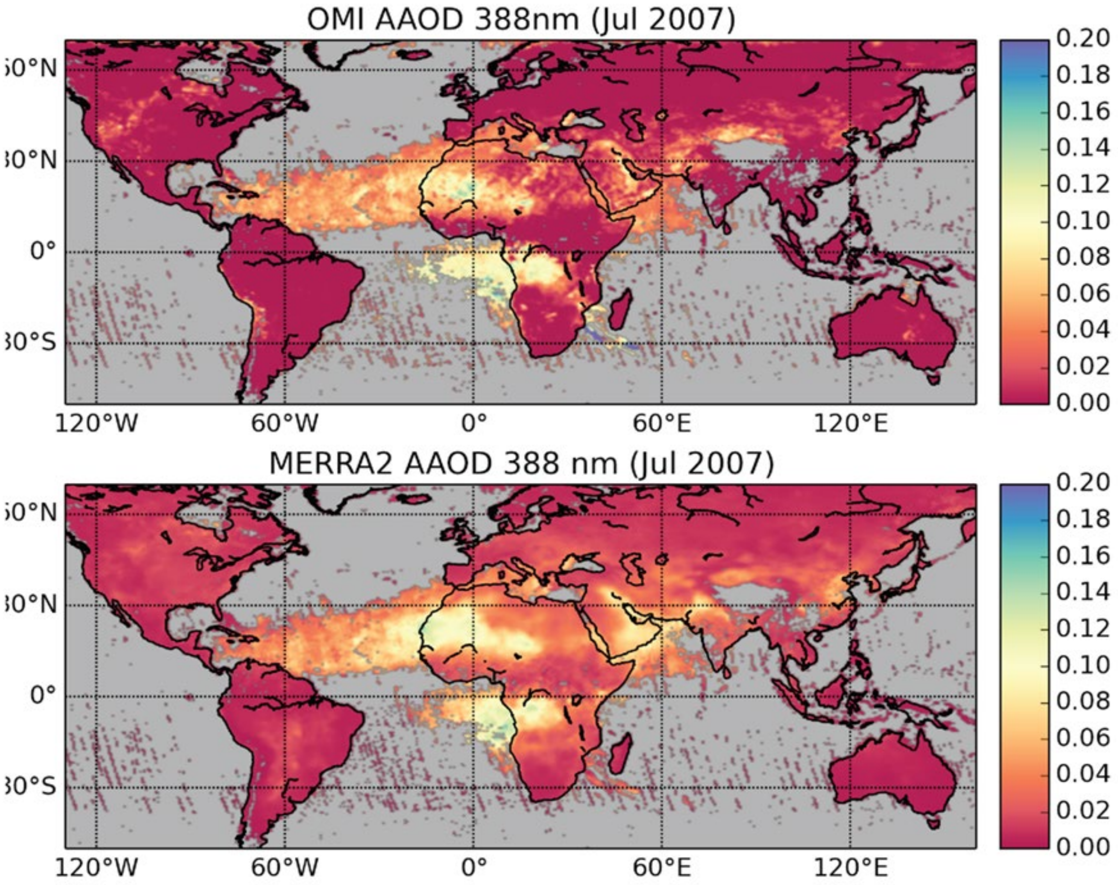
Assimilating AOD improves the vertical distribution of aerosols with respect to daytime lidar observations



Adapted from Figure 2 of Buchard et al. 2017

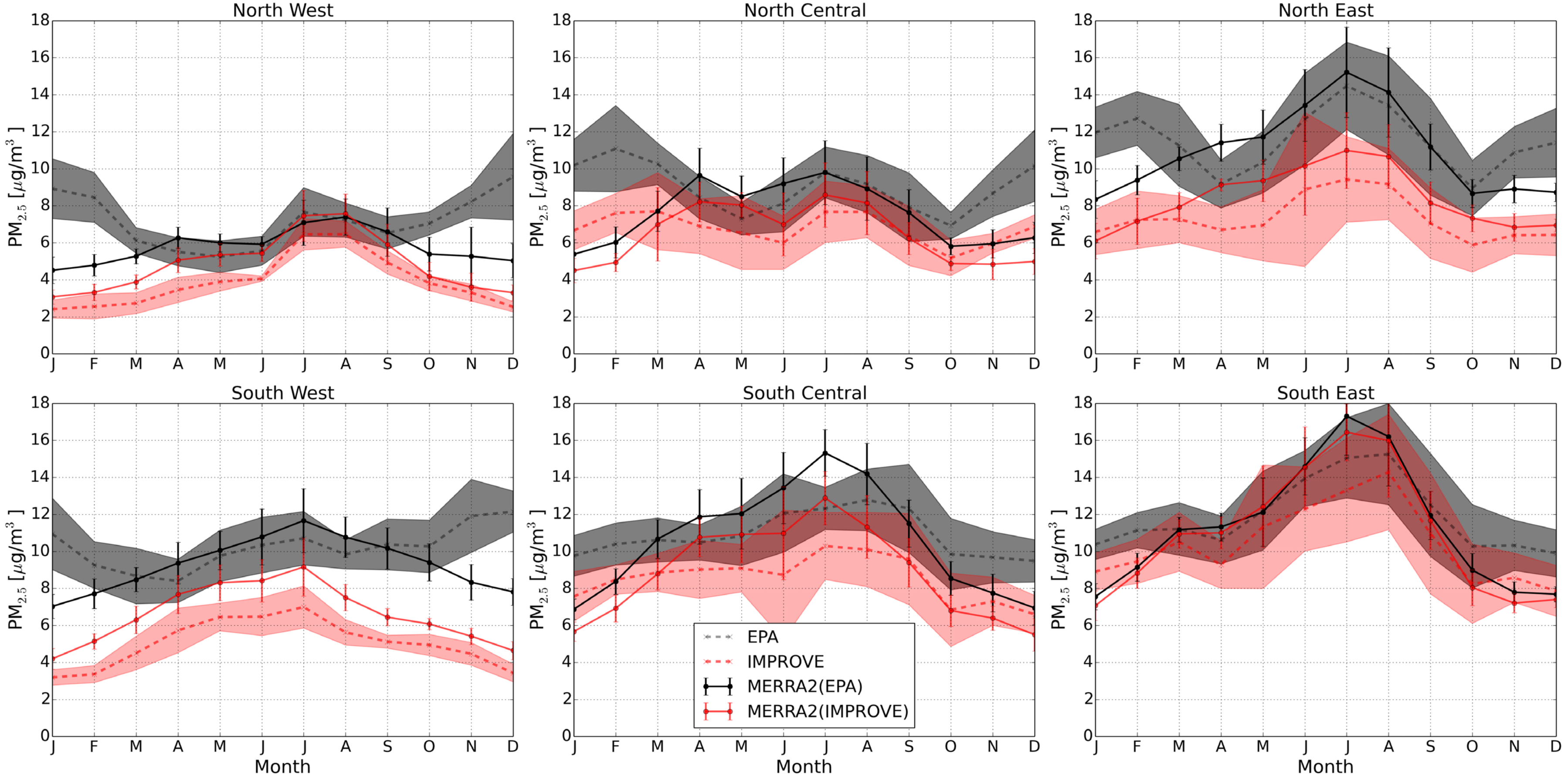
MERRA-2 Evaluation: Aerosol Absorption

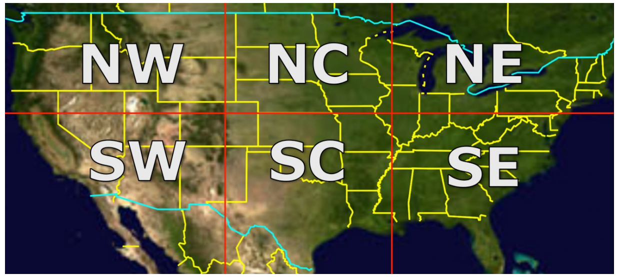
- Comparison of MERRA-2 Absorption Optical Depth (AAOD) with OMI retrievals
- Good agreement for African dust and smoke
- North American biomass burning underestimated according to OMI

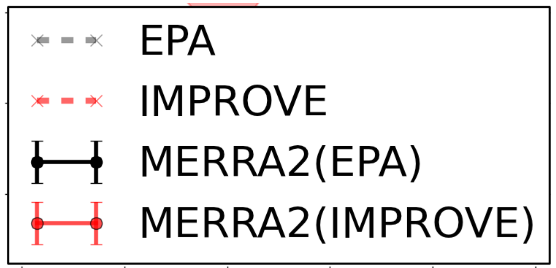


Adapted from Figure 1 of Buchard et al. 2017

MERRA-2 Evaluation: Regional $PM_{2.5}$ Climatology

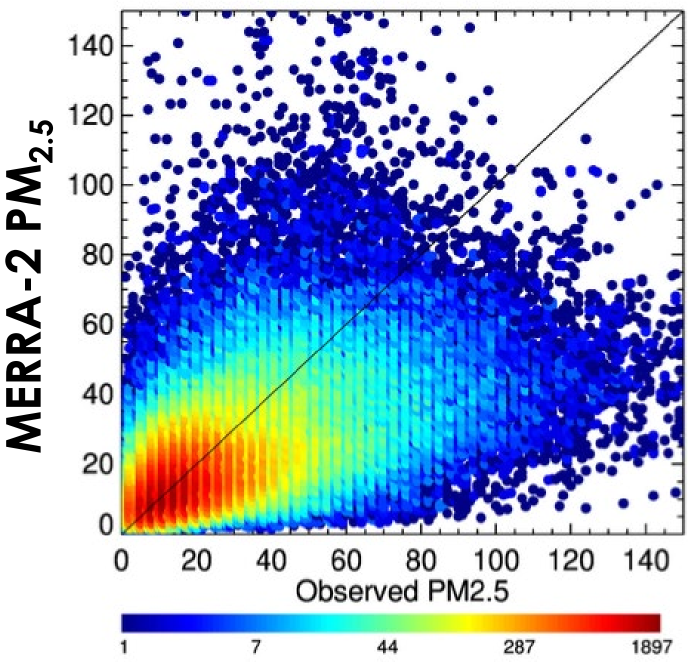




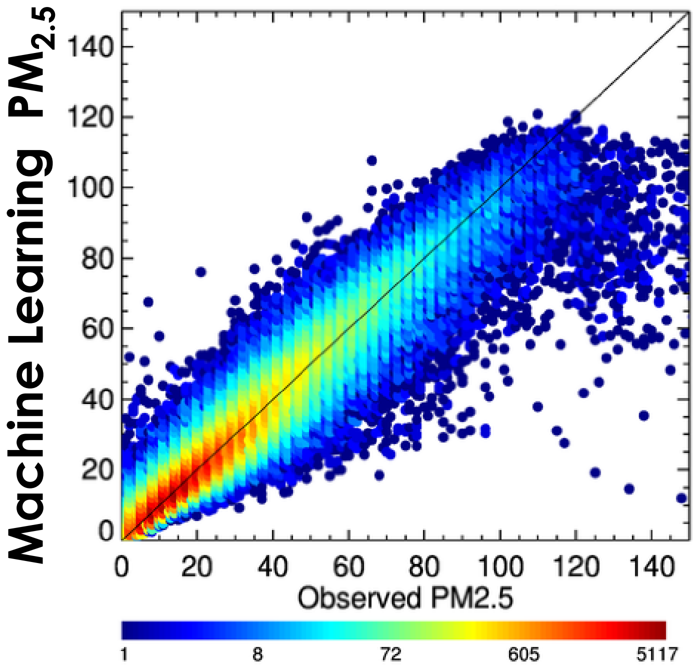


Case Study: MERRA-2 as a Dataset for Machine Learning

- In a very recent study, Gupta et al. (2021) use MERRA-2 output, along with ground observations of PM2.5 to train a machine learning model to predict $PM_{2.5}$ in Thailand
- The machine learning predicted PM2.5 shows better correlation and reduced bias with respect to observations
- This algorithm can be used to bias correct the entire MERRA-2 time period, creating a more accurate long-term dataset for this region



R = 0.53
Mean Bias = -3.8
RMSE = 16.7

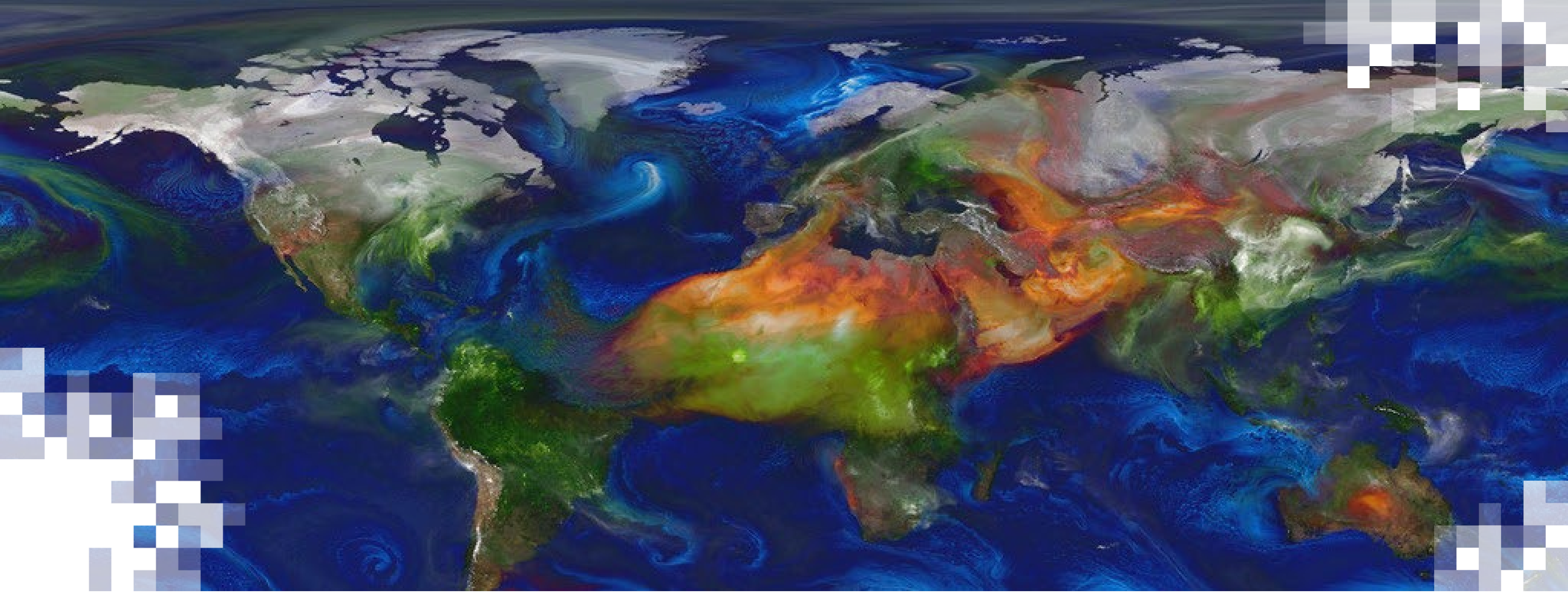


R = 0.95
Mean Bias = 0.03
RMSE = 5.9

Gupta et al. (2021) Machine Learning Algorithm for Estimating Surface $PM_{2.5}$ in Thailand, https://doi.org/10.4209/aaqr.210105

References

- GOCART
  - Chin, M., P. Ginoux, S. Kinne, O. Torres, B. Holben, B. Duncan, R. Martin, J. Logan, A. Higurashi, and T. Nakajima (2002), Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements, J. Atmos. Sci., 59(3), 461–483. https://doi.org/10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2
  - Colarco, P., A. Da Silva, M. Chin, and T. Diehl (2010), Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth, J. Geophys. Res.-Atmos., 115, –, doi:10.1029/2009JD012820. https://doi.org/10.1029/2009JD012820
  - Liu, F. et al. (2018). A new global anthropogenic SO2 emission inventory for the last decade: A mosaic of satellite-derived and bottom-up emissions - https://doi.org/10.5194/acp-18-16571-2018
  - Janssens-Maenhout, G. et al. (2015). HTAP_v2.2: A mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution - https://doi.org/10.5194/acp-15-11411-2015
  - Ginoux, P. et al (2001) Sources and global distributions of dust aerosols simulated with the GOCART model, J. Geophys. Res., https://doi.org/10.1029/2000JD000053
  - Carn, S. (2019). Multi-satellite volcanic sulfur dioxide L4 long-term global database V3. https://doi.org/10.5067/measures/so2/data404
- GEOS-CF
  - Keller et al. (2021) Description of the NASA GEOS Composition Forecast Modeling System GEOS-CF v1.0, https://doi.org/10.1029/2020MS002413
  - File Specification Document – https://gmao.gsfc.nasa.gov/pubs/docs/Knowland1204.pdf
- GEOS FP
  - Reinecker et al. (2008), The GEOS-5 Data Assimilation System Documentation of Versions 5.0.1, 5.1.0, and 5.2.0,
  - File Specification Document - https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi1203.pdf
- MERRA-2 - https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/docs/
  - File Specification Document - https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf
  - Aerosol Assimilation – Technical Document - https://gmao.gsfc.nasa.gov/pubs/docs/Randles887.pdf
  - Randles et al. (2017) The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part I: System Description and Data Assimilation Evaluation - https://journals.ametsoc.org/view/journals/clim/30/17/jcli-d-16-0609.1.xml
  - Buchard et al. (2017) The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part II: Evaluation and Case Studies - https://journals.ametsoc.org/view/journals/clim/30/17/jcli-d-16-0613.1.xml



GMAO Fluid Website Demo

Thank You!