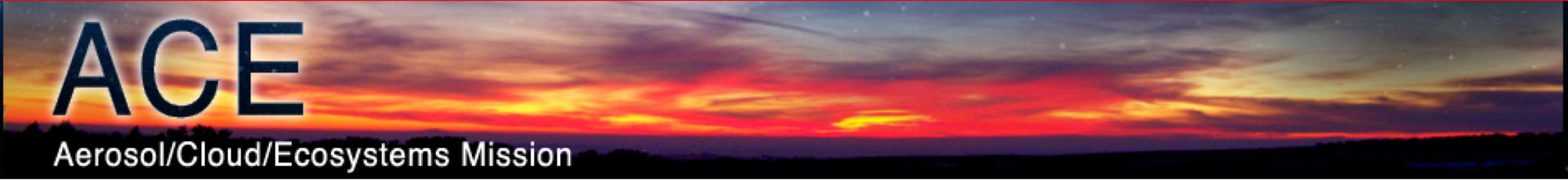


Observing System Simulation Experiments (OSSEs)

Relevance to ACE and A-CCP

Arlindo da Silva

Global Modeling and Assimilation Office

NASA Goddard Space Flight Center

With contributions from W. Putman, P. Castellanos, Gala Wind, P. Colarco, V. Buchard, P. Norris, and many others …

Outline

- OSSE elements and some framing remarks
- Nature runs and types of simulators
- Parameterizing subgrid variability in clouds
- Retrieval vs radiance error modeling
- Some OSSE examples and on-going activities
- Concluding remarks

O.S.S.E

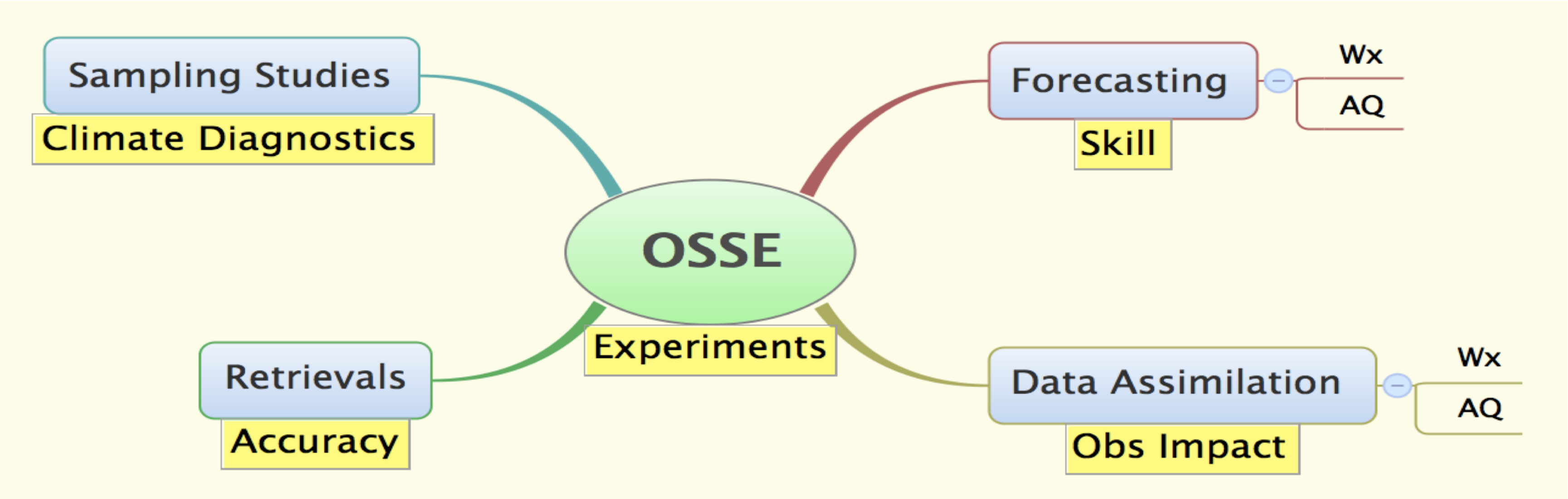
- **Observing System**
- **Simulation**
- **Experiment**

Model-based OSSE

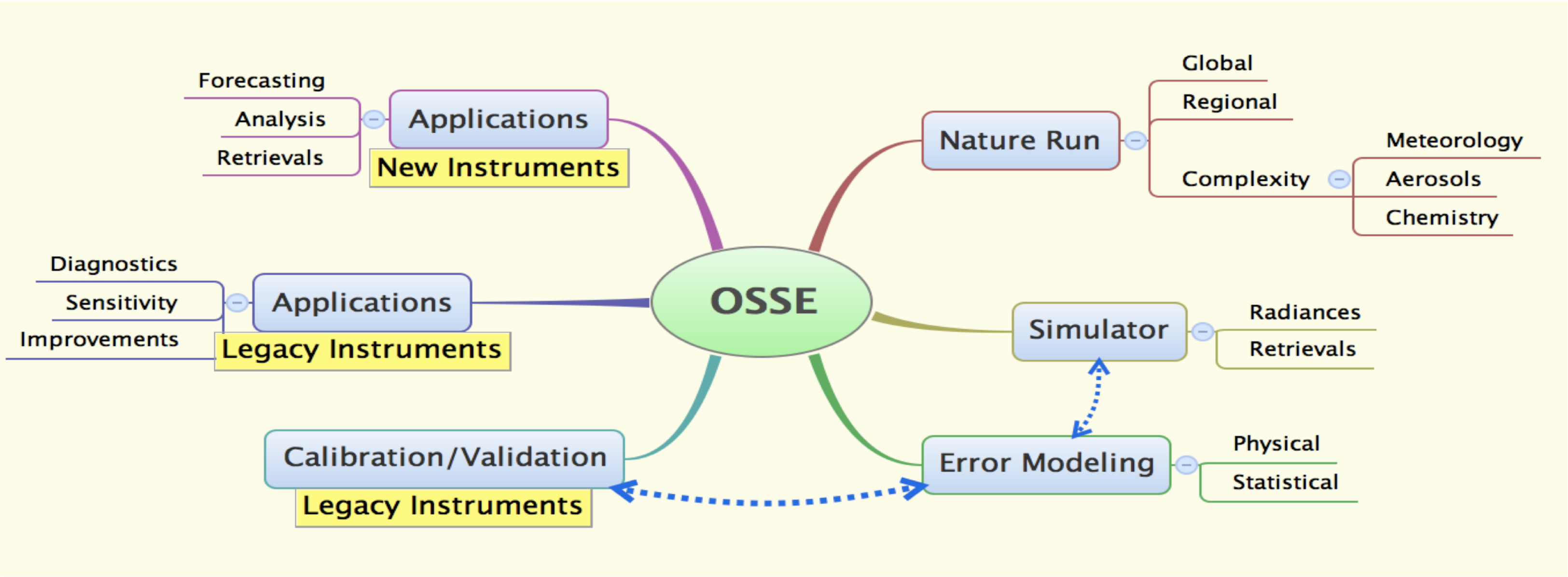
A framework for numerical experimentation in which *observables* are simulated from fields generated by an earth system model, including a *parameterized* description of the *observational error* characteristics.

Simulations are performed in support of an experimental goal.

The "E" in OSSE


Sampling Studies
Climate Diagnostics
Forecasting
Wx
AQ
Skill
OSSE
Experiments
Retrievals
Accuracy
Data Assimilation
Wx
AQ
Obs Impact

Elements of an OSSE System



The Validation Imperative

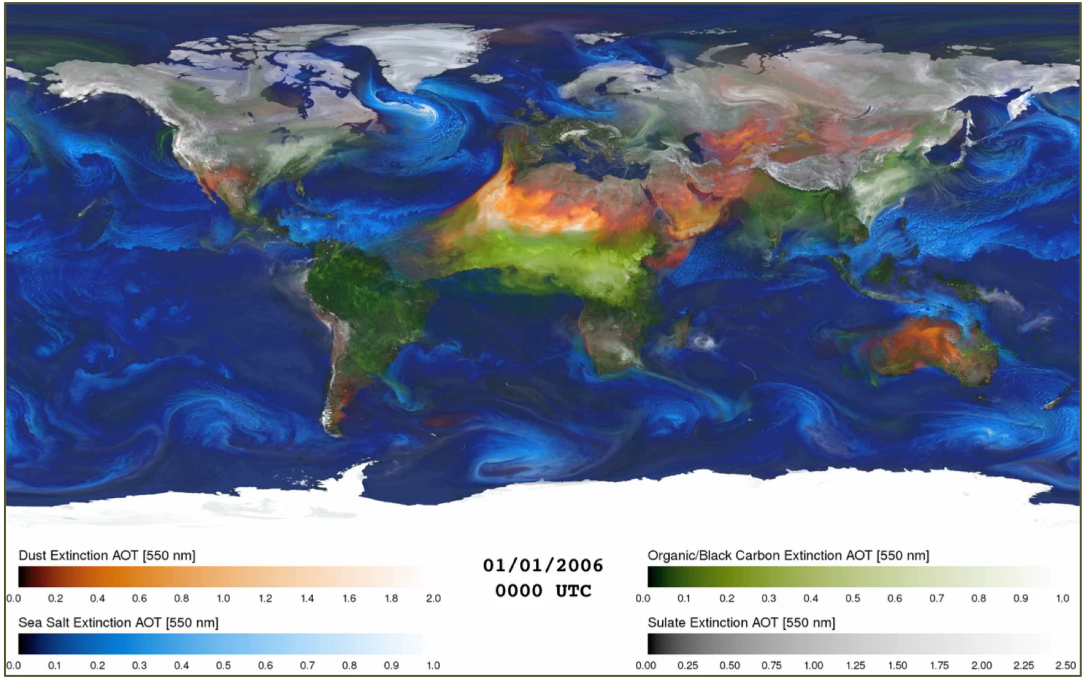
- As with any simulation, OSSE results apply to new instruments only to the degree they have been validated with existing legacy instruments.
- OSSE credibility is first determined by carefully comparing a variety of statistics that can be computed in both the real and OSSE simulated contexts.

OSSEs need to be validated as a *System*.

Global Aerosols

GEOS-5 7 km Nature Run (G5NR) Global Mesoscale Simulation

Aerosols play an important role in both weather and climate. They are transported around the globe far from their source regions, interacting with weather systems, scattering and absorbing solar and terrestrial radiation, and modifying cloud micro- and macro-physical properties. They are recognized as one of the most important forcing agents in the climate system.



For details of simulation, including Documentation and live data access see: https://gmao.gsfc.nasa.gov/global_mesoscale/7km-G5NR/

OSSE: Types of Simulators

Level 1 simulators

- Detailed radiative transfer calculation in the presence of clouds, aerosols, ice, etc.
- Instrument characteristics
- Example Observables: polarized radiances, backscatter

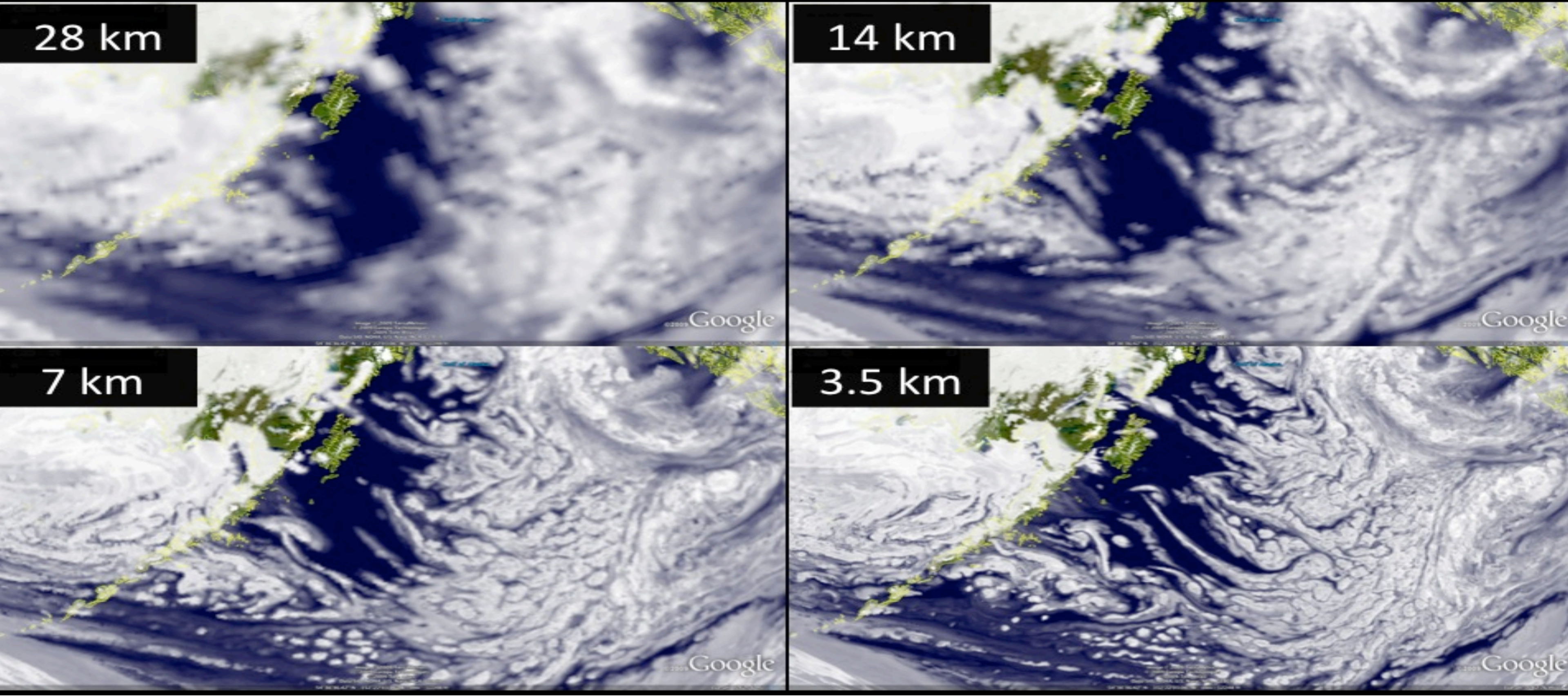
Level 2 simulators

- Retrieved quantities at observation location
- Averaging kernels, error characteristics

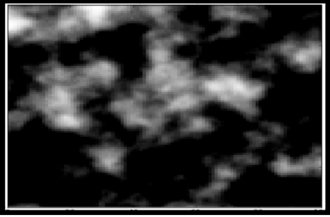
Level 3 simulators

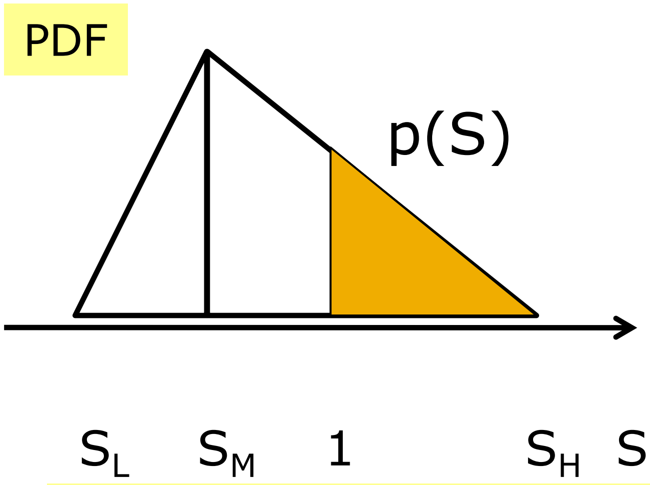
- Hourly to seasonal mean statistics sampled at the instrument footprint

Sub-Grid Variability



Clouds & Sub-grid Variability





$$S = (q_v + q_L + q_I) / q_S(T)$$

Norris and da Silva, 2016

- PDF-based cloud parameterizations provide very useful information about sub-grid variability
- Given a PDF of total water one can generate sub-columns consistent with that PDF
- Observation simulators can account for representativeness error by operating on these sub-columns

Radiance Error Modeling

Radiance errors (detector noise) ϵ are defined as

$$y = f(z^t, b^t) + \epsilon$$

where

y	radiance measurement
f	forward function (radiative transfer function)
z^t	true state ($z^t = Hw^t$)
b^t	true forward model parameters (e.g., spectral line data, calibration parameters)
ϵ	detector noise + error of representativeness

The real physics of the radiative transfer is often too complex or its details unknown. In practice, a forward model (F) is used

$$f(z^t, b^t, b'^t) = F(z^t, b^t) + \delta f(z^t, b^t, b'^t)$$

Retrieval Error Analysis

Ignoring the transfer function bias for the prior states, the retrieval error reads:

$$\begin{aligned} \epsilon^r &\equiv z - Hw^t \\ &= (I - A)\epsilon^p + D_y F_b \epsilon_b + D_y \delta f + D_y \epsilon \end{aligned}$$

where $\epsilon^p = z^p - z^t = H(w^p - w^t)$, etc., and

A	averaging kernel ($= D_y F_z$)
$(I - A)\epsilon^p$	smoothing error (prior error)
$D_y F_b \epsilon_b$	forward model parameter error
$D_y \delta f$	forward model error
$D_y \epsilon$	instrument + representativeness error

Simulating Retrievals

FROM RADIANCES

- Synthetic retrievals
  - Simulate radiances by radiative transfer
  - Model radiance errors
  - Apply retrieval code

BY MODEL SAMPLING

- Sample and perturb
  - Interpolate geophysical to obs location
  - Model retrieval errors
  - Done.

While interpolating a model simulated geophysical quantity to observation location is much more straightforward than performing a full RT calculation, modeling retrieval errors is far more complex than modeling radiance errors.

MODIS Cloud & Aerosol Retrieval Simulator

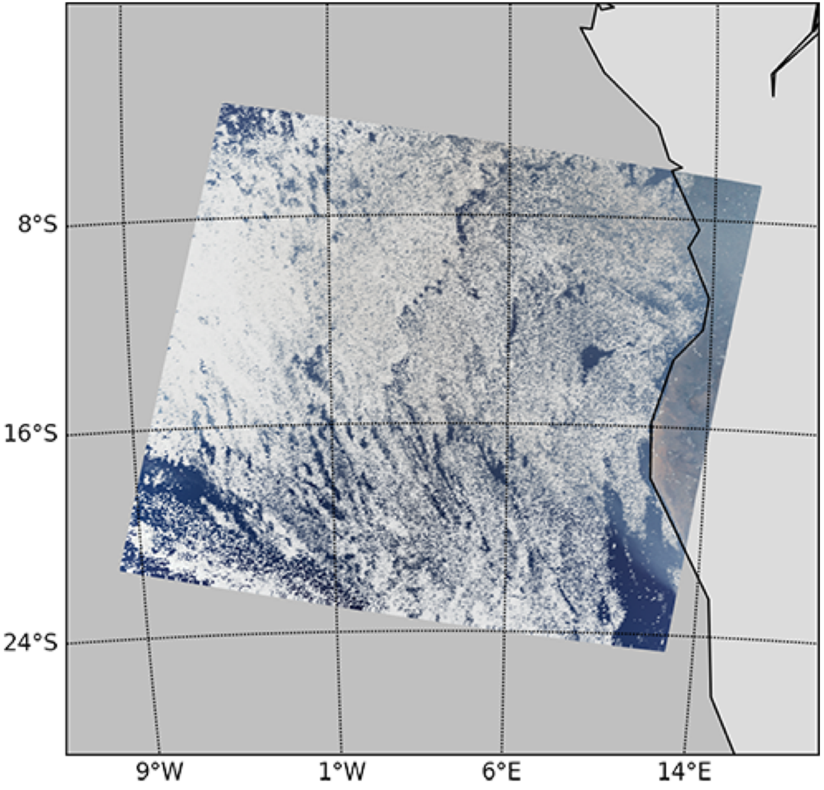
- Algorithm proofing sandbox
- 1km MODIS sensor geometry + 7km GEOS-5 Nature Run + Total Water PDF sampling to go from 7km to 1km
- 25 MODIS channels (410nm – 14.2μm)
- Correlated-k atmospheric transmittance model
- DISORT-5 radiative transfer core
- Output to standard 1-km MODIS radiance file

- Any data product code runs as if presented with real data, no awareness of radiance source
- Can examine retrieval code in fine detail
- Supercomputing application (400 processors, 8.5 hours wall-clock-time, 32 streams per granule)

Wind et al. 2013, 2016.

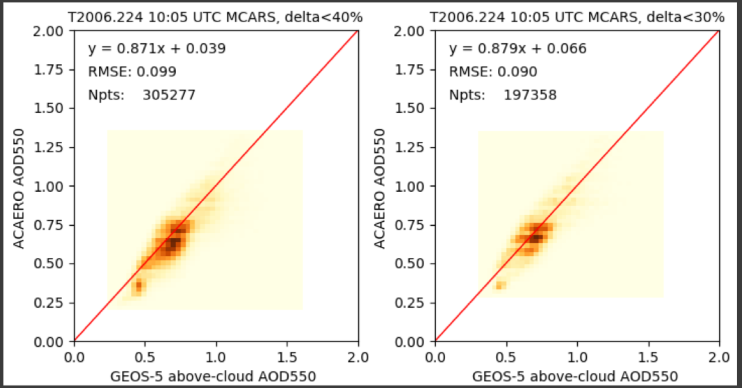
MODIS ACAERO Algorithm Evaluation

- MODIS Above-Cloud Aerosol Optical Properties by K. Meyer
- Returns aerosol optical depth, cloud optical thickness and cloud effective radius with pixel-level uncertainty at 1km resolution
- Uses 6 MODIS channels (440nm – 2.1μm)
- MODIS Dark-Target operational absorbing aerosol model
- Above-cloud retrievals over marine boundary layer clouds
- Uses MODIS Cloud product for cloud top pressure and cloud thermodynamic phase information
- Ran during ORACLES campaign as a near-real-time (NRT) product

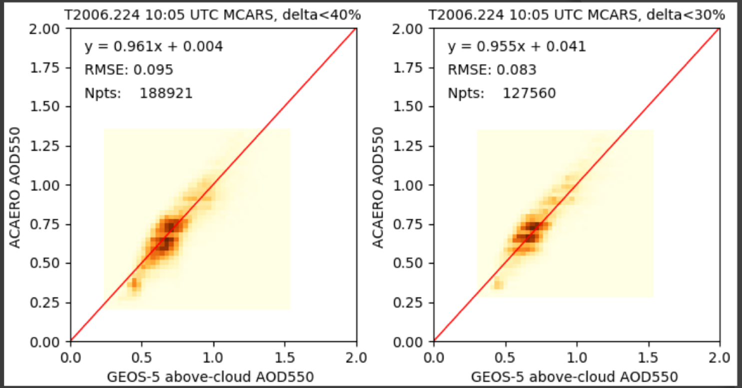


Wind et al. 2018, *in preparation*.

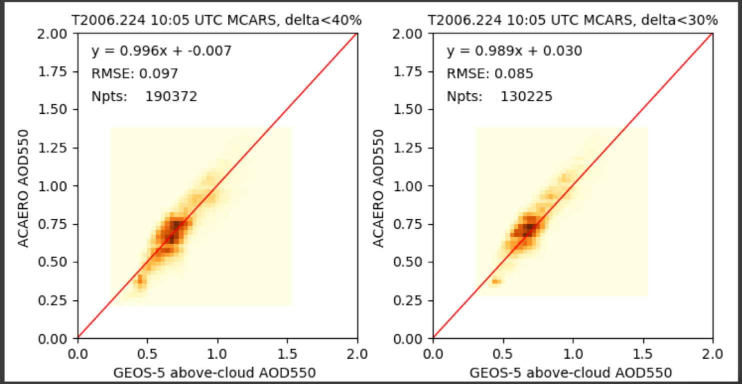
Add screening by sensor zenith < 30 degrees



by sensor zenith < 20 degrees



Add GEOS-5 input as ancillary



Recipe

Assimilate points with:

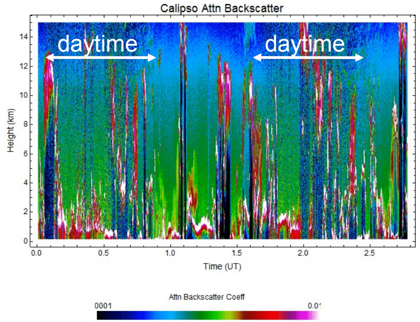
1. Pixel-level uncertainty < 40%
2. Cloud optical thickness > 4
3. Avoid the rainbow scattering angle
4. Select pixels with sensor zenith < 20^{o}

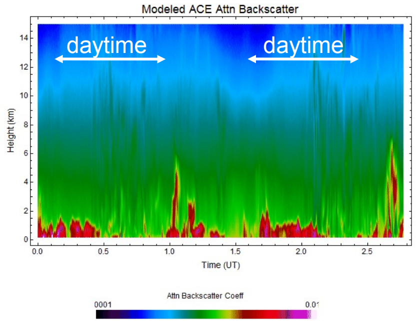
ACE Lidar Simulator

Approach

- Simulate HSRL lidar measurements for full Calipso orbit July 15, 2009 at 10 s resolution
  - Density and 3+2 aerosol optical profiles from GEOS-5
  - Radiance values from RT model (VLIDORT)
- Study yields for microphysical retrievals considering both 3+1 and 3+2 configurations
- Study microphysical inversions using these data. Compare with GEOS-5 references.

Calipso track for July 15, 2009.





Courtesy David Whiteman (GSFC)

RMS Difference Between GEOS-5 Microphysics and LIDAR Inversions

	Case A: η > 0.75 (Fine Mode Predominance)																			
	Errors 0-15 %				Errors 15-20 %				Errors 20-30 %				Errors 30-40 %				Errors 40-50 %			
	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532
R_{eff}	48.6	59.3	44.9	73.3	51.2	53.2	45.6	64.9	54.4	55.2	49.7	67.2	53.2	56.4	47.3	68.3	48.9	59.8	43.1	61.2
V	16.2	18.4	22.1	30.5	16.7	17.7	21.5	28.4	19.6	19.4	23.5	28.4	17.7	17.6	21.2	26.2	16.5	17.3	20.1	25.3
S	34.3	39.3	35.3	52.3	33.7	37.3	35.0	49.3	37.0	38.7	34.5	48.9	35.5	37.0	32.0	47.2	36.0	35.9	28.2	43.4
m_r	0.03	0.03	0.04	0.03	0.03	0.03	0.04	0.03	0.04	0.03	0.04	0.03	0.04	0.03	0.04	0.03	0.04	0.04	0.05	0.03
m_i	9E-3	8E-3	8E-3	8E-3	8E-3	8E-3	8E-3	7E-3	8E-3	8E-3	7E-3	7E-3	8E-3	8E-3	8E-3	8E-3	7E-3	7E-3	7E-3	7E-3
	Case B: 0.25 < η < 0.75 (Mixture)																			
	Errors 0-15 %				Errors 15-20 %				Errors 20-30 %				Errors 30-40 %				Errors 40-50 %			
	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532
R_{eff}	40.5	22.6	52.1	23.0	42.4	24.4	52.3	24.8	44.4	24.2	54.2	25.6	52.0	28.2	58.2	25.4	52.1	33.1	57.8	25.8
V	28.1	19.9	52.5	40.1	30.7	22.0	52.0	39.7	32.1	22.6	53.7	40.7	37.0	23.5	57.2	44.0	36.6	28.8	56.5	39.8
S	43.6	22.1	16.0	26.2	46.7	23.4	17.0	27.3	49.7	23.8	17.4	27.6	60.0	24.9	19.1	29.5	67.8	25.6	21.9	28.8
m_r	0.11	0.13	0.08	0.08	0.11	0.13	0.08	0.08	0.10	0.12	0.08	0.08	0.11	0.13	0.09	0.08	0.11	0.12	0.09	0.08
m_i	2E-3	2E-3	5E-3	4E-3	2E-3	2E-3	5E-3	4E-3	3E-3	2E-3	5E-3	4E-3	2E-3	2E-3	5E-3	4E-3	2E-3	2E-3	5E-3	4E-3
	Case C: η < 0.25 (Coarse Mode Predominance)																			
	Errors 0-15 %				Errors 15-20 %				Errors 20-30 %				Errors 30-40 %				Errors 40-50 %			
	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532	Reg.	Reg. 3b1a 532	LE	LE 3b1a 532
R_{eff}	58.2	65.6	70.0	68.4	59.2	65.8	70.2	68.1	55.9	64.8	70.3	69.2	55.3	65.1	70.8	70.2	49.1	67.6	70.9	72.1
V	60.7	68.3	71.7	71.8	61.2	67.2	71.5	71.2	57.6	67.5	71.3	72.5	56.7	69.3	72.5	74.7	50.9	72.1	71.1	77.5
S	14.7	21.8	14.4	17.8	14.9	20.2	14.0	16.3	15.2	21.1	14.9	17.7	12.5	21.3	13.1	19.0	12.0	22.1	11.5	20.5
m_r	0.14	0.15	0.11	0.11	0.14	0.15	0.11	0.11	0.15	0.15	0.11	0.11	0.15	0.15	0.11	0.12	0.12	0.23	0.12	0.22
m_i	4E-3	3E-3	3E-3	3E-3	4E-3	3E-3	4E-3	3E-3	4E-3	3E-3	4E-3	3E-3	5E-3	4E-3	4E-3	4E-3	5E-3	4E-3	4E-3	4E-3

Fine/Coarse separation done based on GEOS-5 optical data

Whiteman et al. 2016

Lidar-Polarimeter Simulator

- **Surface:**
  - MODIS RTLS bi-directional reflectance
  - BPDF from Maignan et al. (2009)
    - Polarized reflectance that is a function of IGBP land use and NDVI
    - Fits POLDER measurements, spectrally flat
  - Possible to add on GISS Cox-Munk surface reflectance for ocean scenes if there is interest
- **Atmosphere**
  - 7 km Global GEOS-5 Nature Run (GOCART)
  - Rayleigh scattering
  - Optical properties are RH dependent
- **Orbits, Angles, and Wavelengths are speciable, for example:**
  - CALIPSO, ISS, 425 km orbit, etc…
  - VZA: 3.66, 11., 18.33, 25.66, 33, 40.33, 47.66, 55.0
  - Wavelengths: 354, 388, 410, 440, 470, 550, 670, 865, 1020, 1650, 2130
  - Observables: intensity, DoLP
- **RTM: VLIDORT v2.7**
- **Test simulation files can be found here.**
  - https://portal.nccs.nasa.gov/datashare/G5NR/c1440_NR/OBS/POLAR_LIDAR/CALIPSO/

Other OSSE Activities of Relevance

- **Several GEO-CAPE related activities (P. Castellanos)**
  - G5NR-chem, a Nature Run with full tropospheric tropospheric chemistry
  - Radiance simulator for several golden days (aerosol channels):
    - GOES-R, GEMS, TEMPO, SENTINEL-4
  - CO and AOD (forecast) OSSEs (David Edwards, J. Barré – NCAR)
- **OMI/OMPS related activities**
  - OMI AI simulator (V. Buchard, P. Colarco, S. Gassó, O. Torres)
  - OMPS volcanic SO_2 retrieval OSSEs
- **AERONET retrieval OSSEs**
- **GRASP-ACE: joint lidar-polarimeter Retrieval OSSEs (D. Ramirez, O. Duovik)**

Concluding Remarks

- **A *credible* OSSE system requires well validated modeling components:**
  - Nature run
  - Physical simulation of measurements
  - Instrument characterization and error modeling
- **However, it must be validated as a *System*, by exercising it with the existing legacy observing system.**
- **OSSE applications such as *Retrieval OSSEs and sampling studies are as relevant to ACE and A-CCP as the classical analysis and forecast skill metric***
- **While computationally demanding, there exists capacity for realistic simulations of many designated observables for multiple orbits and seasonal conditions, globally**
  - refinement of methodology and components is a joint venture involving many expertises

Some Remarks on Global Aerosol and Cloud Data Assimilation

Relevance to ACE and A-CCP

Arlindo da Silva

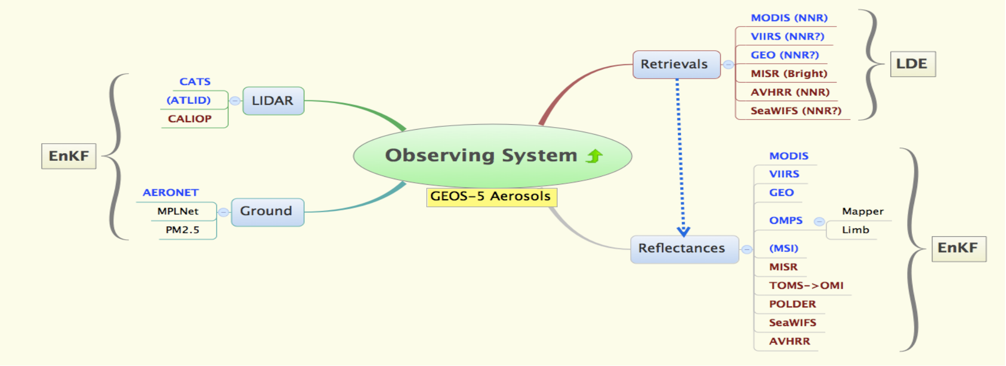
Global Modeling and Assimilation Office

NASA Goddard Space Flight Center

With contributions from V. Buchard, P. Norris, and many others …

Aerosol Observing System

- **Aerosol Optical Depth (AOD) is the most commonly available observable**
  - Vertically integrated mass weighted by extinction coefficient, summed over multiple species: *low observability*
  - Available multi-spectral AOD measurements are not really measured
- **Radiance assimilation:**
  - Vector scattering calculations needed for UV-VIS measurements are not cheap
  - Surface BRDF characterization is a challenge
- **Surface PM 2.5**
  - Single level
  - Often plagued by representativeness



- **Lidar measurements provide vertical info**
  - Spatially coverage is poor (pencil thin)
  - Attenuated backscatter again requires optical assumptions which are not directly measured
    - HSRL concept does provide additional info.

Aerosol Data Assimilation

State representation

- Multiple 3D concentrations
  - ✓ Mass
  - ✓ Number (modal schemes)
  - ✓ Bin sizes (sectional schemes)
- Number of tracers: tens to hundreds

Emissions:

- ✓ Dynamic: dust, marine, biogenic aerosols
- ✓ Remotely sensed: biomass burning
- ✓ Inventories: anthropogenic

Observation operators

- Intrinsic aerosol optical properties needed for remotely sensed data:
  - ✓ Mass extinction coefficient, single scattering albedo, phase matrix
  - ✓ These are often poorly known but assumed to be known due to identifiability issues:

$$\tau = \beta \bullet \mathbf{M}$$

Aerosol Analysis: Splitting

2D AOD ANALYSIS

- Observable 550 nm AOD is 2D
  - Constrains column averaged optics
  - Cannot constrain speciation or vertical distribution
- Analysis in observation space:

$$\tau^a \equiv Hq^a = H\left(q^b + \delta q^a\right)$$
$$= \tau^b + \delta\tau^a$$

GOING TO 3D CONCENTRATIONS

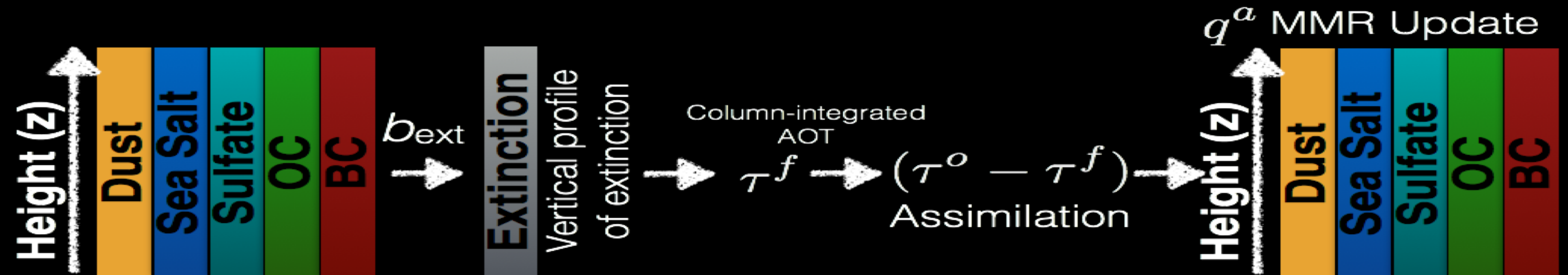
- Based on error covariances:

$$\delta q^a = BH^T\left(HBH^T\right)^{-1}\delta\tau^a$$

- Using ensemble perturbations,

$$\delta q^a = XY^T\left(YY^T\right)^{-1}\delta\tau^a$$

- NRT GEOS-5 uses Local Displacement Ensembles (LDE), in 1D
- Developing EnKF for Aerosols

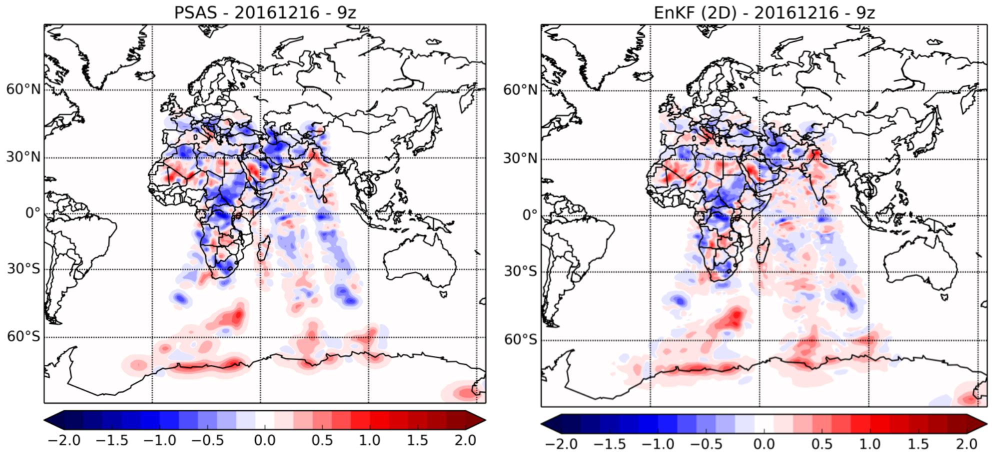


In Development: Aerosol EnKF

- As part of GMAO's hybrid system, aerosol ensemble members are produced as a matter of routine
- The same Whitaker-Hamill EnKF used for the hybrid Meteorological assimilation has been adapted for aerosols
- Target observation systems
  - Multi-spectral AOD: 470, 550 and 870 nm
  - Lidar attenuated backscatter
  - Sensors: MODIS, VIIRS, GEO, CATS/CALIOP, TropOMI

PSAS AOD inc (current method)

EnKF AOD inc Ensmean



Assimilation of Clouds in Global Models

- **The so-called assimilation of microwave *all-sky radiances* are intended to improve the representation of temperature and moisture**
  - Retrieved cloud properties are often discarded
- **Visible, NIR cloud retrievals are usually at a resolution still much higher the global models**
  - Markov Chain/Monte-Carlo (MCMC) methods to exist to use these measurements to estimate parameters of the total water (vapor+condensate) p.d.f. (e.g., Norris and da Silva, 2016.)

Concluding Remarks

- Aerosol Data assimilation provide an integrative approach to synthetize aerosol information from
  - Passive LEO/GEO sensors
  - HSRL and backscatter lidars
  - Polarimeters, ocean color instruments, etc.
- Prescribed optical properties are often a challenge for observation operators
  - Systematic airborne in-situ measurements such as those in the SAM-CAAM concept may become a key element for constraining these optical properties.
- Assimilation of multi-spectral aerosol retrievals based on prescribed aerosol models poses many challenges:
  - Internal consistency, homogenization of observing system
- Model assimilated/forecast aerosol fields downscaling to ground sites may be key for AQ applications
  - Model Output Statistics by means of machine learning trained on historical ground monitor data