Aerosol and Meteorological OSSEs:

For Forecasting and Retrieval Applications

Arlindo da Silva(1)

Arlindo.daSilva@nasa.gov

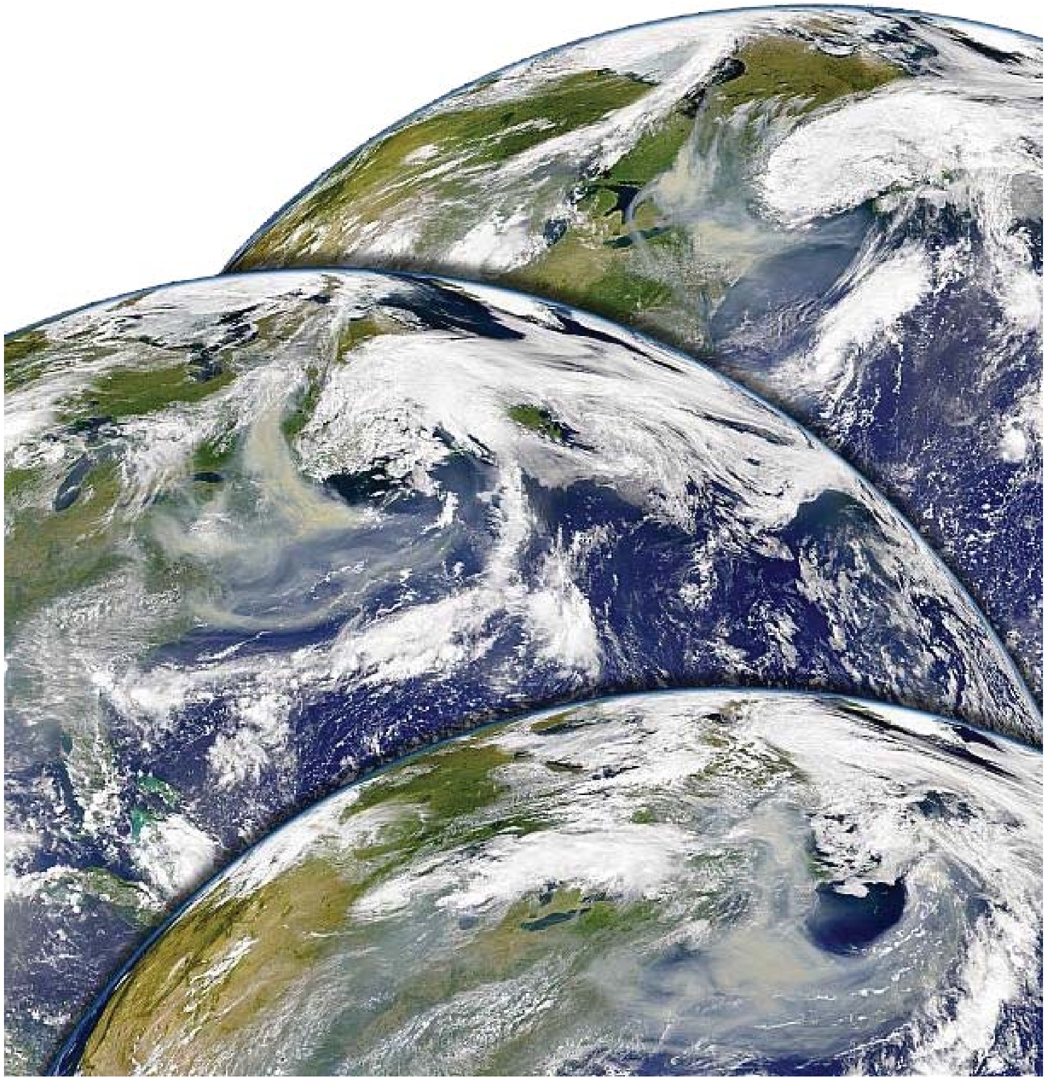
Gala Wind(5,6) ,Ron Errico (1,3) , Nikki Prive(1,3) , W. Putman(1) , P. Norris(1,4) , P. Colarco(2) and the GEOS-5 Development Team

(1) Global Modeling and Assimilation Office, NASA/GSFC Code 610.1
(2) Atmospheric Chemistry and Dynamics Lab, NASA/GSFC Code 614
(3) GESTAR/Morgan State University
(4) GESTAR/USRA
(5) Science Systems and Applications, Inc.
(6) Climate and Radiation Lab, NASA.GSFC Code 613

ICAP 2014 Workshop
21-24 October 2014, Boulder, Colorado

Outline

- OSSE:
  - Scope and Definition
- The Nature Run
- Simulators
- Error Characterization
- The "E" in OSSE
- OSSE Activities at NASA/GMAO
- Concluding Remarks



O.S.S.E.

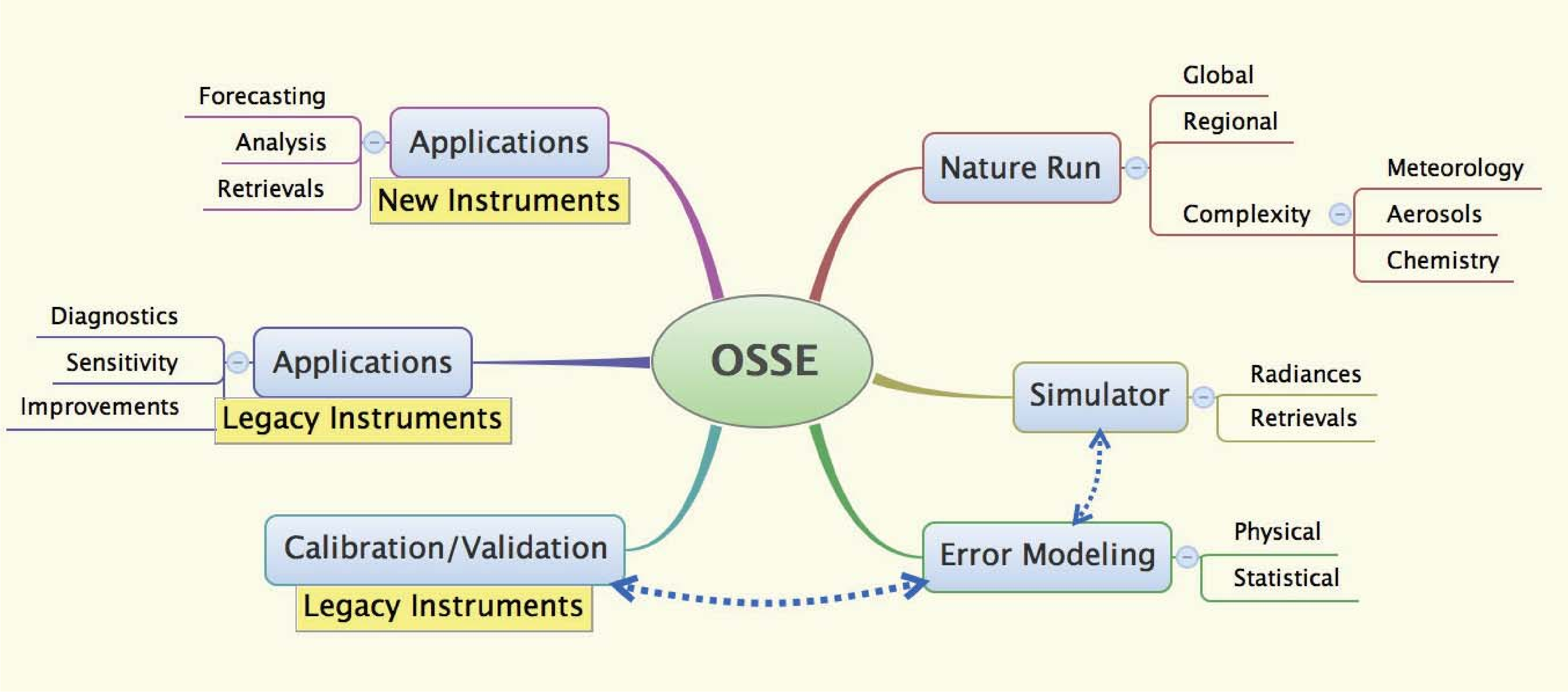
- Observing System
- Simulation
- Experiment

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.

Elements of an OSSE System



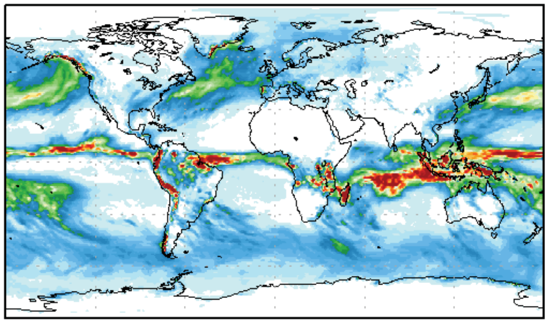
The Validation Imperative

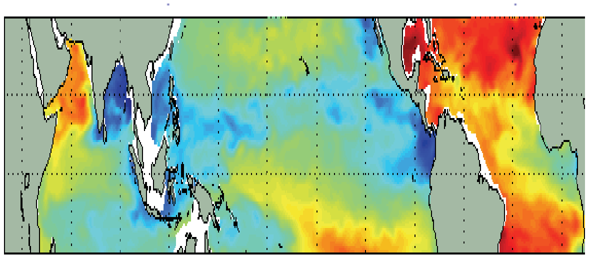
- As with any simulation, OSSE results apply to new instruments only to the degree they have been validated with existing legacy instruments.
- OSSE credibility is first determined by carefully comparing a variety of statistics that can be computed in both the real and OSSE simulated contexts.

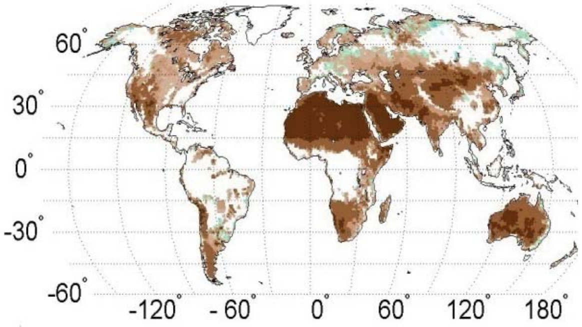
OSSEs need to be validated as a *System*.

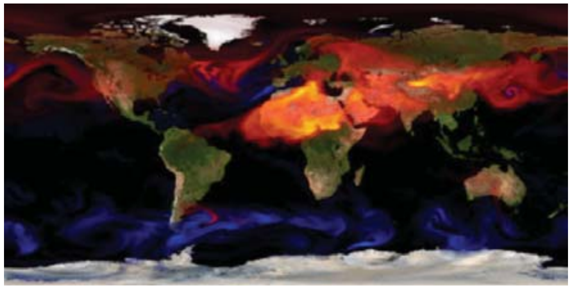
Nature Run

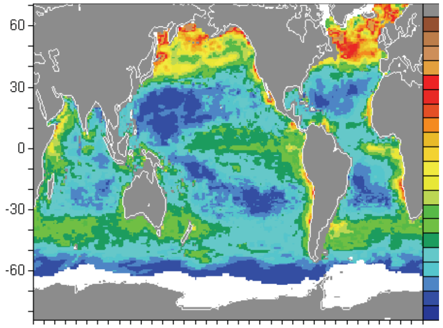
IESA: Integrated Earth System Modeling

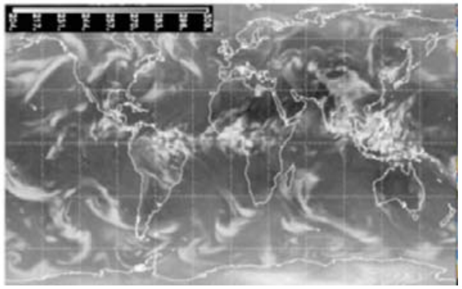


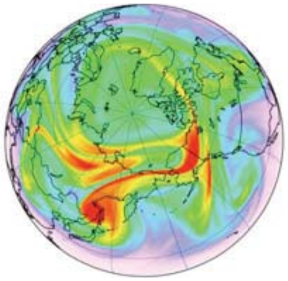


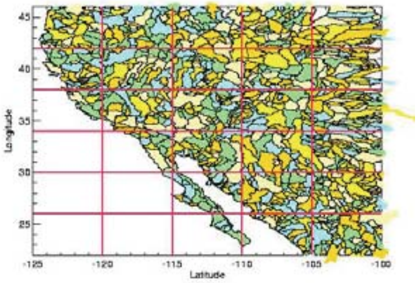


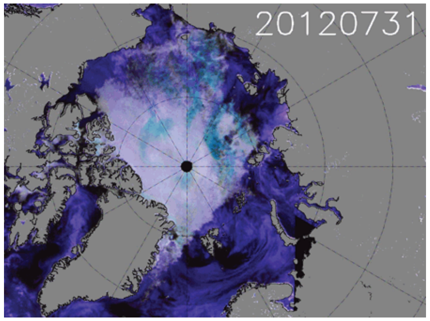






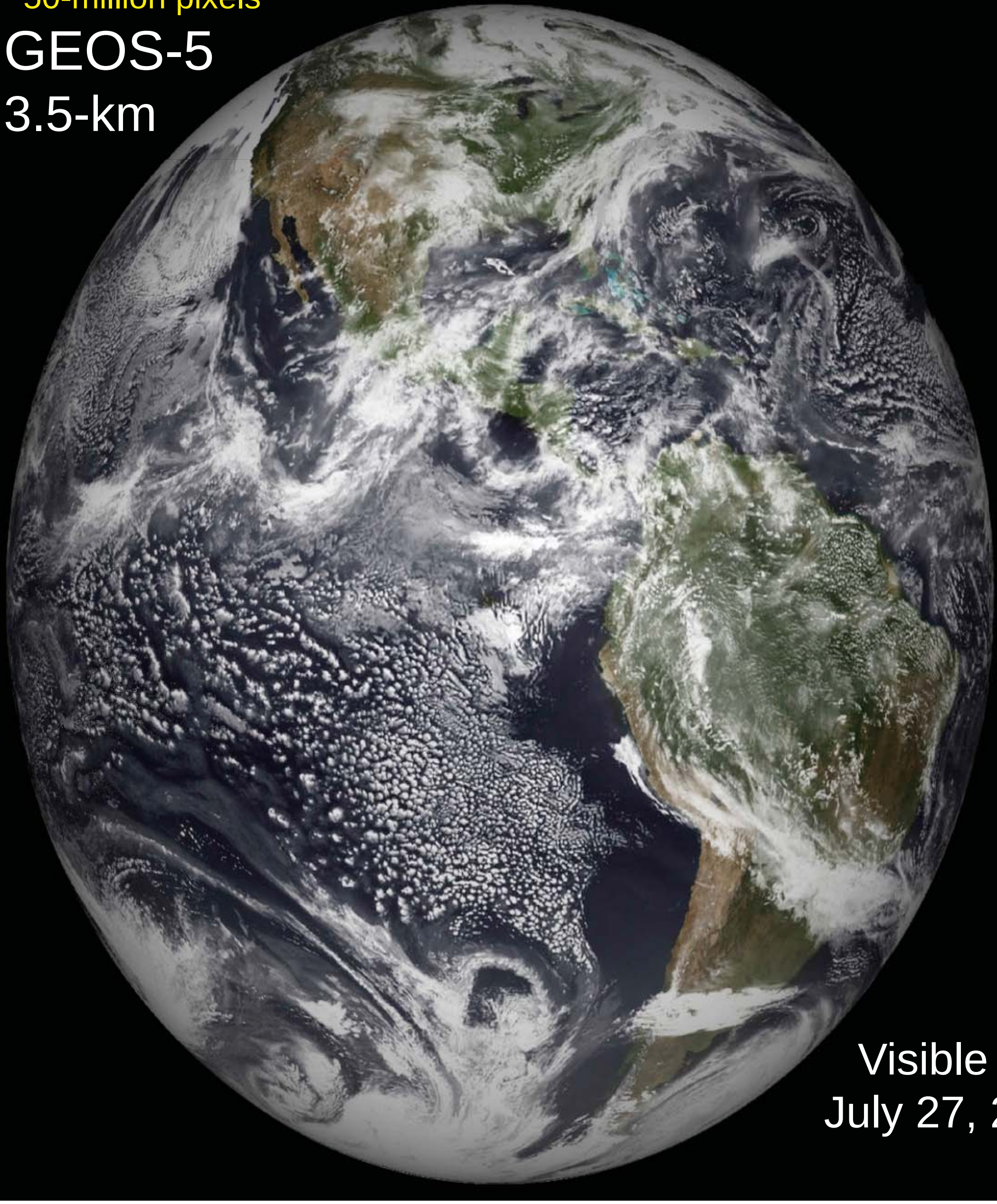






Balancing act: Complexity x Resolution x Spatial/Temporal Coverage

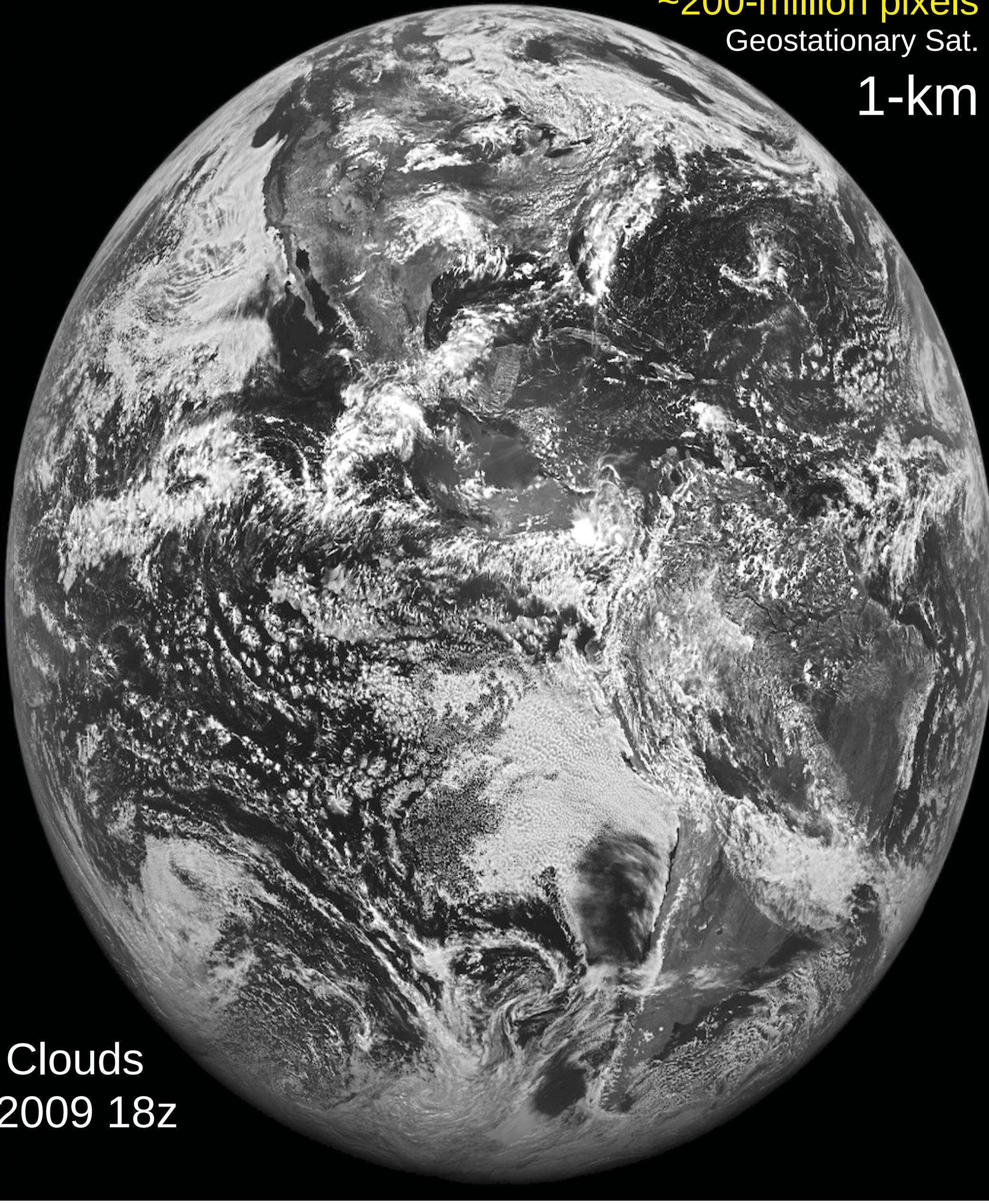
Nature Run Realism



~50-million pixels

GEOS-5

3.5-km



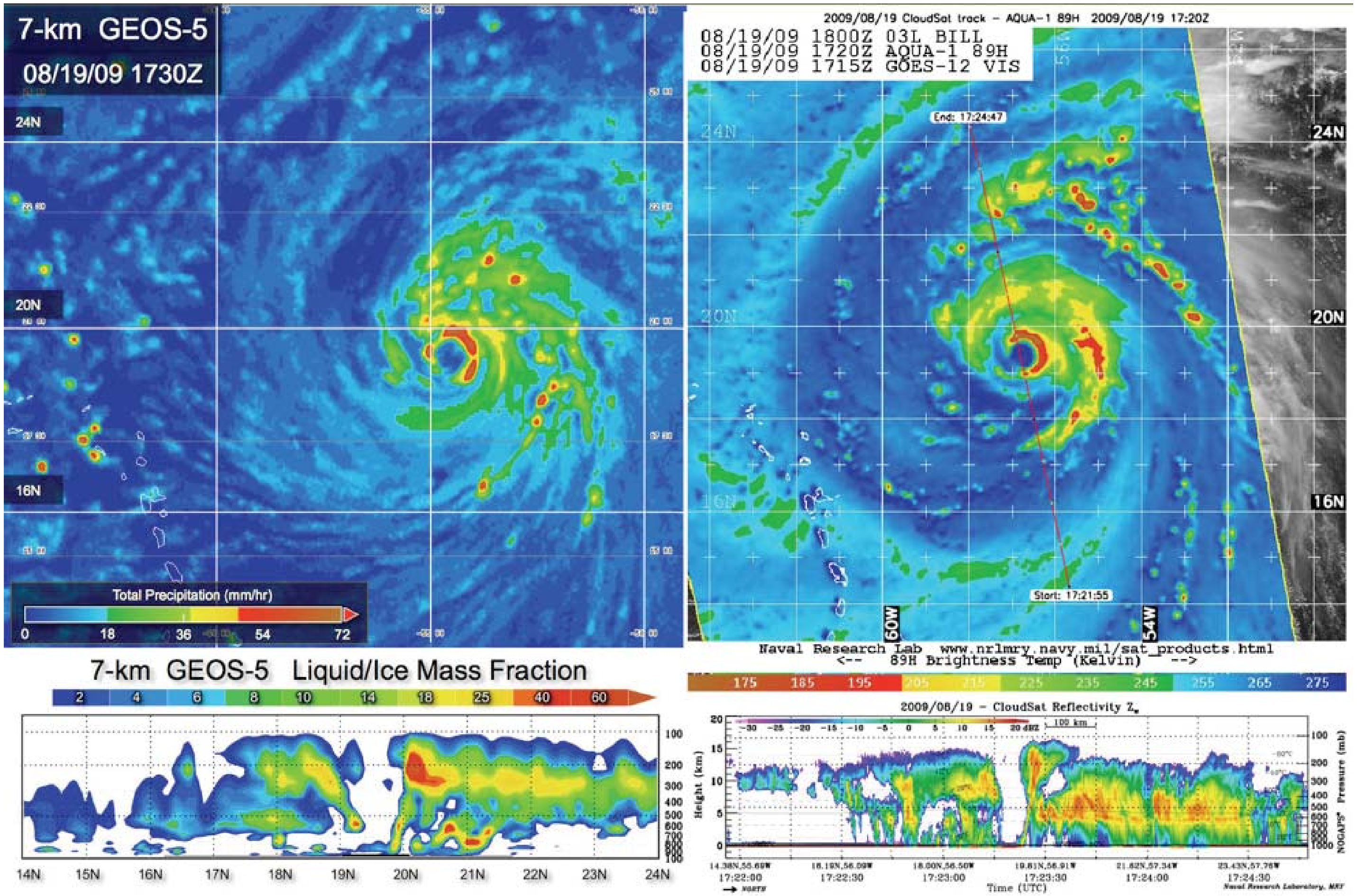
~200-million pixels

Geostationary Sat.

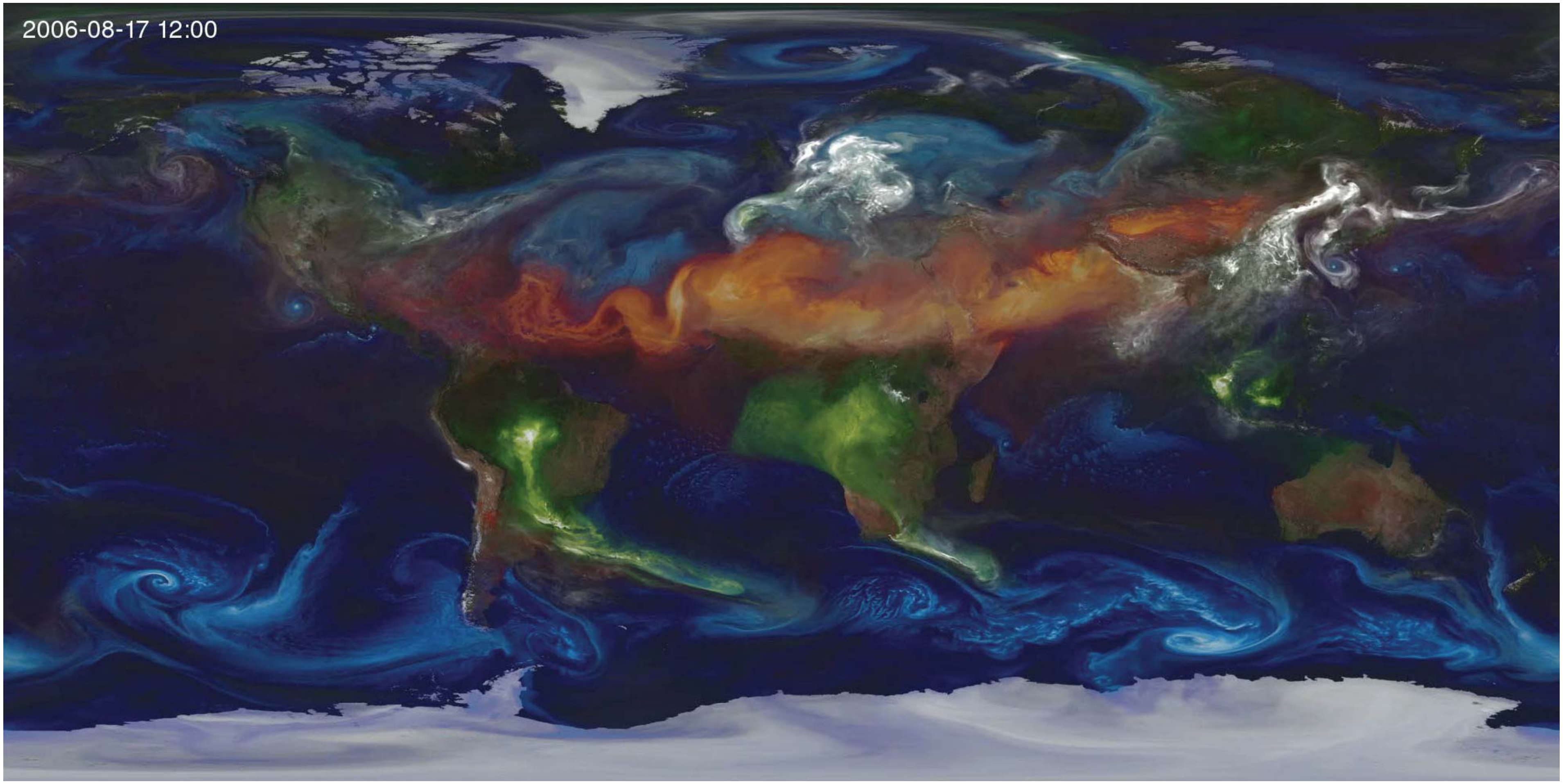
1-km

Visible Clouds
July 27, 2009 18z

Nature Run Validation



Atmospheric Composition NR: Simulated or Observed Meteorology?



GEOS-5 10km Global Mesoscale Simulation: SST, aerosol emissions

Simulators and Observation Error Modeling

OSSE:
Types of Simulators

- Level 1 simulators
  - Detailed radiative transfer calculation in the presence of clouds, aerosols, ice, etc.
  - Instrument characteristics
  - 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

Observation Error

at Instrument Footprint

Measurement equation:

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

where

y	measurement
z^t	true state at instrument footprint
f	observation operator
ϵ	observation error w.r.t. z^t (detector noise)

Error of Representativeness

Let x^t be the true state in model grid point space:

$$z^t = \mathcal{I}(x^t) + \epsilon'$$

where

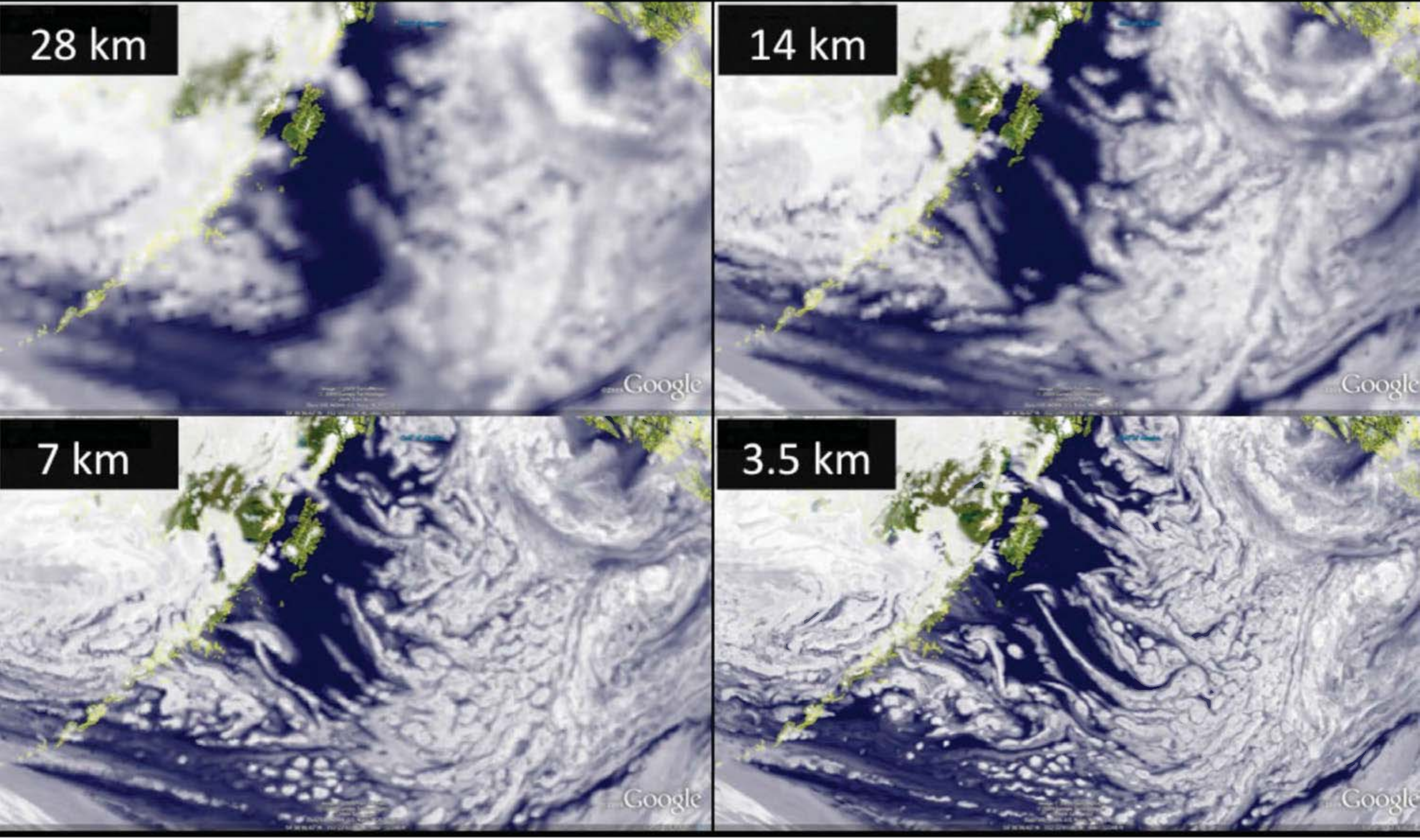
$\mathcal{I}$ *remapping* operator
ϵ' representativeness error

In OSSE studies the Nature Run is assumed to provide the ground truth, but it does so in grid-point space: x^t

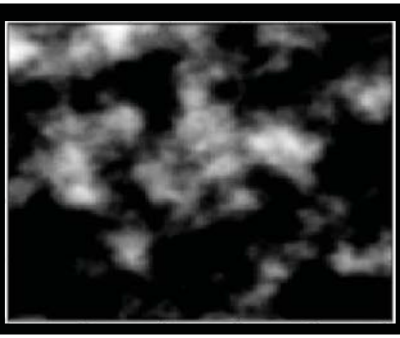
$$\begin{aligned} y &= f(\mathcal{I}(x^t) + \epsilon') + \epsilon \\ &= h(x^t) + \epsilon + F_z\epsilon' + .. \end{aligned}$$

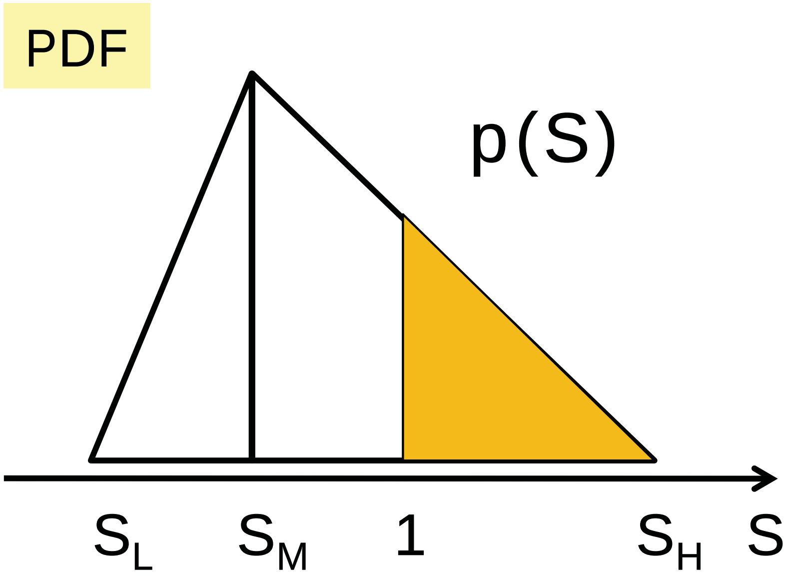
It is critical to account for error of representativeness errors when simulating observables with a footprint much smaller than the Nature Run resolution.

Sub-Grid Variability



Clouds & Sub-grid Variability





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

- 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

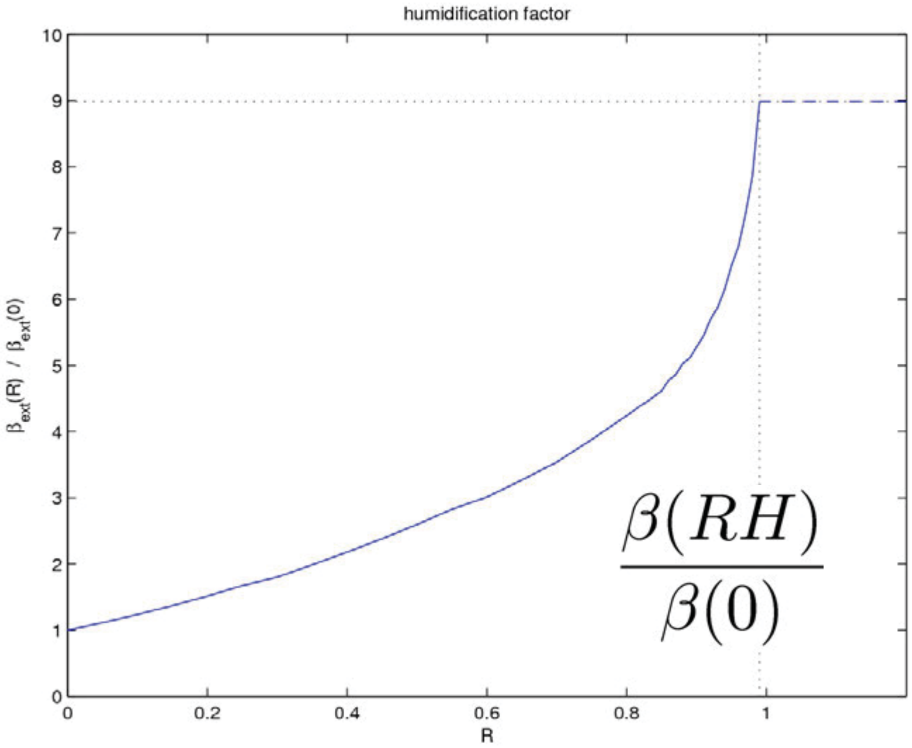
Hygroscopic Aerosols

GOCART prognosticate aerosol dry mass mixing ratio q_{dry}, with humidification effects being included diagnostically prior to computing optical depth

$$\tau = \beta(RH; p) \cdot q_{\mathrm{dry}} \cdot \rho_a \delta z$$

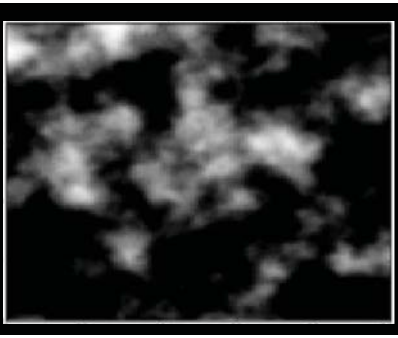
The normalized mass extinction efficiency

$$\hat{\beta} = \frac{\beta(RH)}{\beta(0)} \sim 1 - 10$$



$\hat{\beta}$ saturates at 99%

PDF-based Humidification



Recall that aerosol satellite retrievals are only available under *clear sky* conditions:

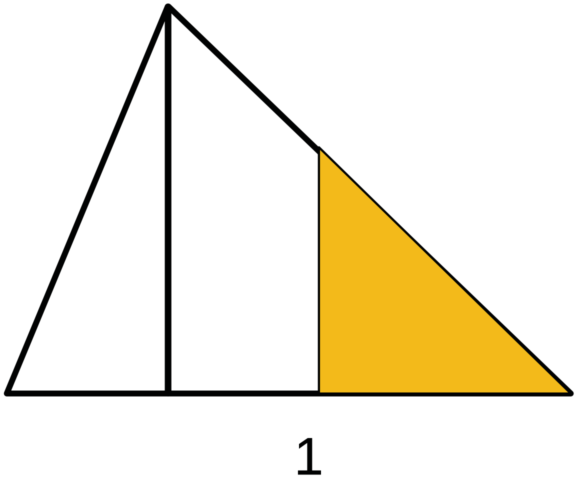
$$\tau_{\text{clear}} = \beta_{\text{clear}}(RH) \cdot q_{\text{dry}} \cdot \rho_a \delta z$$

PDF-based cloud schemes as in GEOS-5 can be used to estimate the mean humidification effect on a GCM gridbox

$$\begin{aligned} <\hat{\beta}> &= \int_0^\infty p(S)\hat{\beta}(S)dS \\ &= \int_0^1 p(S)\hat{\beta}(S)dS + \int_1^\infty p(S)\hat{\beta}(S)dS \\ &= (1-f)\cdot <\hat{\beta}>_{\text{clear}} + f\cdot <\hat{\beta}>_{\text{cloudy}} \end{aligned}$$

where the *cloud fraction* f is given by

$$f = \int_1^\infty p(S)ds$$



A PDF of water vapor + condensate is provided in each gridbox

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 Modeling

- Retrievals are produced as result of a nonlinear estimation,

  $$z = D(y, b, z^p) = D(F(z^t, b^t) + \delta f + \epsilon, b, z^p)$$

  where z^p is a background state (prior) used in the retrieval.

- In order to obtain a basic understanding we linearize this equation around a known state, say the prior z^p

  $$\begin{aligned} z - z^p &= [D(F(z^p, b), b, z^p) - z^p] \\ &\quad + D_y F_z (z^t - z^p) \\ &\quad + D_y F_b (b^t - b) \\ &\quad + D_y \epsilon \end{aligned}$$

  where $T(z^p, b) = D(F(z^p, b), b, z^p)$ is the so-called Transfer Function.

- The term in brackets is the transfer function bias and should be small for any reasonable retrieval,

  $$T(z^p, b) = D(F(z^p, b), b) \approx z^p$$

  i.e., the retrieval should return the first guess when noiseless data consistent with it is input.

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

Retrieval Error Mechanisms

- Averaging Kernel. The retrieved state is a smoothed version of the true state with smoothing functions given by the rows of A

  $$z = Az^t + (I - A)z^p + \dots$$

  Those details of the true state which are smoothed out by the retrieval must be provided by the first guess z^p. Notice that A is state dependent.

- Model parameter errors are usually associated with biases. If retrieval is too sensitive to a parameter then it should be retrieved as well (state augmentation).

- Model error. Hard to evaluate. If correct physics is known then it can be modeled (e.g., approximations introduced for computational efficiency.)

- Instrument error. Usually the easiest to evaluate. It can have a strong dependence on the state through D_y, although the detector noise ϵ is usually assumed stationary and state independent in clear sky. Cloud clearing introduces additional complications. However, error of representativeness is often state dependent.

Simulating Retrievals

FROM RADIANCES

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

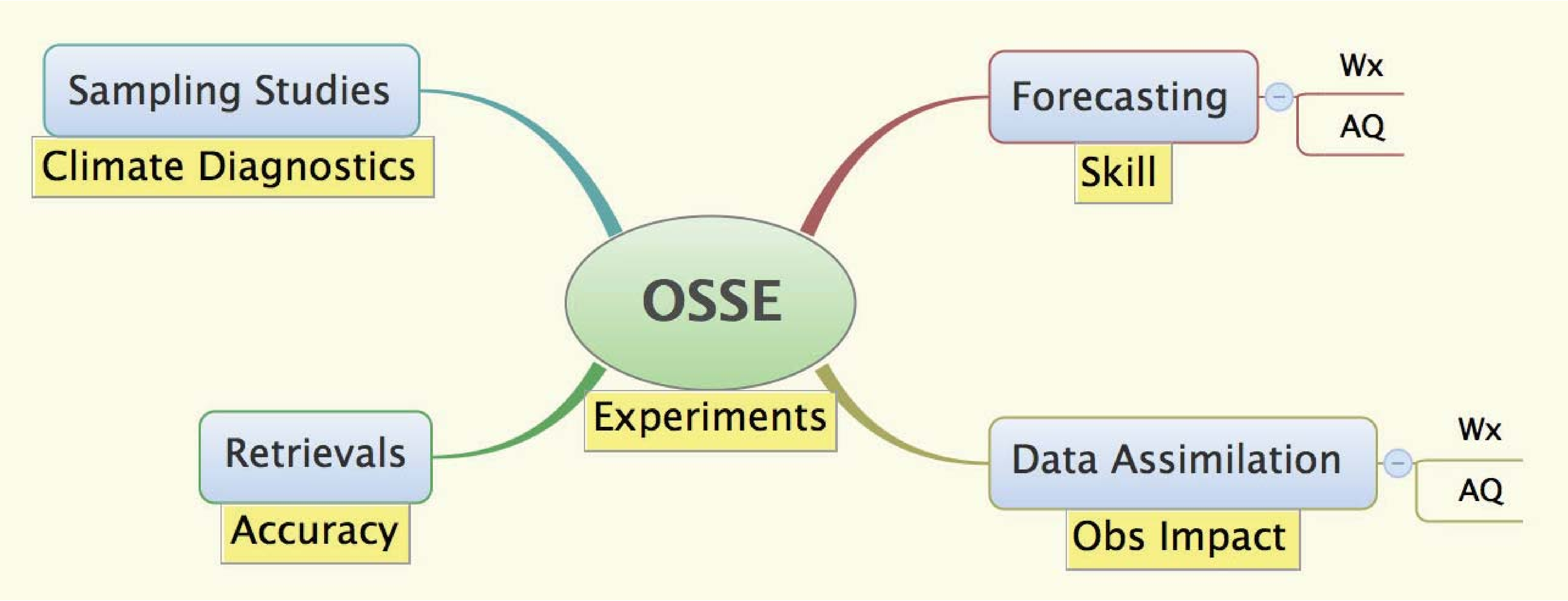
BY MODEL SAMPLING

- 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.

OSSE Experiments

The "E" in OSSE



OSSE

Experiments

Sampling Studies

Climate Diagnostics

Forecasting

Skill

- Wx
- AQ

Retrievals

Accuracy

Data Assimilation

Obs Impact

- Wx
- AQ

OSSE: Types of *Experiments*

- Forecasting/data assimilation impact
  - Goal: assess impact of new instrument
  - Metric: analysis or forecast *impact*
  - Nature Run: a high-resolution model simulation by another model
  - Calibration: observational errors for existing instruments are tuned to produce same impact as real Data Assimilation System (DAS)

OSSE: Types of *Experiments*

- Retrieval algorithm evaluation
  - Goal: given radiances/backscatter at a variety of atmospheric conditions (clouds, rain, ice, etc.) assess the quality of
    - Retrievals themselves
    - Retrieval perceived errors/averaging kernels
  - Example of metrics: direct comparison of retrievals and *Nature Run*, full set of statistics
  - Example of Nature Run: global aerosol transport model driven by “observed” meteorology
  - Calibration: when possible tune observational errors based on similar current instruments

OSSE: Types of *Experiments*

- Sampling Studies
  - Goal: assess how a given orbit, scanning characteristic would impact the ability of the measurements to reproduce key climate parameters
  - Example of metrics: hourly to seasonal time means of Level 2 parameters at a variety of spatial resolutions
  - Example of Nature: global aerosol transport model driven by “observed” meteorology
  - Calibration: noisy and perfect observations

OSSEs at NASA's Global Modeling and Assimilation Office

7-km GEOS-5 Nature Run

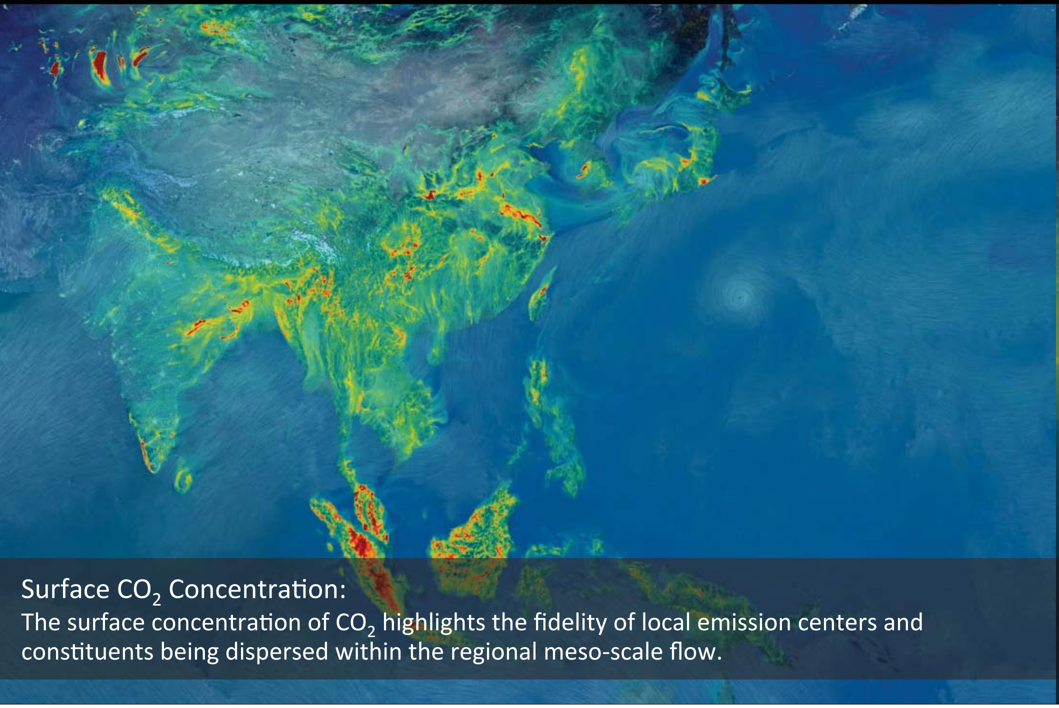
Global Modeling and Assimilation Office

Nature Run Details

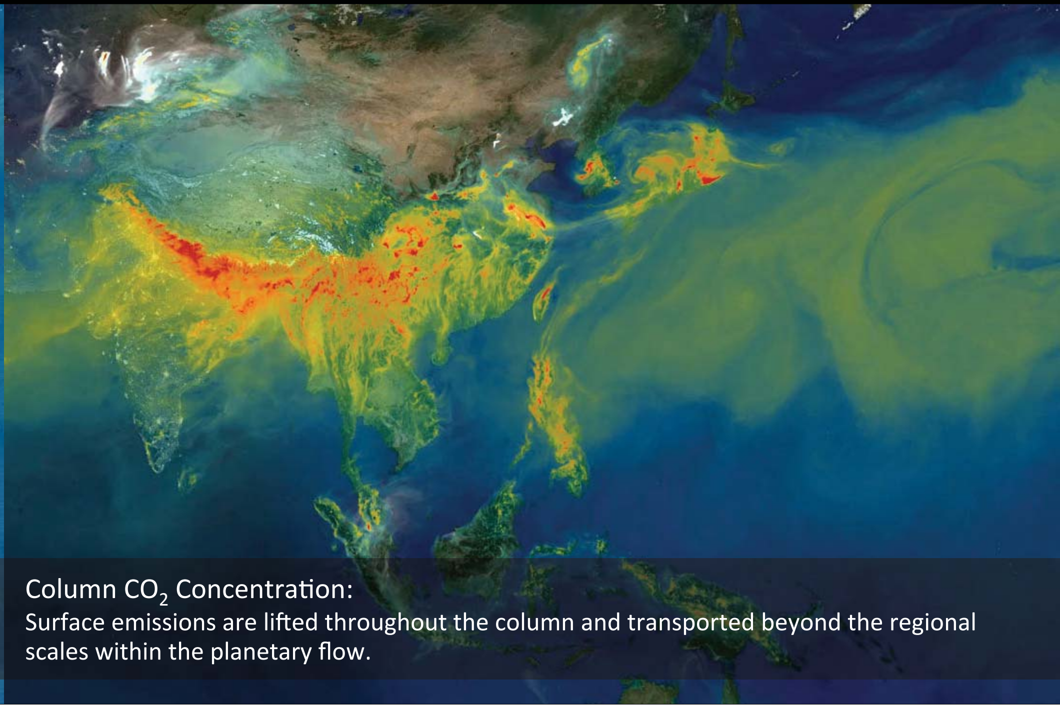
- 2-years : *June 2005 – June 2007*
- 7-km Global Resolution *[72 vertical levels : 0.01mb top]*
- 5-minute physics time step *[5 seconds for dynamics]*
- Non-Hydrostatic Dynamics *[Finite-Volume Cubed-Sphere]*
- Limited deep convection *[RAS with stochastic Tokioka limiter]*
- Resolve mesoscale weather *[storm-scale and cloud clusters]*
- High-resolution constituent transport *[GOCART]*
- Executed on “Discover” at NCCS *[7200 Cores : 11-days/day]*

Surface Boundary Conditions & Emissions

- SST and Sea-Ice *[1/4-degree Reynolds/OSTIA]*
- CO/CO_2 Fossil fuel emissions *[10-km EDGAR inventory]*
- Land CO_2 fluxes *[CASA-GFED at 10-km with MODIS EVI]*
- Biomass burning *[daily QFED emissions]*
- Volcanic SO_2 *[AEROCOM emissions and injection heights]*
- GOCART *[mixing, chemistry, and deposition]*
  - *Aerosol species [sulfates, dust, sea-salt and black carbon]*
  - *aerosols are radiatively coupled with the dynamics*



Surface CO_2 Concentration:
The surface concentration of CO_2 highlights the fidelity of local emission centers and constituents being dispersed within the regional meso-scale flow.



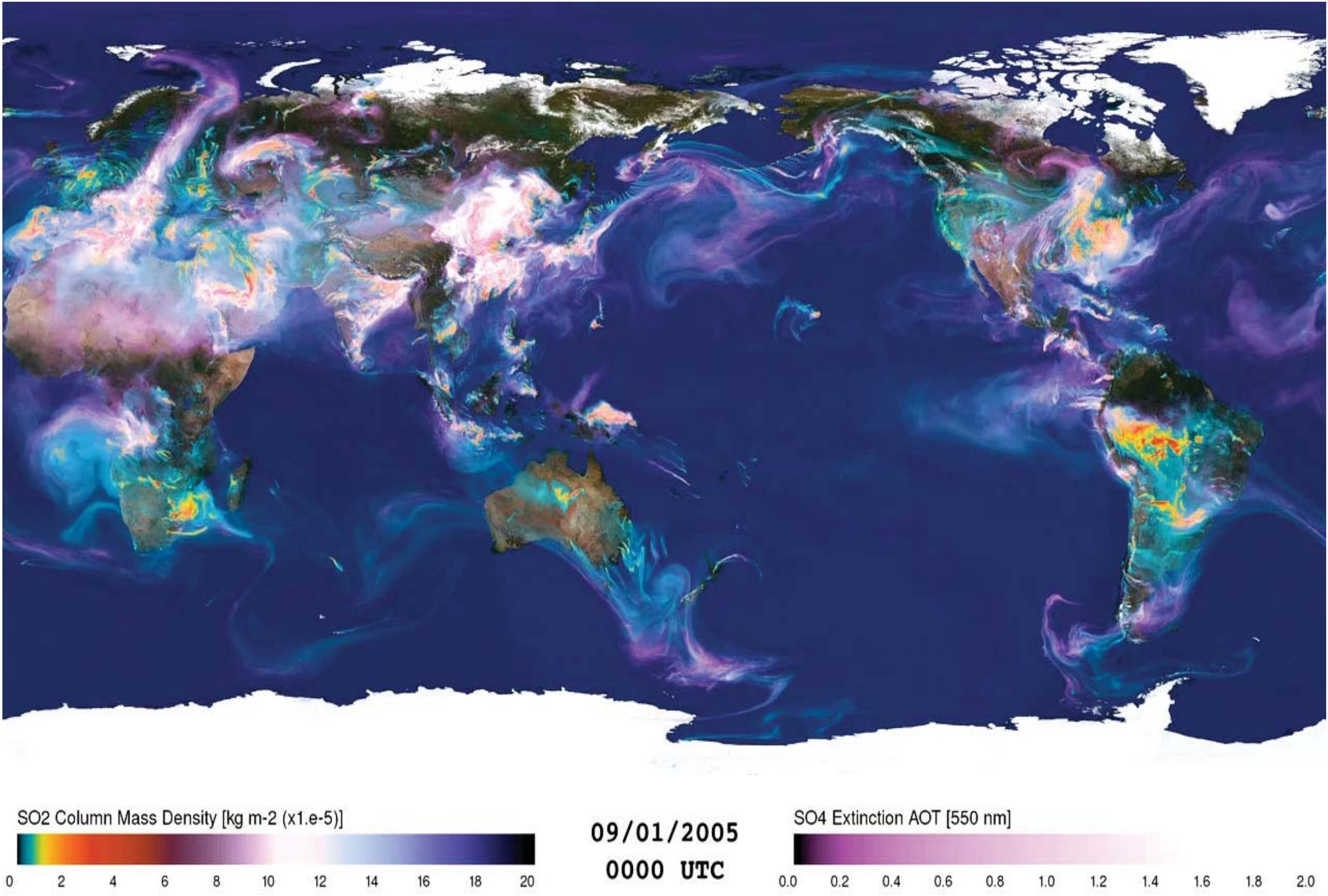
Column CO_2 Concentration:
Surface emissions are lifted throughout the column and transported beyond the regional scales within the planetary flow.

7-km GEOS-5 Nature Run

Global Modeling and Assimilation Office

Sulfur dioxide (SO_2), produced during the burning of fossil fuels and from volcanic eruptions, is a short lived gas which can act as pollutant near the surface with detrimental health and acidifying effects. With a mean life time of just a couple of days in the troposphere, emitted SO_2 is quickly converted to sulfate aerosol (SO_4) through oxidation by OH or by reaction with H_2O_2 within clouds. The resulting SO_4 exerts a direct radiative effect on the atmosphere and it can also have an indirect radiative effect by inducing changes in cloud and precipitation microphysics.

Column Concentrations of Sulfur Dioxide and Sulfate Aerosols

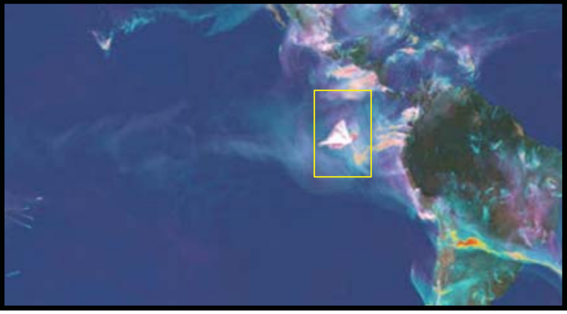


October 2005

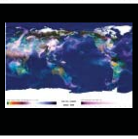
Sierra Negra Volcano

MODIS →

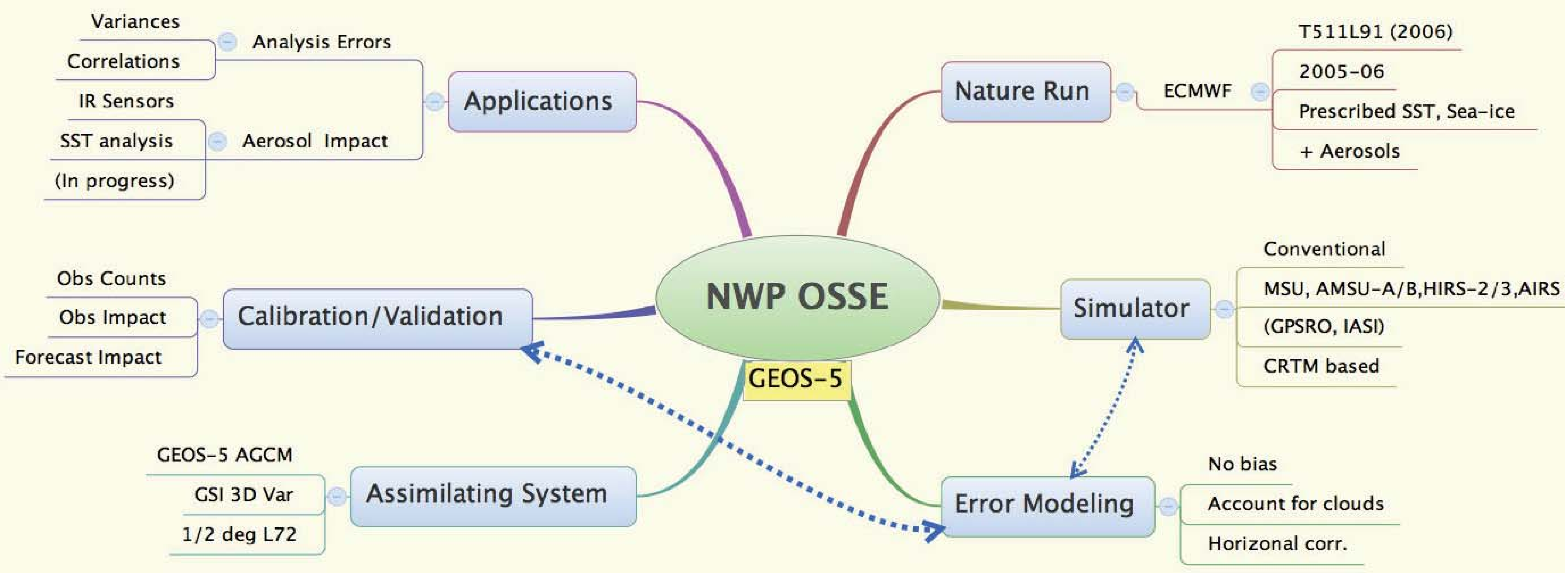
GEOS-5 ↘



The October 2005 eruption of the Sierra Negra Volcano on Isabella Island in the Galapagos Islands. A large plume of gases and steam was observed in this photo with NASA satellite imagery (top-left) on October 25, 2005. Image courtesy of MODIS, NASA. The plume of SO_2 and sulfate is seen being dispersed in the GEOS-5 7-km Nature Run (bottom-left)

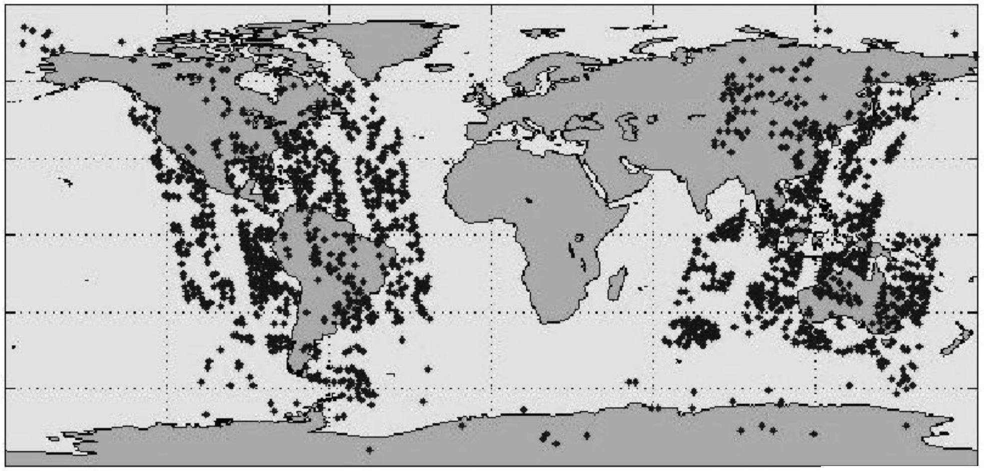


GMAO's NWP OSSE System

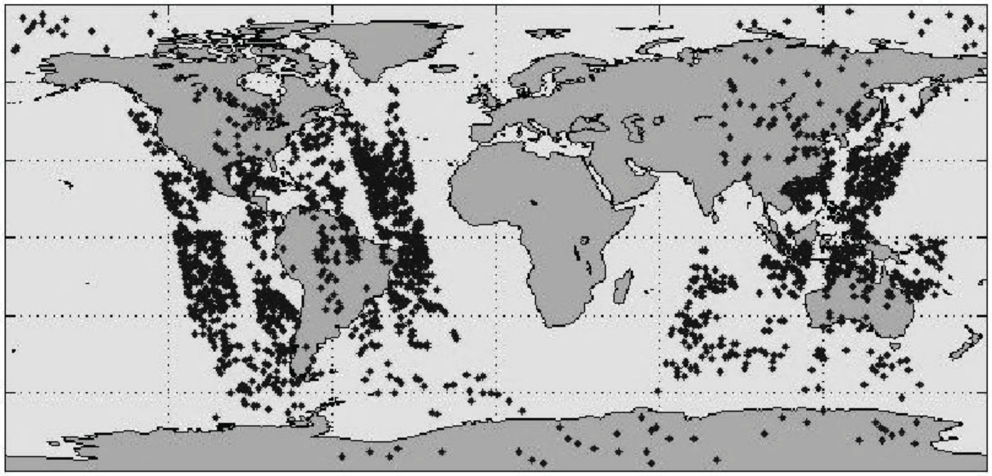


Ron Errico, Nikki Prive.

AIRS Ch 295 QC-Accepted



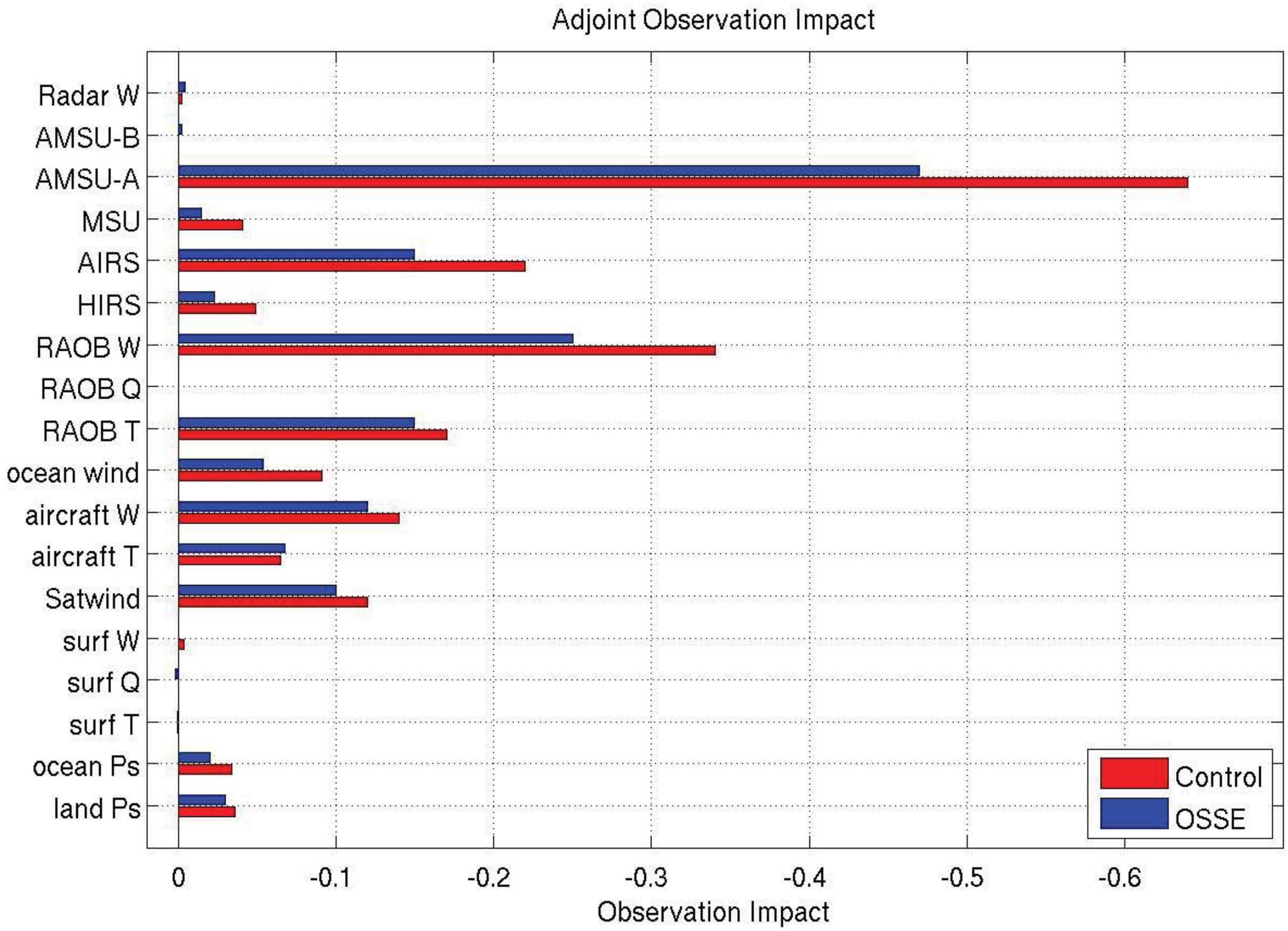
Simulated



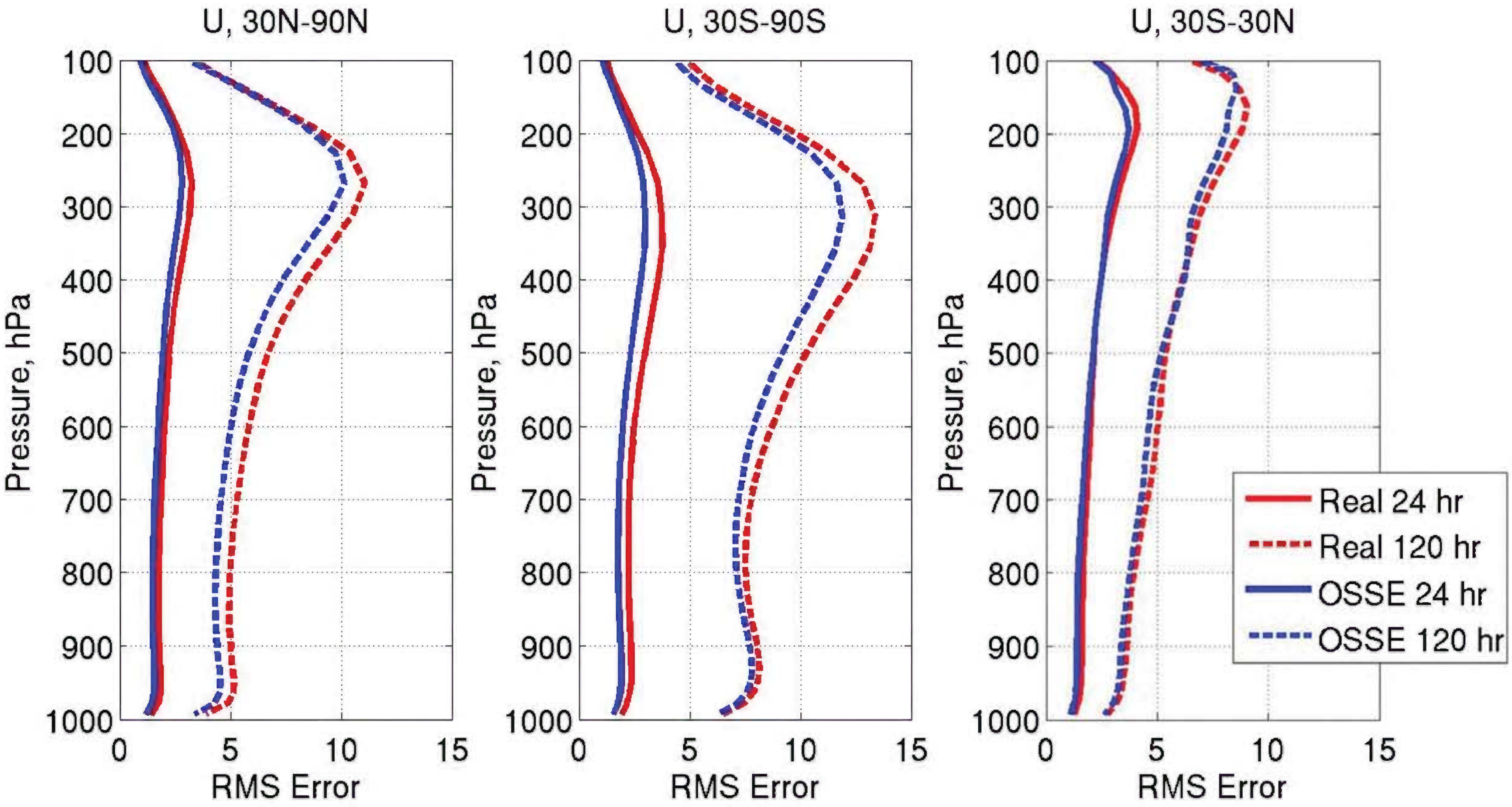
Real

18 UTC 12 July

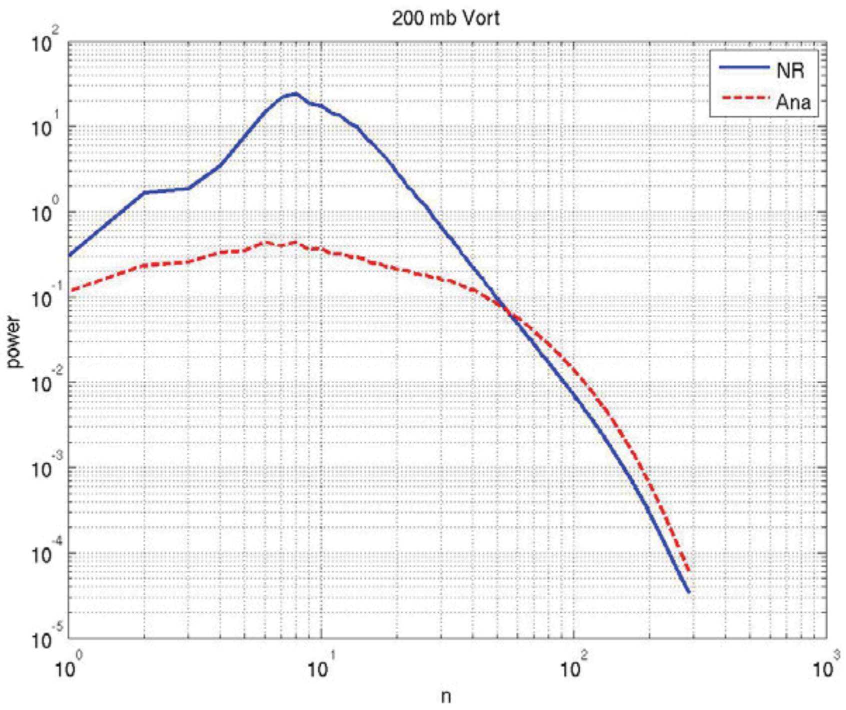
Observation Impact

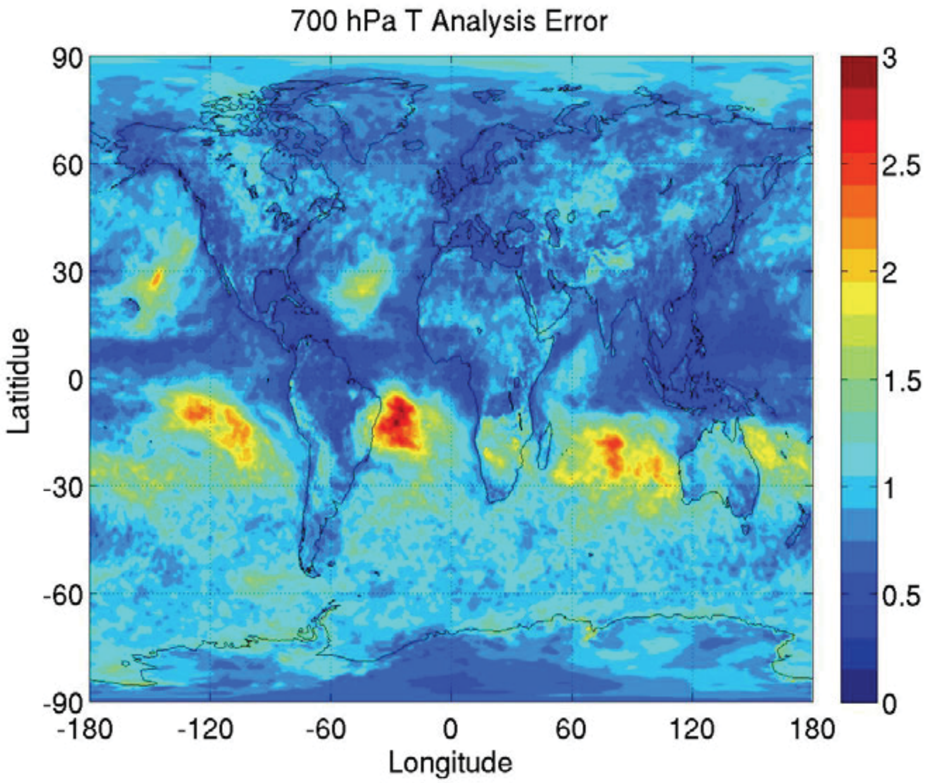


U-Wind RMS error: July



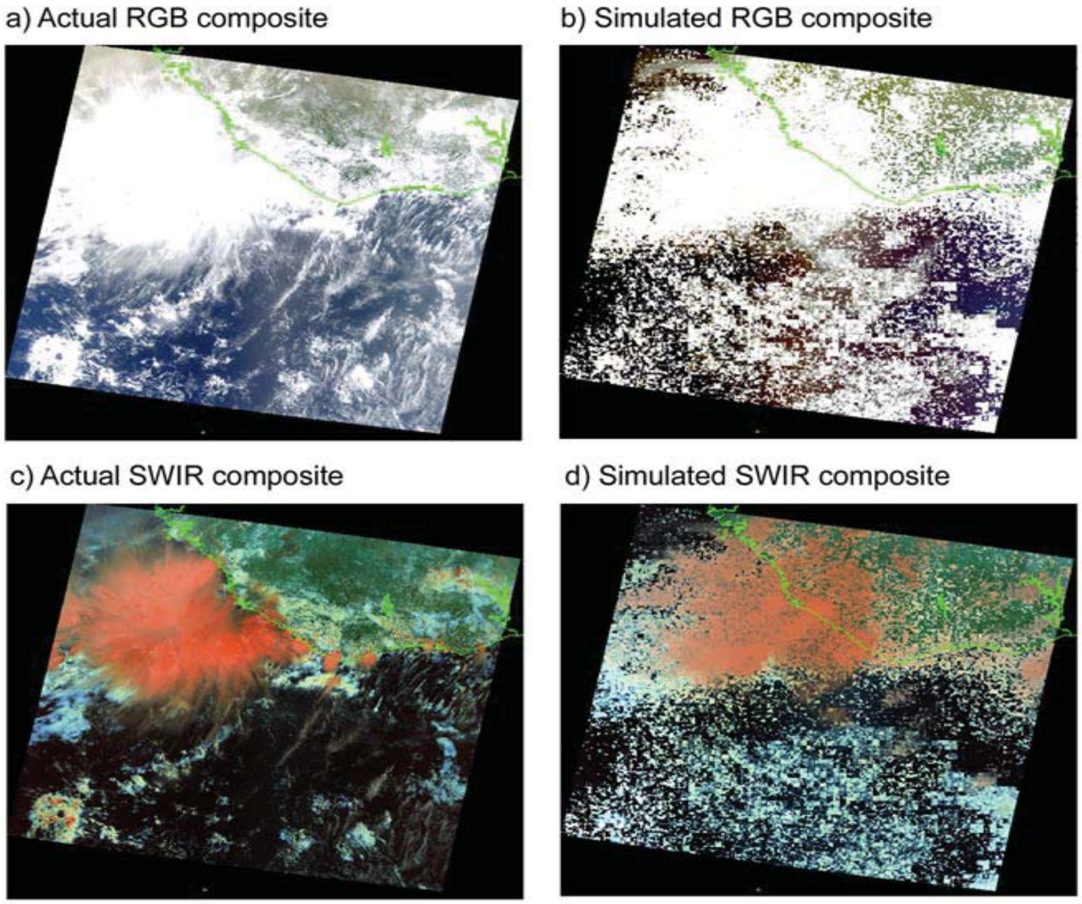
Application: Analysis Diagnostic





MODIS Cloud & Aerosol Retrieval Simulator

- PDF based sub-grid sampling of GEOS-5 fields (ICA)
- Spatial “clumping”
- Radiances for 27 MODIS channels
- Cloud and aerosol extinction, ssa, phase function
- Operational Retrievals
  - Clouds: MOD06
  - Aerosols: MOD04



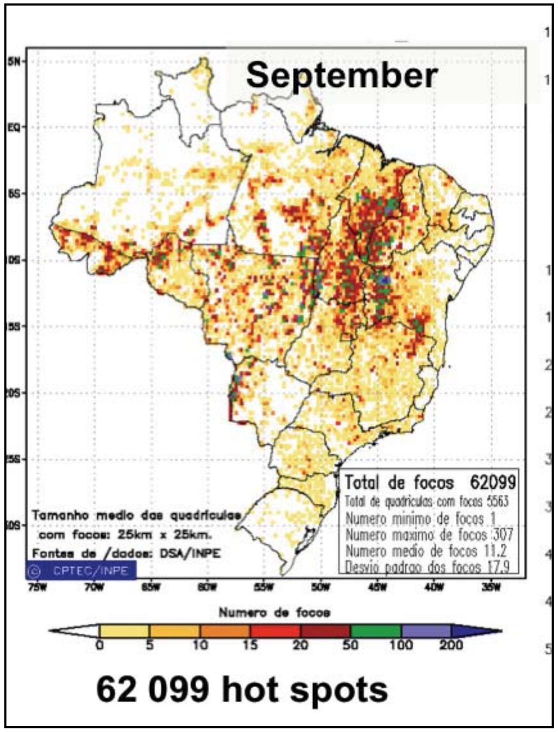
Wind et al., 2013, GMD

Case Studies

from WGNE Forecasting Exercise

1) Dust over Egypt: 4/2012

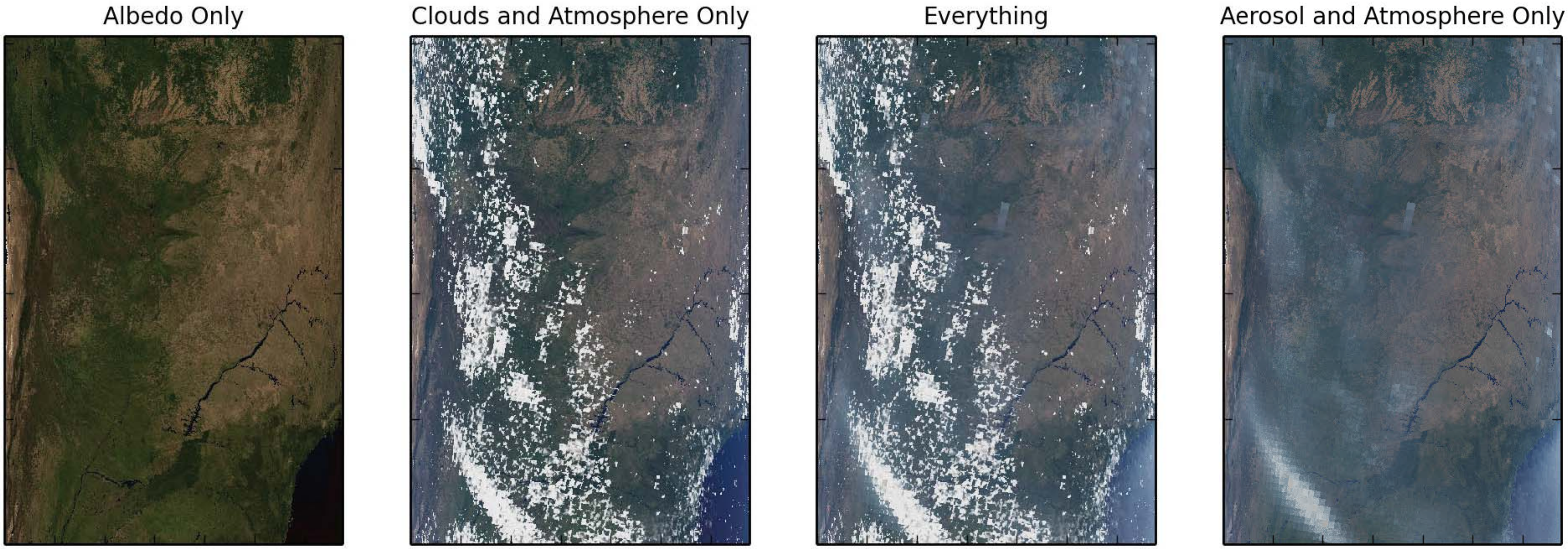
2) Pollution in China: 1/2013



3) Smoke in Brazil: 9/2012

Brazilian Smoke Case

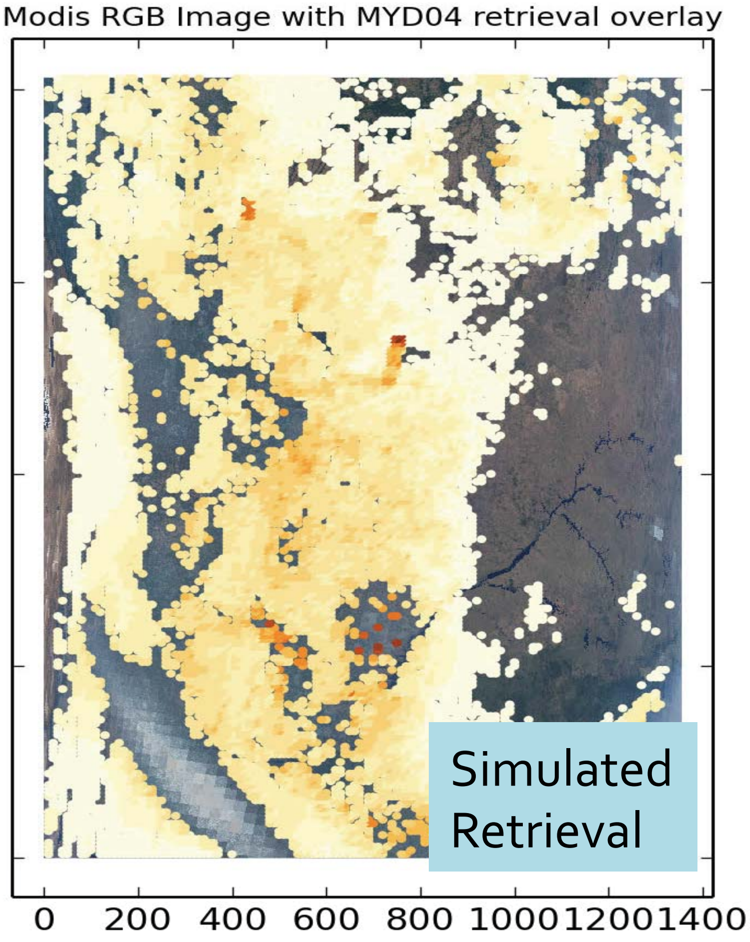
Simulated RGB Images



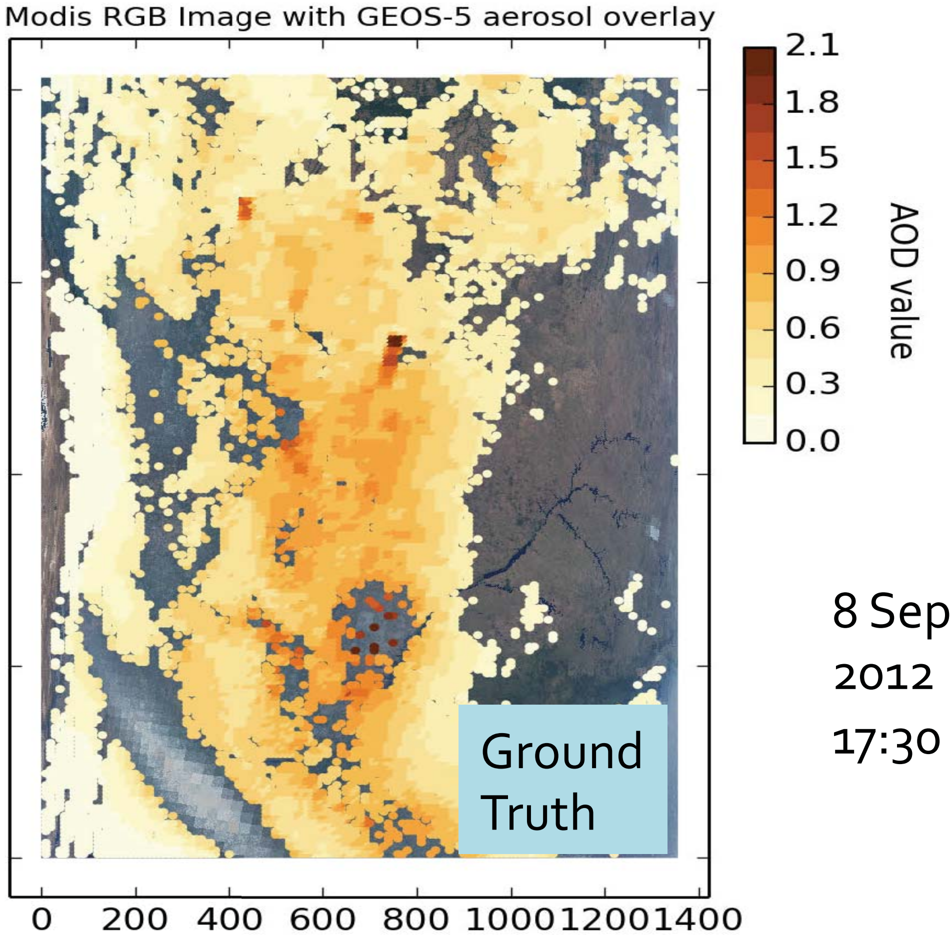
Aqua Granule: 8 September 2012 17:30 UTC

Brazilian Smoke Case

Simulated MODIS Aerosol Retrievals


Modis RGB Image with MYD04 retrieval overlay

Simulated Retrieval


Modis RGB Image with GEOS-5 aerosol overlay

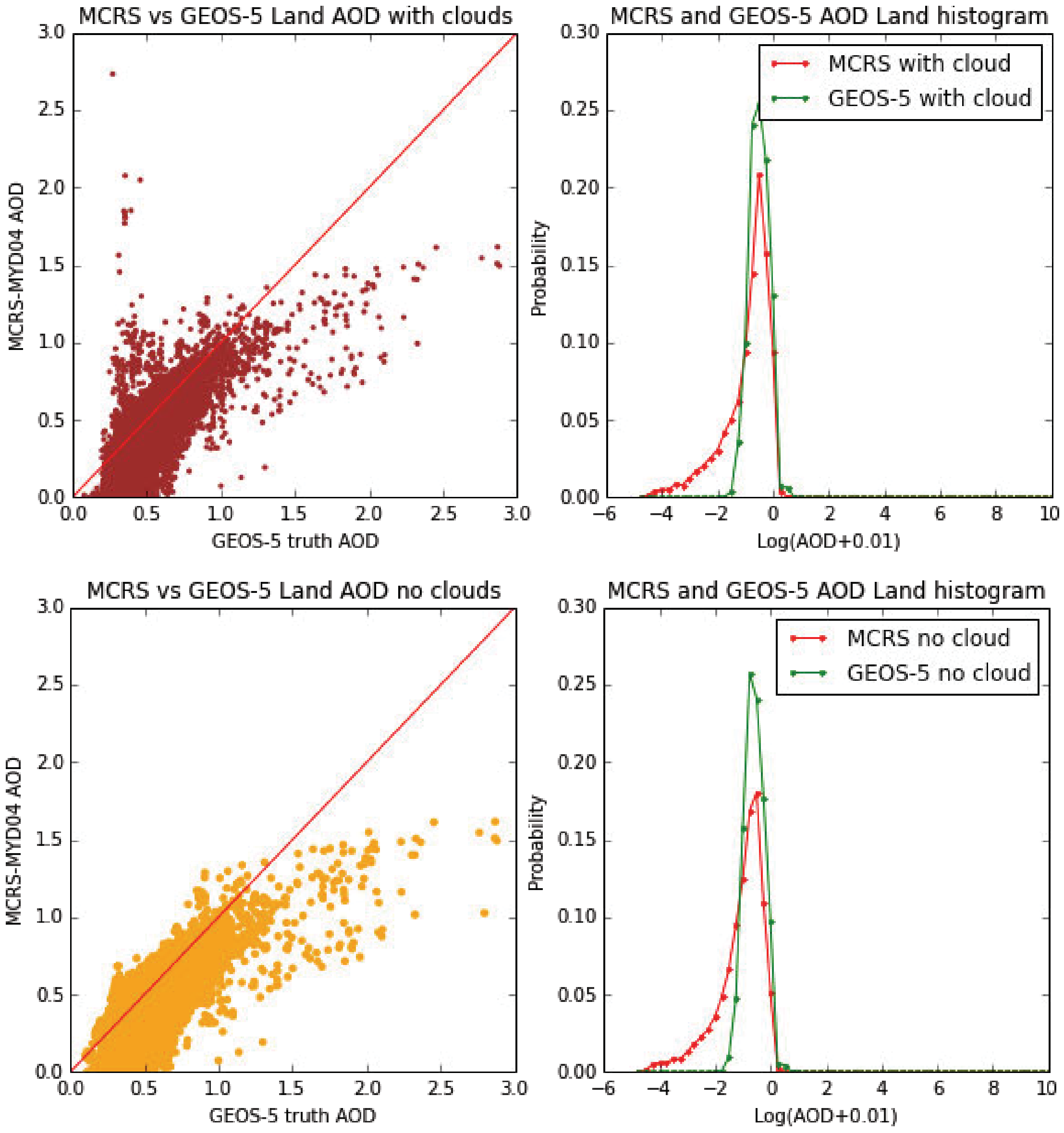
Ground Truth

8 Sept
2012
17:30 UTC

Brazilian Smoke Case

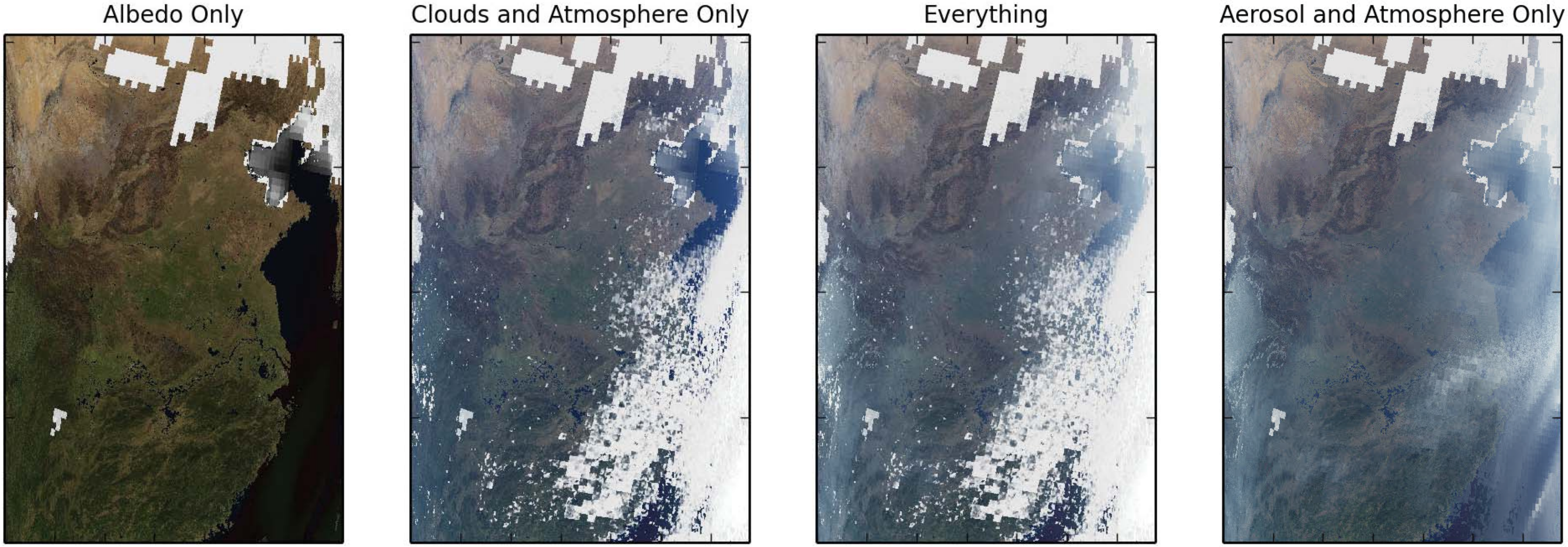
Simulated MODIS Aerosol Retrievals

- Retrievals tend to underestimated AOD
- Cloud contamination while present is not a major factor
- In the no-cloud simulation, surface inhomogeneity leads to data rejection
- Working with MODIS aerosol team on detailed diagnosis



Chinese Pollution Case

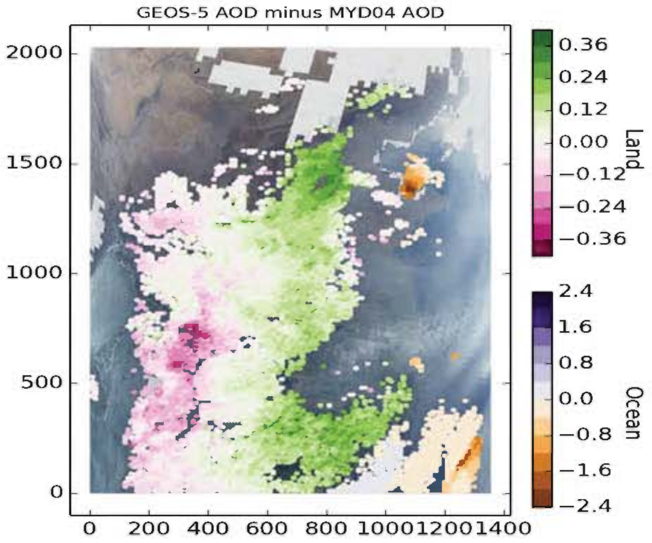
Simulated RGB Images

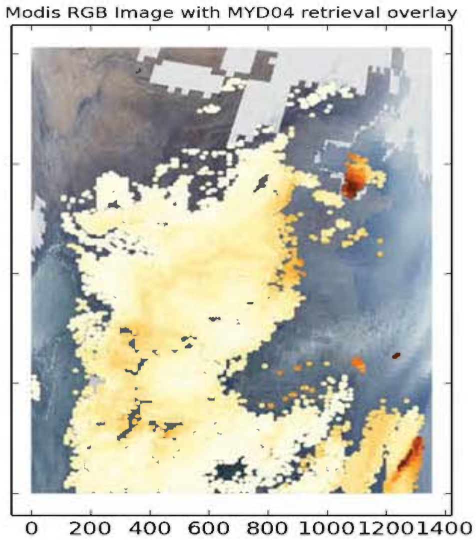


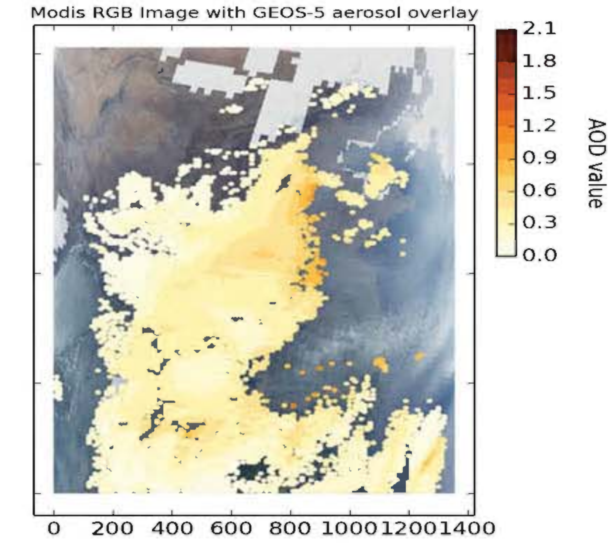
Aqua Granule: 13 January 2013 05:30 UTC

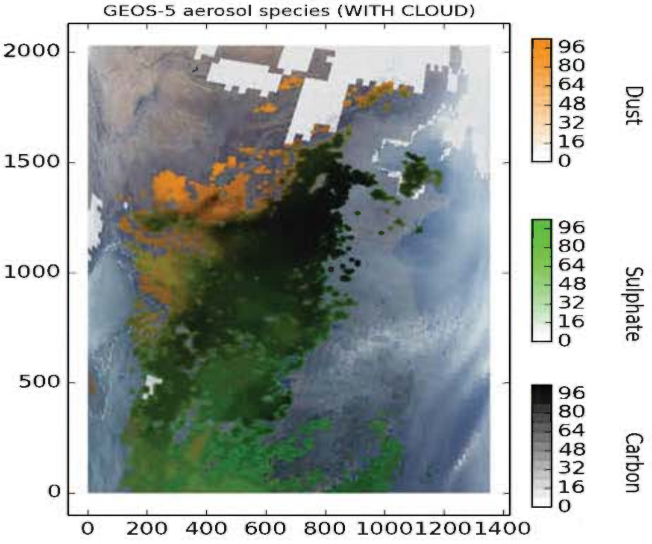
Chinese Pollution Case

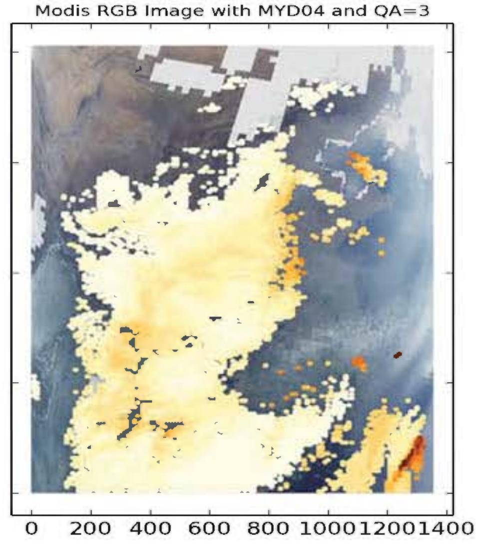
Simulated MODIS Aerosol Retrievals

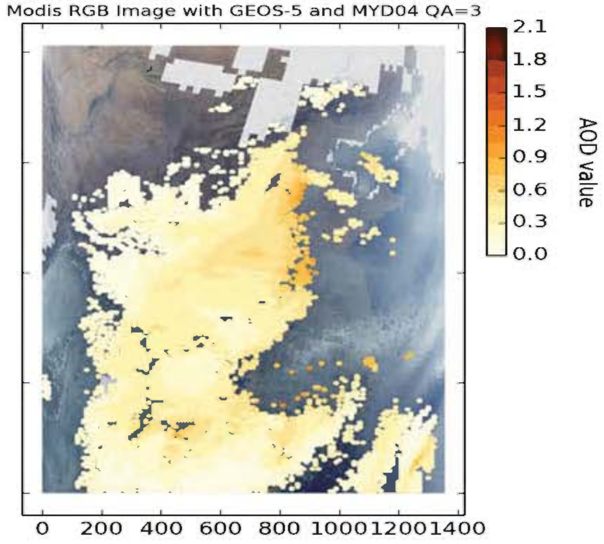

GEOS-5 AOD minus MYD04 AOD


Modis RGB Image with MYD04 retrieval overlay


Modis RGB Image with GEOS-5 aerosol overlay


GEOS-5 aerosol species (WITH CLOUD)

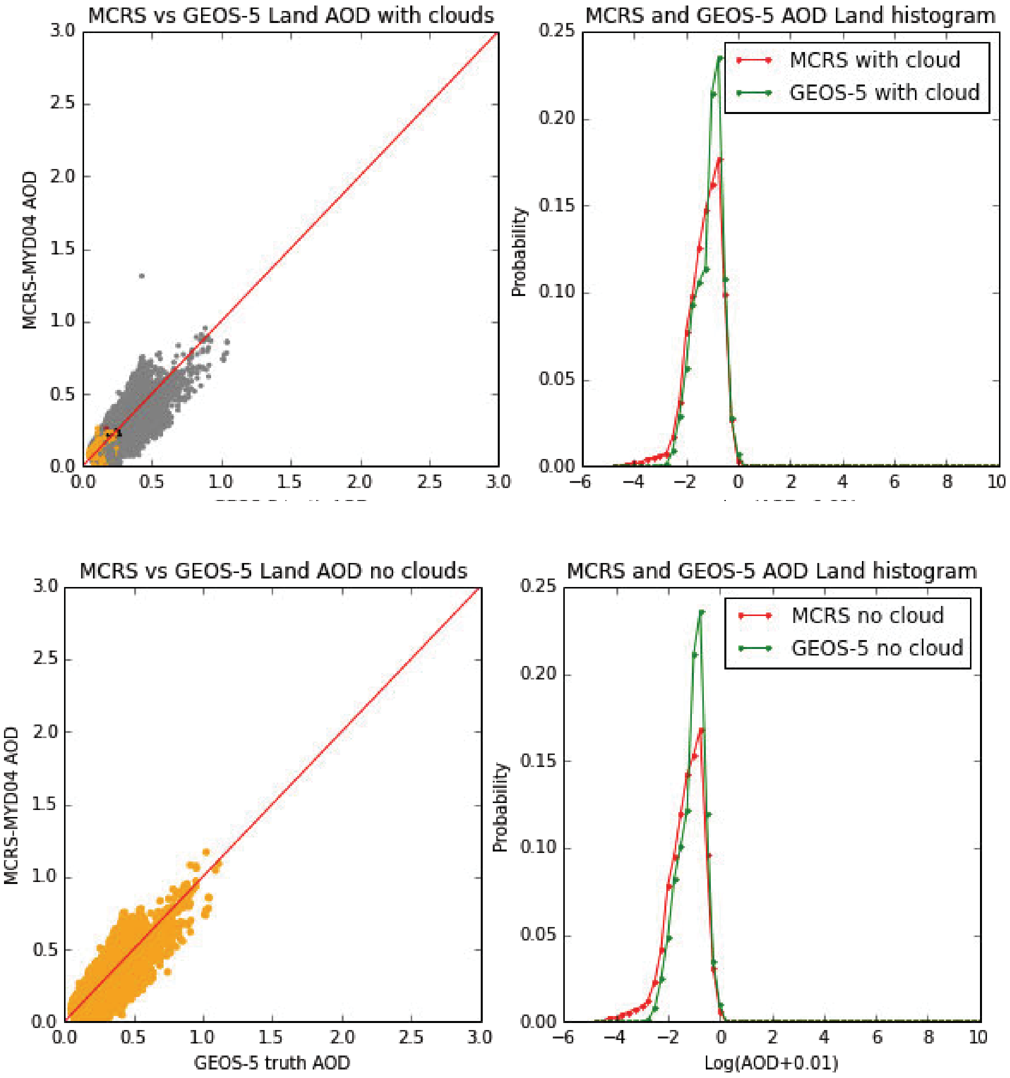

Modis RGB Image with MYD04 and QA=3


Modis RGB Image with GEOS-5 and MYD04 QA=3

Chinese Pollution Case

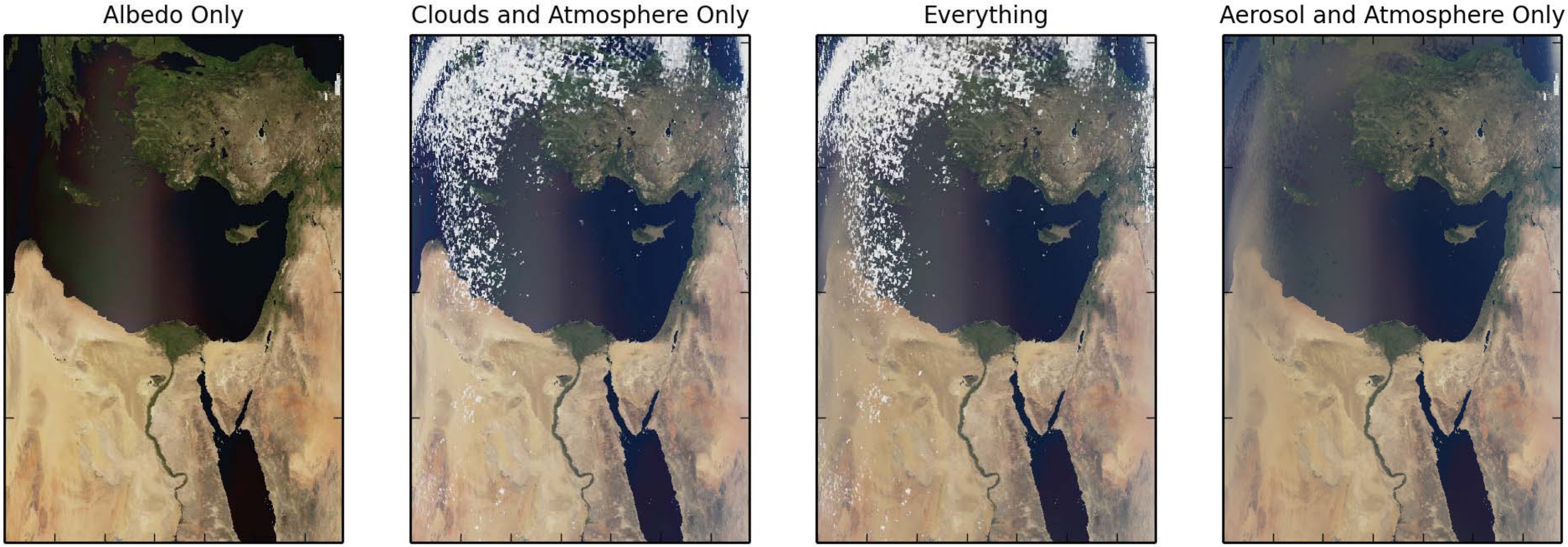
Simulated MODIS Aerosol Retrievals

- On average, relative small bias in retrieved AOD
- However, regional bias associated with species
- Cloud mask very effective in this case



Saharan Dust Case

Simulated RGB Images



Albedo Only

Clouds and Atmosphere Only

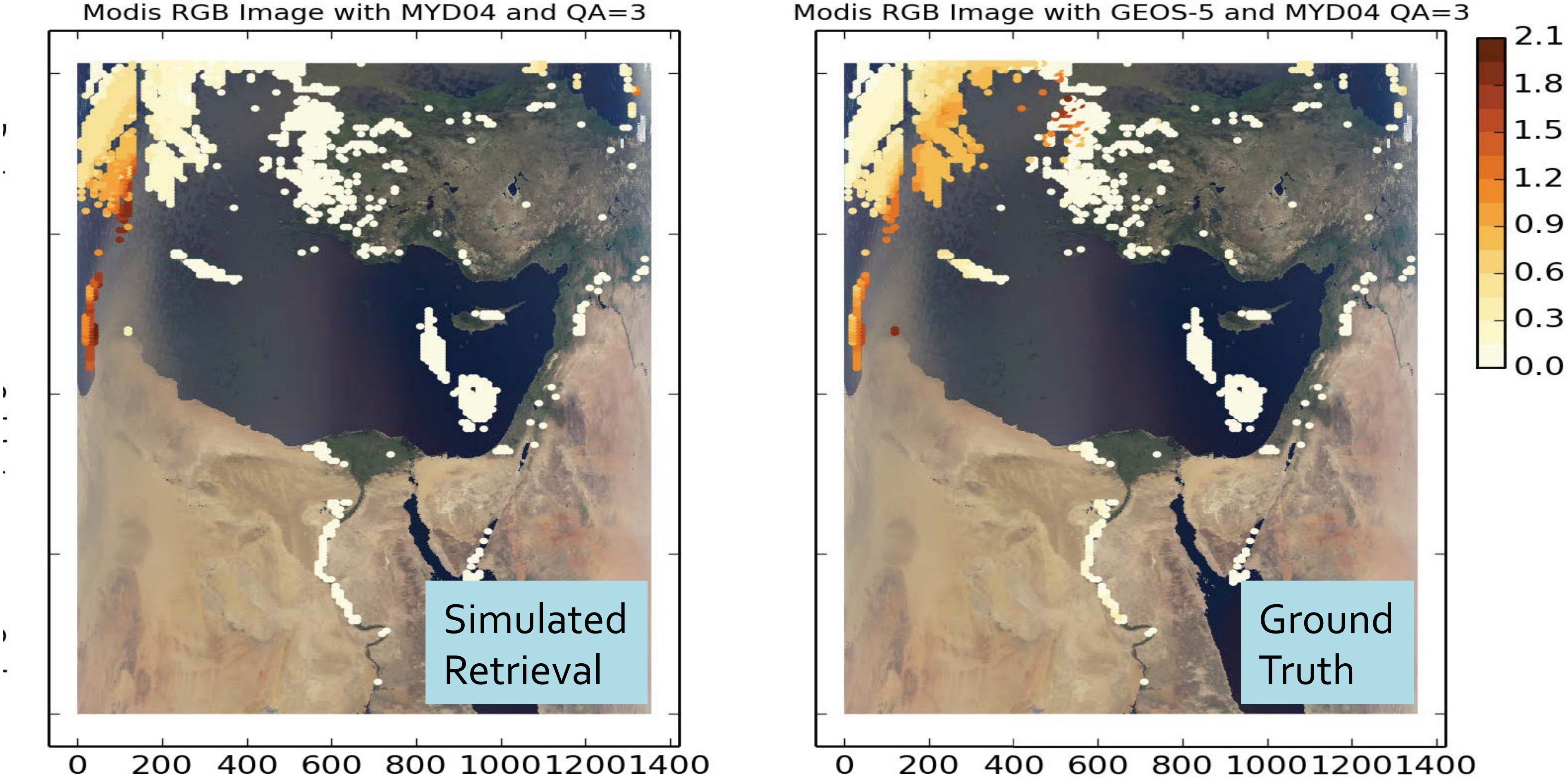
Everything

Aerosol and Atmosphere Only

Aqua Granule: 17 April 2012 11:10 UTC

Saharan Dust Case

Simulated MODIS Aerosol Retrievals



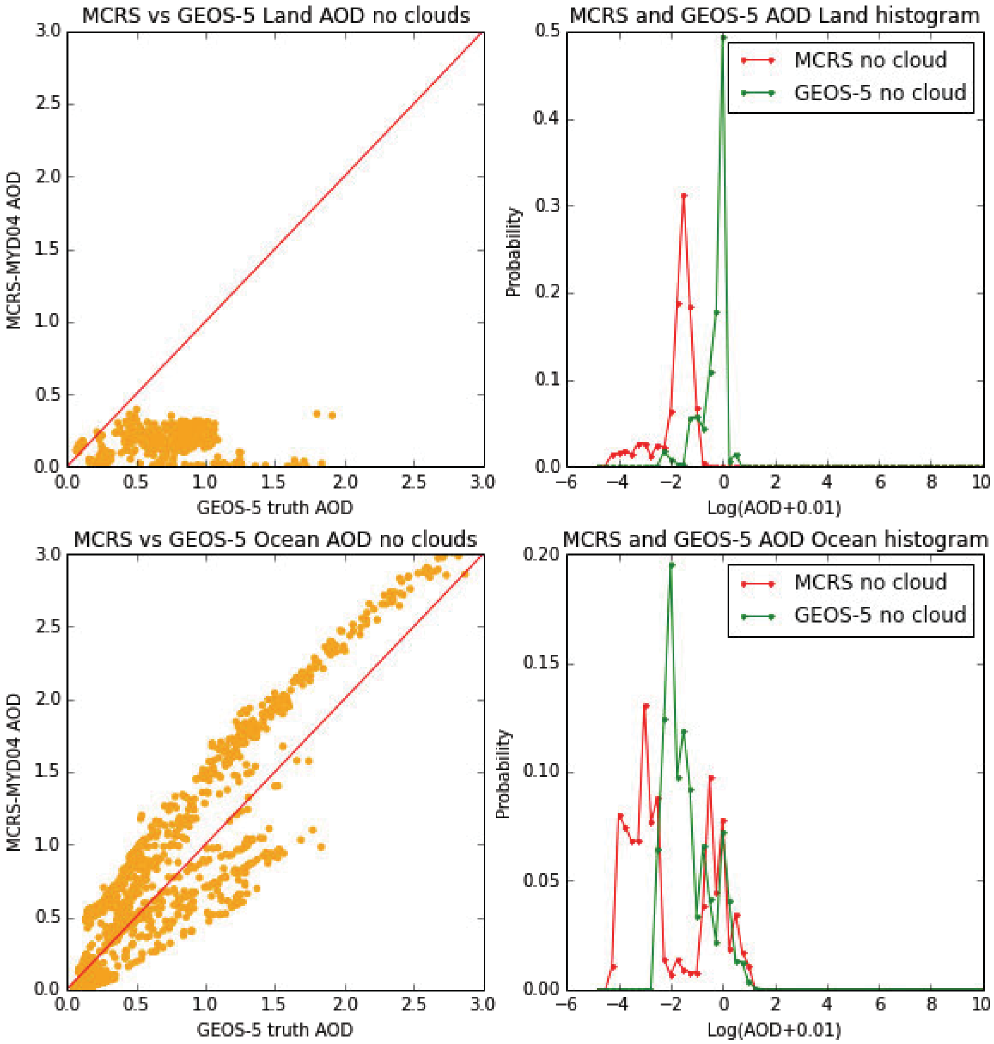
Simulated Retrieval

Ground Truth

Saharan Dust Case

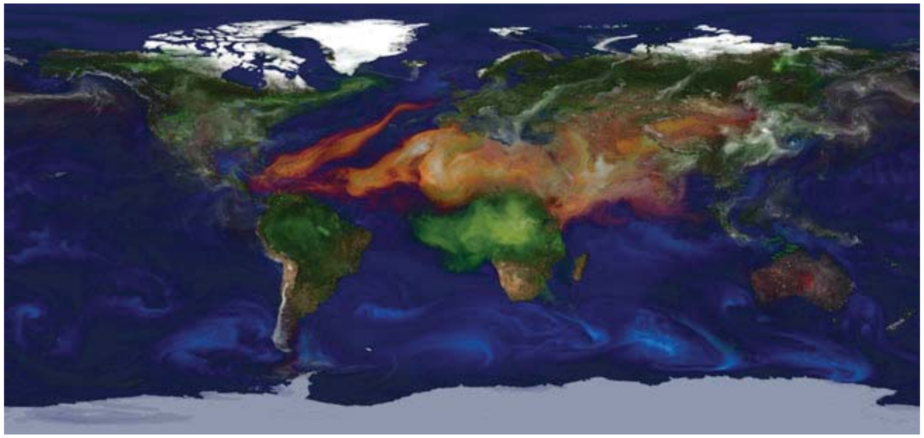
Simulated MODIS Aerosol Retrievals

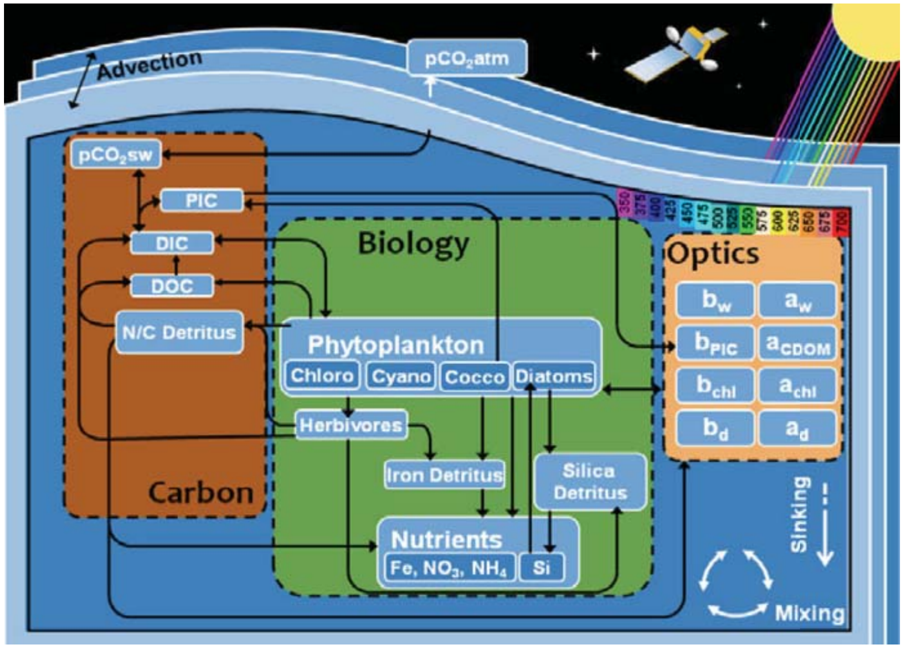
- Low AOD bias over land
- Ocean results: too much or too little
- Cloud contamination is minimum for this case
- Analysis in progress



PACE Simulator

- GEOS-5 non-hydrostatic 7 km atmosphere with GOCART aerosols
- Coupled to GEOS-5 10 km ocean component with biogeochemistry
- Simulation of
  - Water leaving radiances
  - (t.o.a. reflectances)





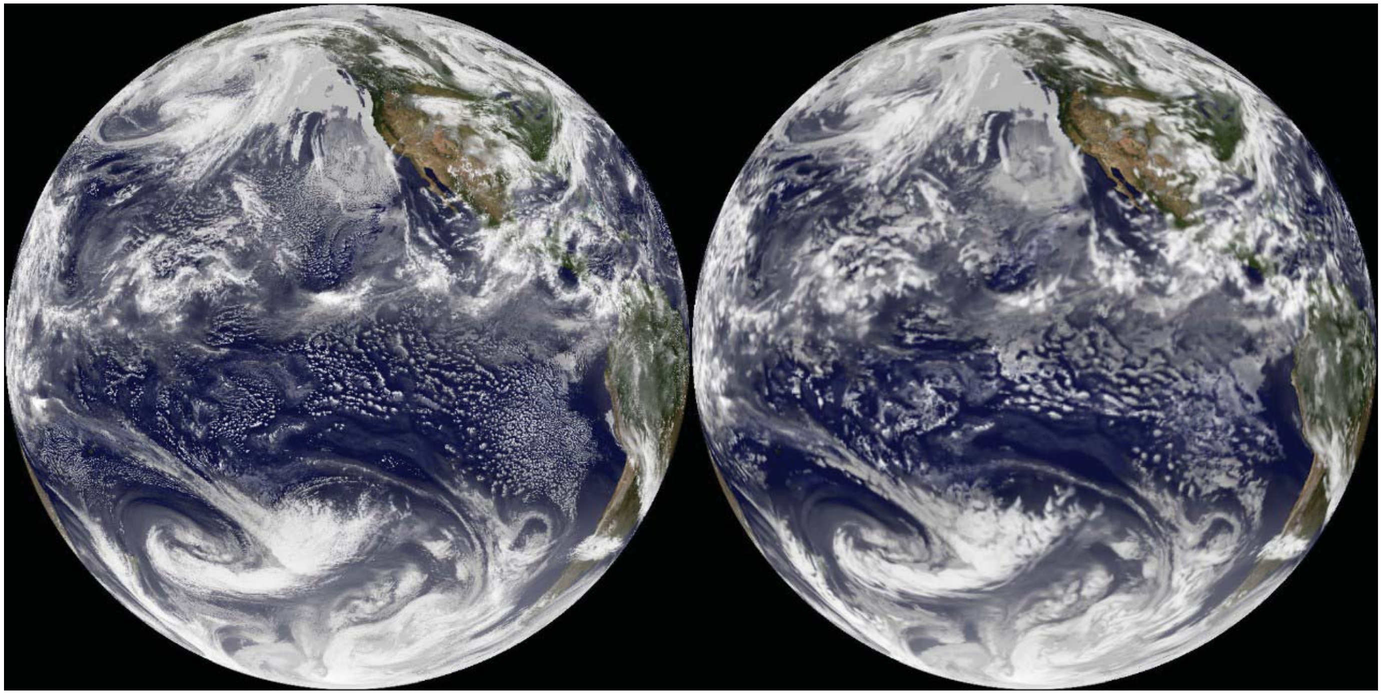
Concluding Remarks

- A *credible* OSSE system requires well validated modeling components:
  - Nature run
  - Physical simulation of measurements
  - *Error modeling*
- However, it must be validated as a *System*, by exercising it with the existing legacy observing system.

Extra Slides

Nature Run Realism

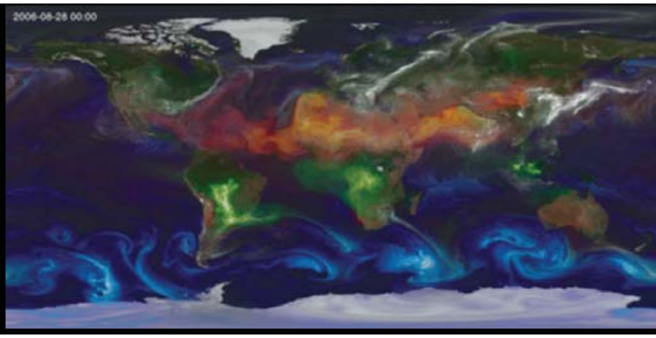
- Global Domain
- Resolutions:



3.5 km

25 km

GEOS-5 Global 7 km Nature Run



- **Components**
  - Atmospheric GCM on cubed-sphere, non-hydrostatic
  - Prescribed SST, sea-ice
  - Constituents
    - Radiatively coupled aerosols
    - Carbon species
    - GMI Combo Chemistry (*)
    - GEOSchem Chemistry
- **Emissions**
  - Prescribed daily biomass-burning emissions (QFED)
  - New dust source function from Ginoux
  - Anthropogenic inventories downscaled to 10km

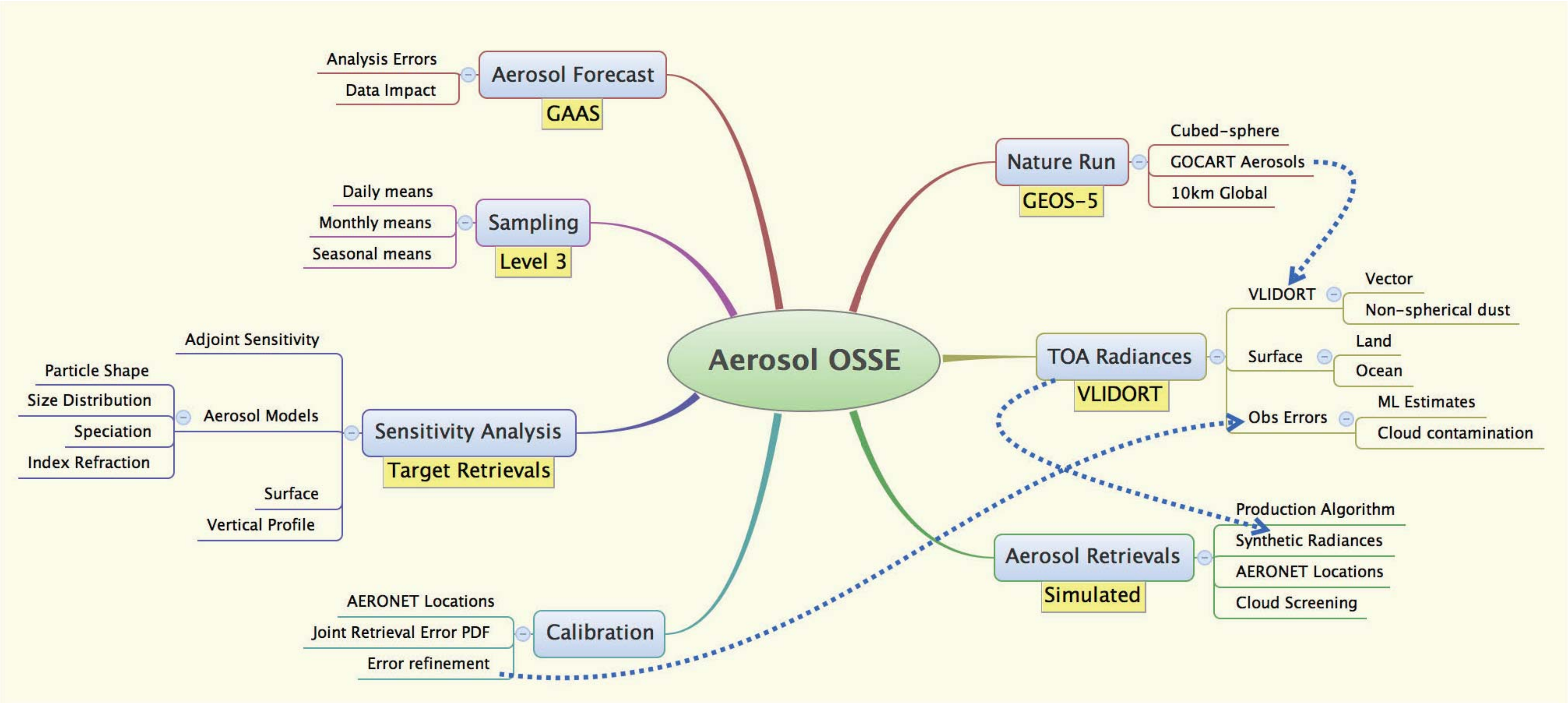
- **GEOS-5 2014 NR**
  - Global, 7 km
  - Aerosol, parameterized Chemistry
    - ~2 years ***simulation***
    - May 2005 – May 2007
  - Aerosol, full chemistry
    - ~ 1 month (TBD)
  - Availability
    - Free, on-line
    - November 2014
- **GEOS-5 2016+ NR**
  - Global, 3.5 km
  - Improved model
  - Cloud-aerosol microphysics, etc.

(*) GMI combo chemistry used for short experiments only.

MODIS Retrieval Simulator

- GEOS-5 output sampled at MODIS Aqua geometry
- Surface albedo from MCD43GF Boston University operational dataset, available from MODAPS
- Ocean surface – MODIS Data Collection 6 Cox-Munk
- 24 spectral channels
- DISORT-5 simulation core
- Correlated-*k* transmittance model
- Rayleigh scattering included
- L1B radiance produced as if Aqua passed over
- IFF and MYD021KM files provided and IFF-compatible aerosol retrieval code executed over IFF files
- H-G phase function for aerosols, but g, ω_0 and τ from GEOS-5

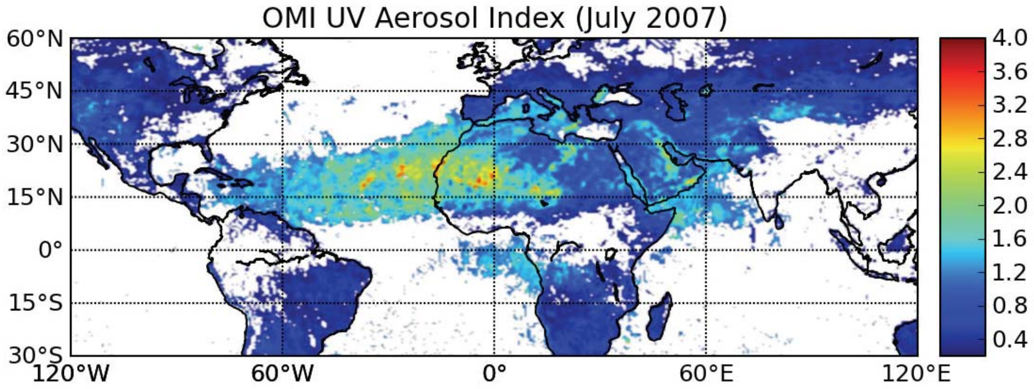
Aerosol OSSE Strategy

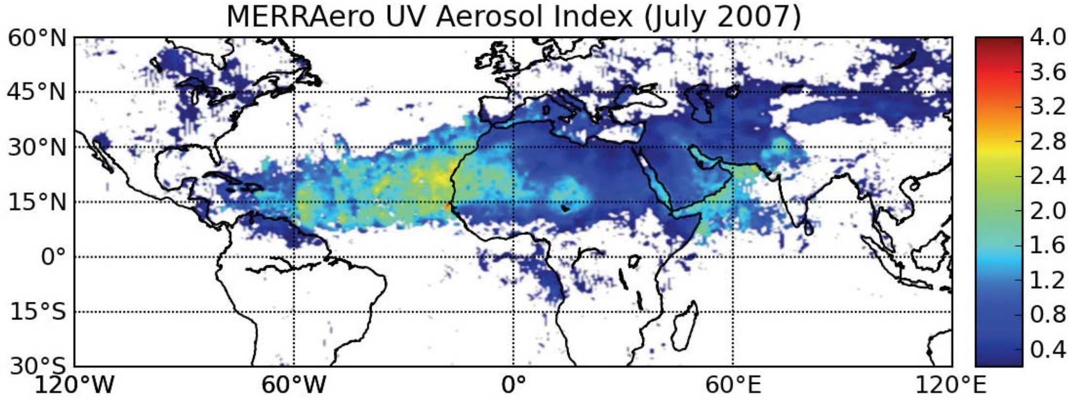


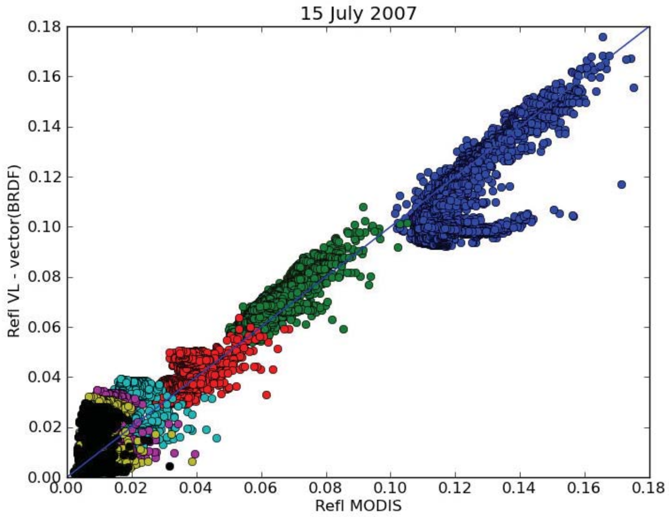
WGNE cases

- 5 cases from different areas dominated by specific aerosol type
- Sahara dust (1 granule)
- Biomass burning in Brazil (2 granules)
- Pollution in China (2 granules)
- Full GEOS-5 input files available for every pixel
  - Contain atmospheric profile, cloud, aerosol information and other valuable data needed to start the simulation
- Radiance data available in following modes as IFF or MYD021KM files for all cases:
  - Albedo only, no atmosphere, cloud or aerosol
  - Clouds and atmosphere only, no aerosol
  - Aerosol and atmosphere only, no clouds
  - All constituents present: atmosphere, clouds and aerosol

Level-1 Simulator



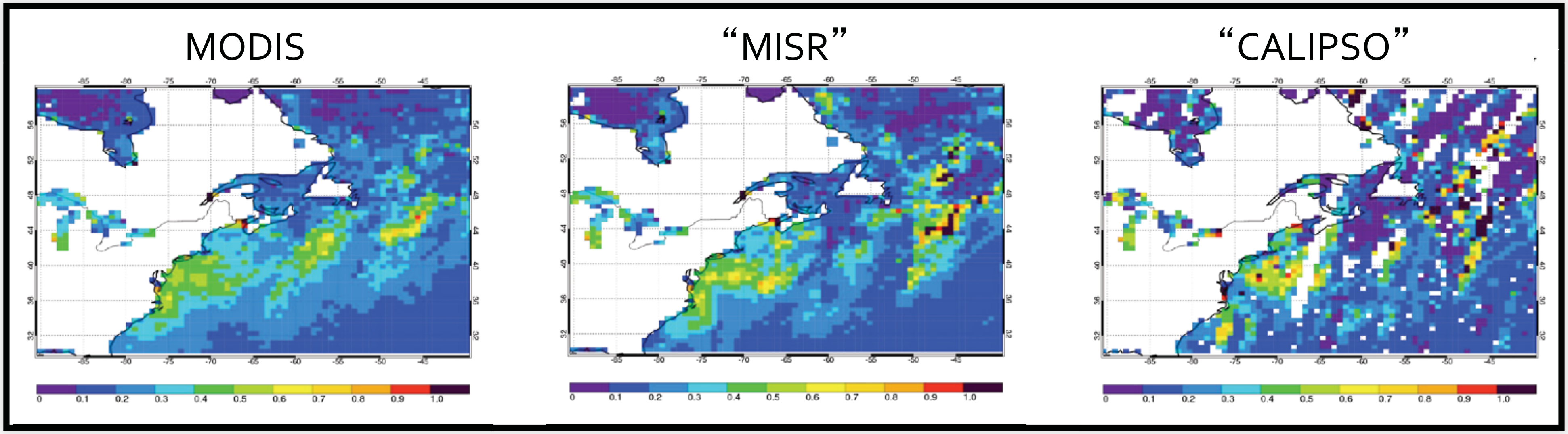




Sampling Studies
Level 2/3 Simulators

Three snapshots of a significant aerosol event:

Canadian forest fire smoke across the eastern US July 7, 2002



MODIS

“MISR”

“CALIPSO”

Even after two months of averaging, the picture these different sensors provide of the regional aerosol load may be quite different.