Bayesian constraint of a gridcolumn statistical model of total water content with high-resolution satellite cloud observations

Peter M. Norris[1,2]

peter.norris@nasa.gov

Arlindo da Silva[1]

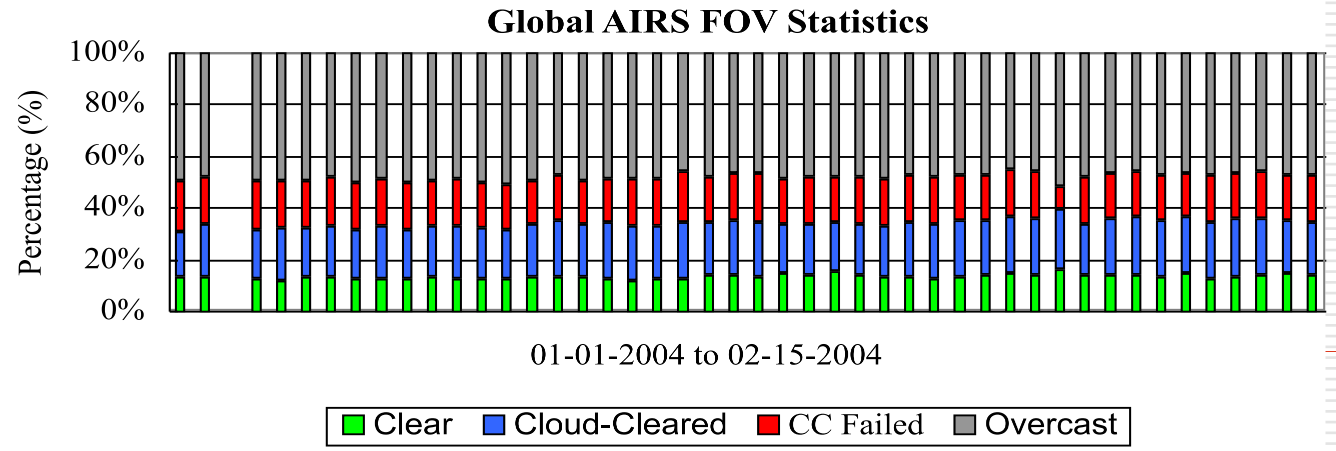
(1) Global Modeling and Assimilation Office, NASA/GSFC
(2) USRA/Goddard Earth Sciences Technology & Research (GESTAR)

NASA Sounder Science Team Meeting
Greenbelt, Maryland, Nov. 13-16, 2012

Motivation

Limited use of satellite radiances due to clouds

- Direct use of cloudy data is currently hampered by complexity and computational expense of cloudy radiative transfer calculations
- Currently, cloud-affected IR radiances are ignored or parameterized by simple gray-body assumption in most operational DA systems, and cloud-affected solar radiances are not assimilated at all.
- Roughly 13% of considered AIRS FOVs are clear, and another 20% can be cloud-cleared successfully… the rest go unused.



Standard 3D-Variational Formulation

$$J(\mathbf{x}_k) = \frac{1}{2}(\mathbf{x}_k - \mathbf{x}_k^b)^T \mathbf{B}^{-1}(\mathbf{x}_k - \mathbf{x}_k^b) + J_c$$
$$+ \frac{1}{2}[\mathbf{h}(\mathbf{x}_k) - \mathbf{y}_k]^T \mathbf{R}_k^{-1}[\mathbf{h}(\mathbf{x}_k) - \mathbf{y}_k]$$

where

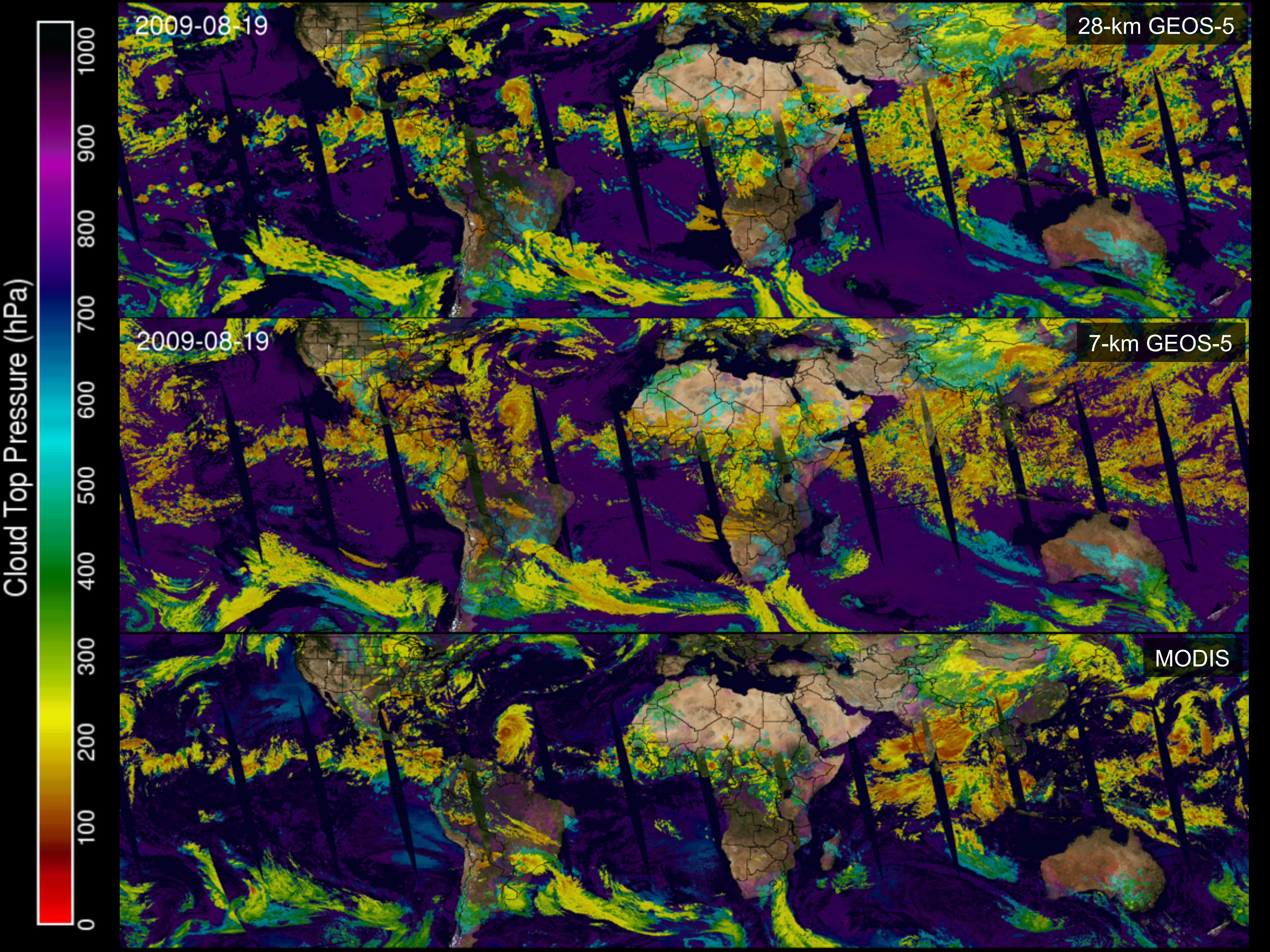
- ▷ $\mathbf{x}_k$ is the 3d state vector (control variable).
- ▷ $\mathbf{y}_k$ and $\mathbf{x}_k^b$ are the observation and background state vectors, respectively.
- ▷ $\mathbf{h}_k$ is the nonlinear observation operator.
- ▷ $\mathbf{B}$ and $\mathbf{R}_k$ are the background and observation error covariances, respectively.

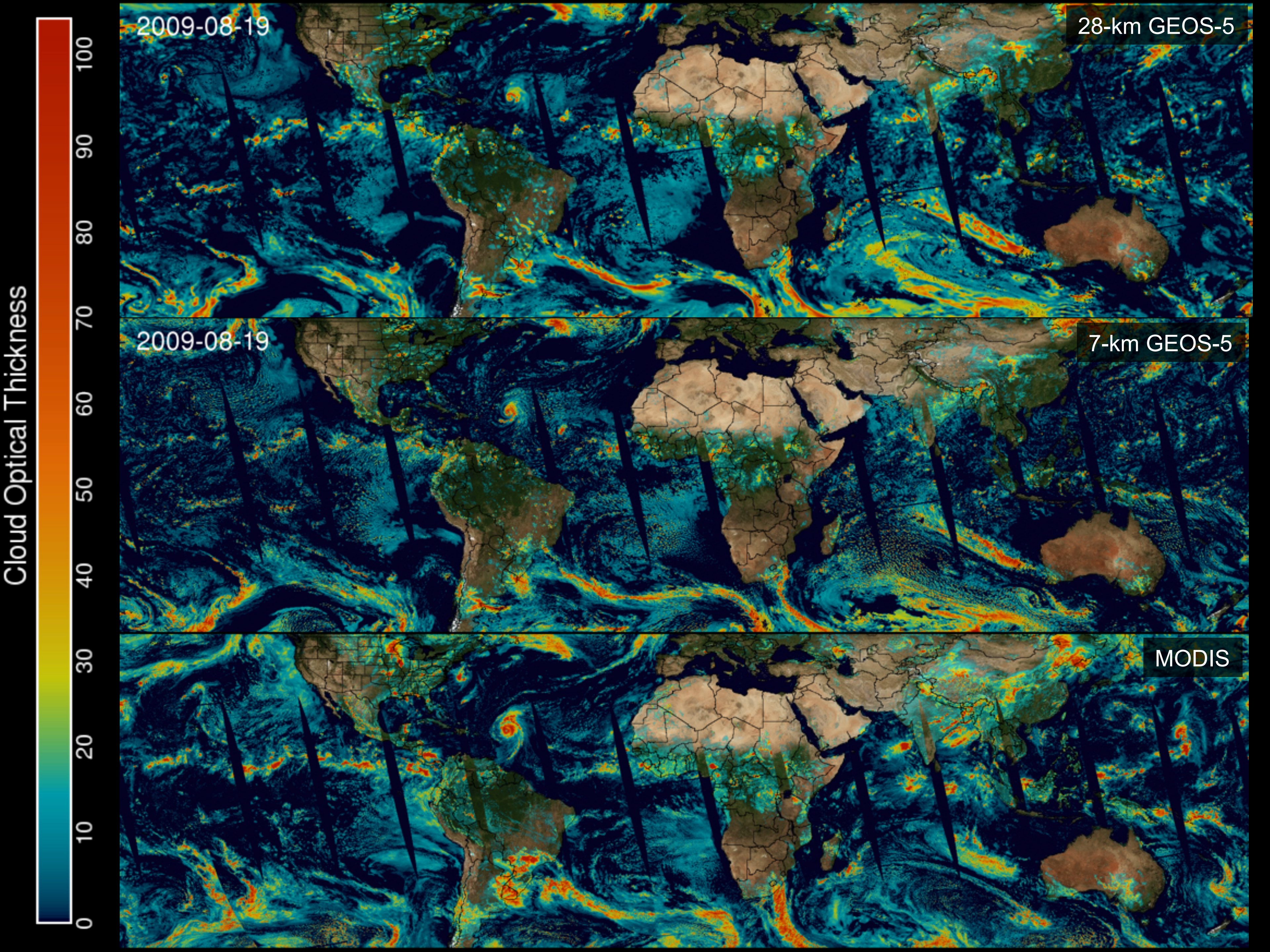
1. Cloud properties are not smooth like temperature and have strong subgrid-scale variability.

2. Cloud observables are not tangent-linear sensitive to *equilibrium* perturbations in the control total water vector for sub-saturated (clear) grid-columns.

Cloud Data Assimilation

- We cannot simply insert clouds in the model
  - We need to convince the model to make clouds.
- Data retention requires a high degree of consistency in cloud representation by GCM and assimilation algorithms.
- Improved background cloud distributions are essential for effective assimilation of cloudy radiances in 3D/4D Var.
- Validation: CloudSat, CERES, SRB.


Cloud Top Pressure (hPa)
0
100
200
300
400
500
600
700
800
900
1000
2009-08-19
28-km GEOS-5
2009-08-19
7-km GEOS-5
MODIS

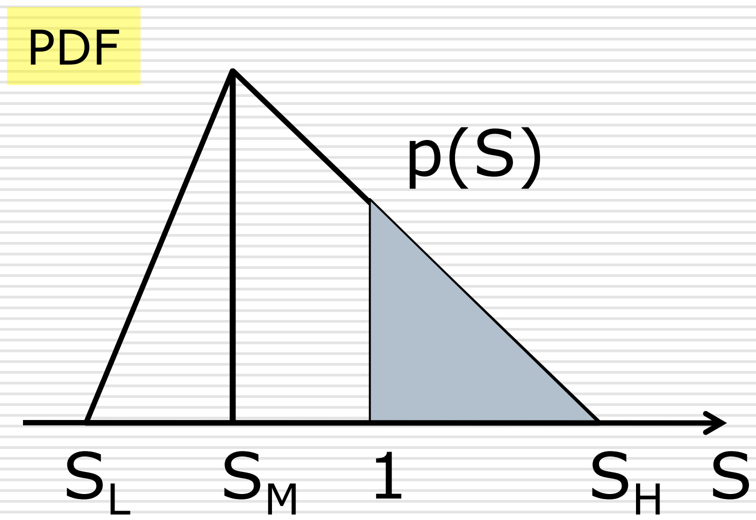


Cloud Optical Thickness
0
10
20
30
40
50
60
70
80
90
100
2009-08-19
28-km GEOS-5
2009-08-19
7-km GEOS-5
MODIS

Statistical Cloud Schemes

- Many recent GCM cloud parameterizations represent sub-gridscale moisture variability using parameterized probability distributions (PDFs) of total water (vapor +condensate).
- Our goal is to use high resolution satellite data to constrain a grid-column of such total water PDFs.
- Cloud optical depth constrains the column integrals.
- IR brightness temperature and cloud top pressure give some information on vertical distribution of moisture in the column.

Total Water Triangular PDF Single Gridbox

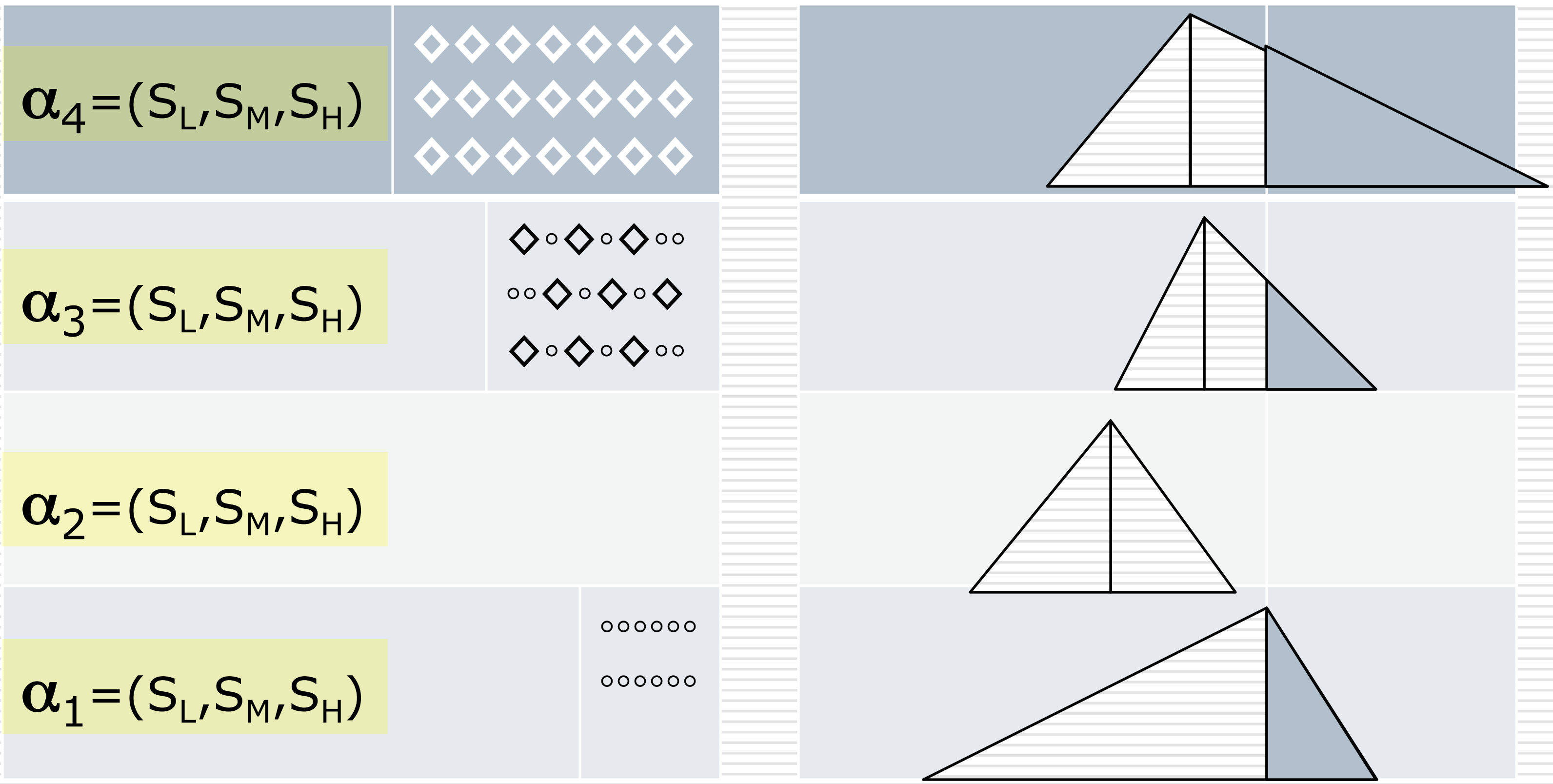


$S = (q_v + q_L + q_I) / q_S(T)$

$p(S) = p(S;\ \boldsymbol{\alpha}=[S_L,S_M,S_H])$

- A simple PDF having skewness
- Given (S_L,S_M,S_H) we can compute
  - c_F, cloud fraction
  - q_V, vapor
  - $q_C = q_L + q_I$
- Conversely, given (c_F,q_V,q_C) we can reconstruct the PDF

Grid Column Statistical Model



- Copula used to couple layer PDFs (Norris et al. 2008)

Multidimensional CDF for a Grid Column

$$P(S_1, ..., S_K) = P(S_1, ..., S_K; \alpha_1, ..., \alpha_K, \beta)$$

where

K	number of vertical layers
S	total water variable
α	PDF parameters for each layer
β	controls vertical coherence of layers

Bayesian Parameter Estimation

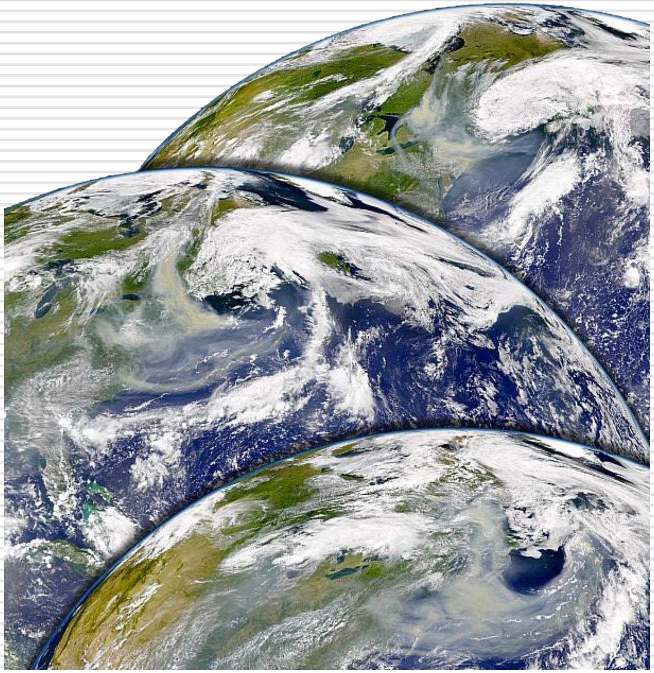
- Within a grid column, consider a set of measurements

  $$\boldsymbol{y} = (y_1,\ldots,y_P)$$

  say MODIS cloud top pressure, cloud optical depth
- Goal:
  - estimate PDF parameters α_k
  - Given the observations $\boldsymbol{y}$
- Bayes theorem:

  $$p(\alpha|\boldsymbol{y}) \sim p(\boldsymbol{y}|\alpha)\, p(\alpha)$$

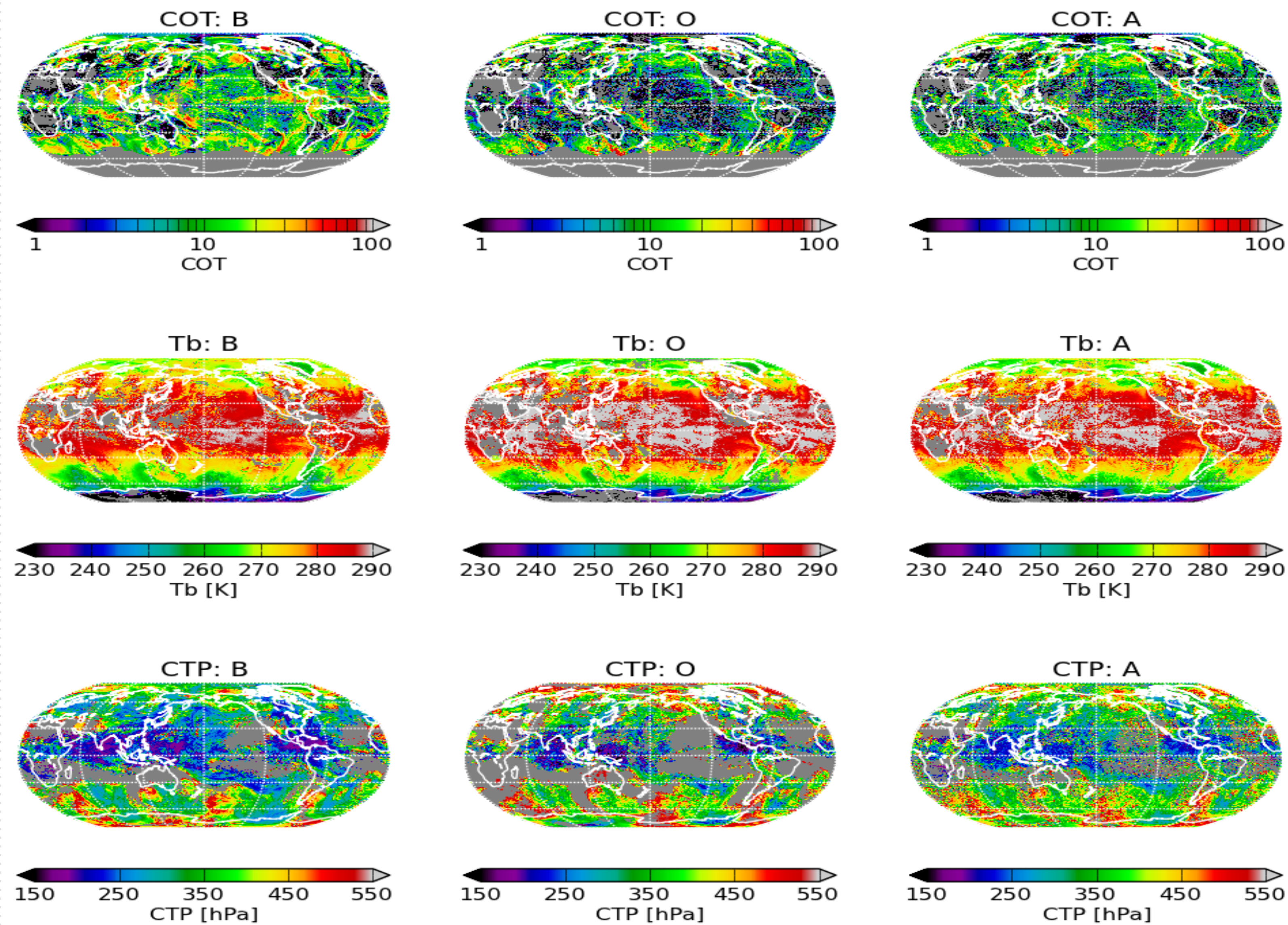
- Maximum *a posteriori* estimation
  - Find α that maximizes $p(\alpha|\boldsymbol{y})$
  - Can also characterize $p(\alpha|\boldsymbol{y})$ form



Parameter Estimation

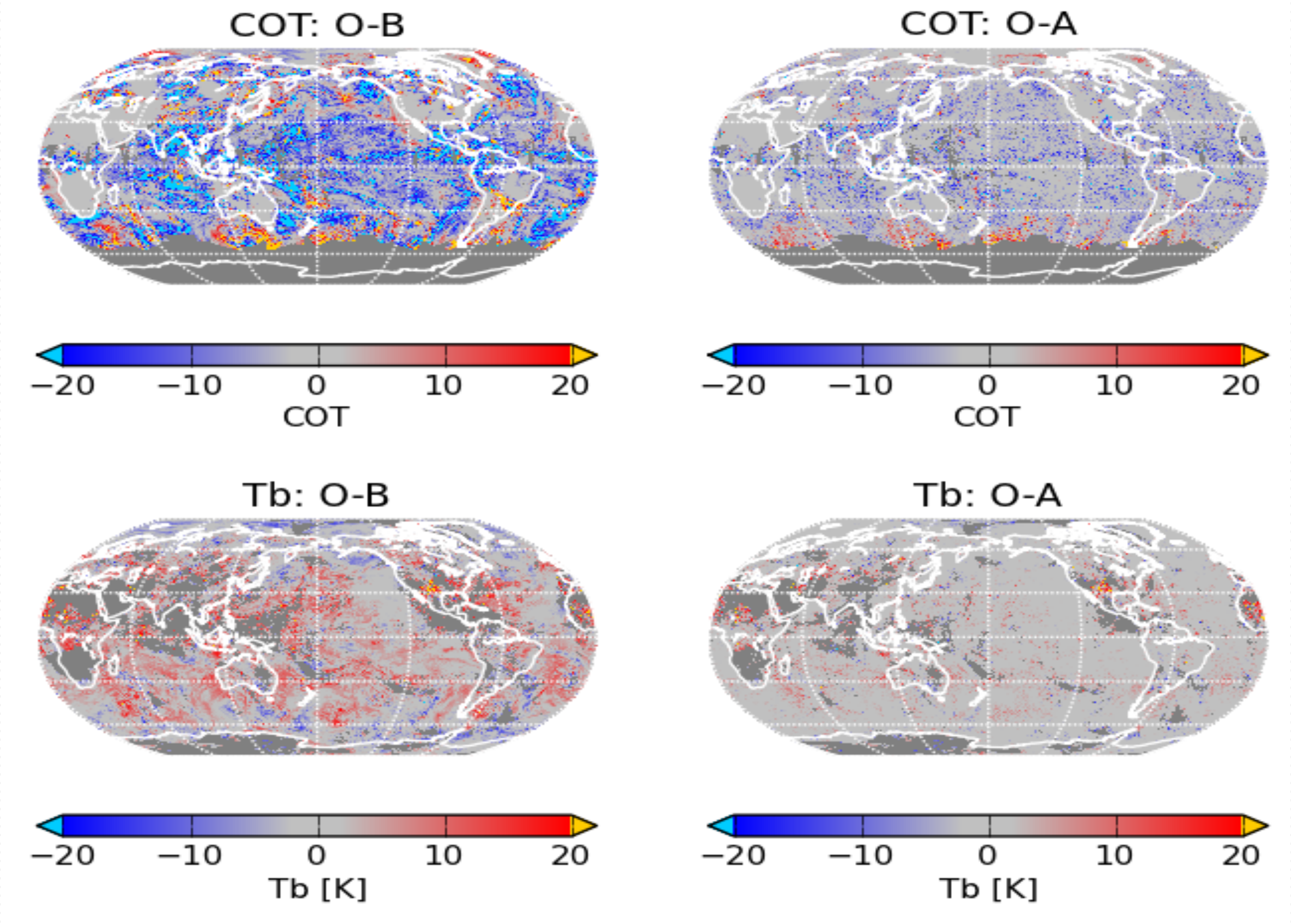
- Evaluating $p(\mathbf{y}|\alpha)$
  - Given α, generate sub-columns by sampling the multidimensional grid-column PDF of S
  - Simulate observables for each sub-column
  - Use these simulated sub-columns to obtain a *Kernel Density Estimate* (KDE) of p at the observational points $\mathbf{y}$
- Optimization
  - Markov Chain Monte Carlo method
    - Modified Metropolis-Hastings algorithm

A one day analysis of MODIS Aqua and Terra

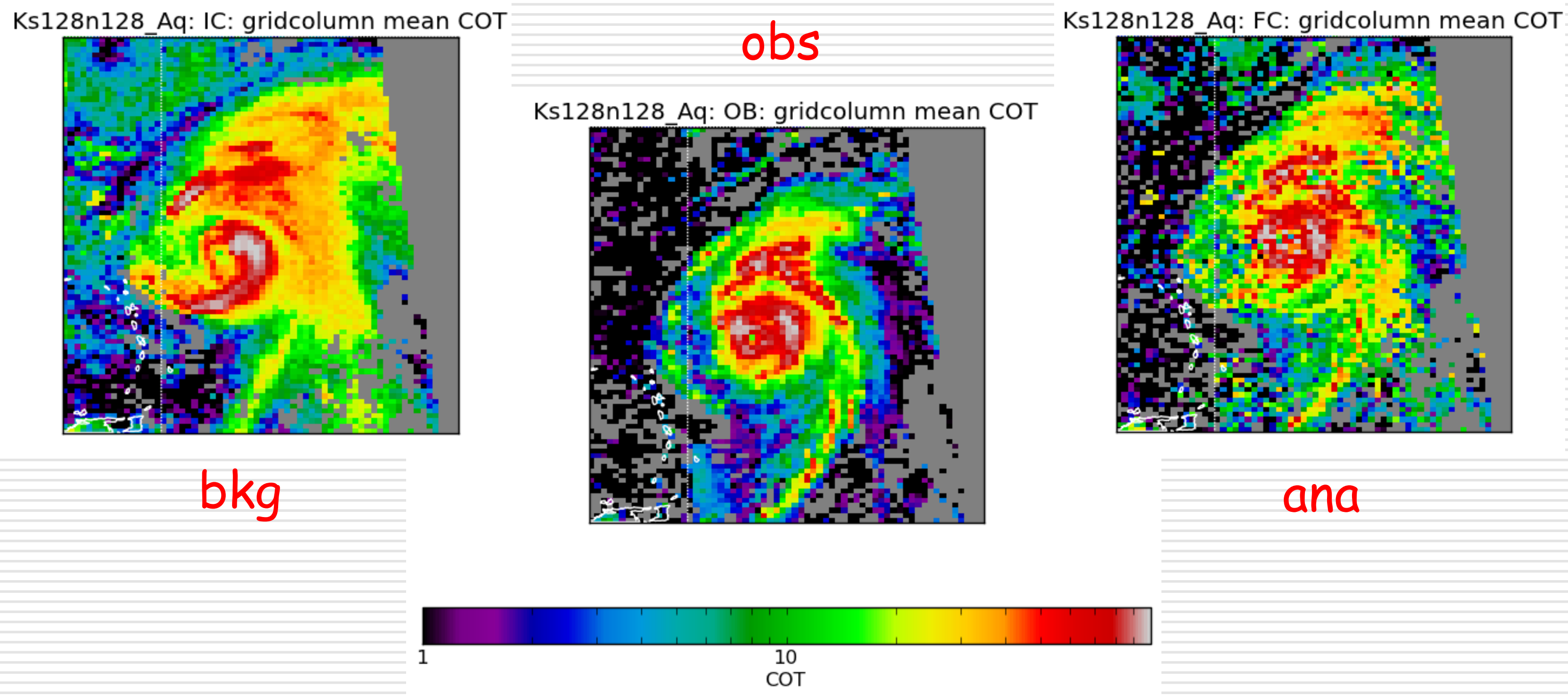


bkg obs ana

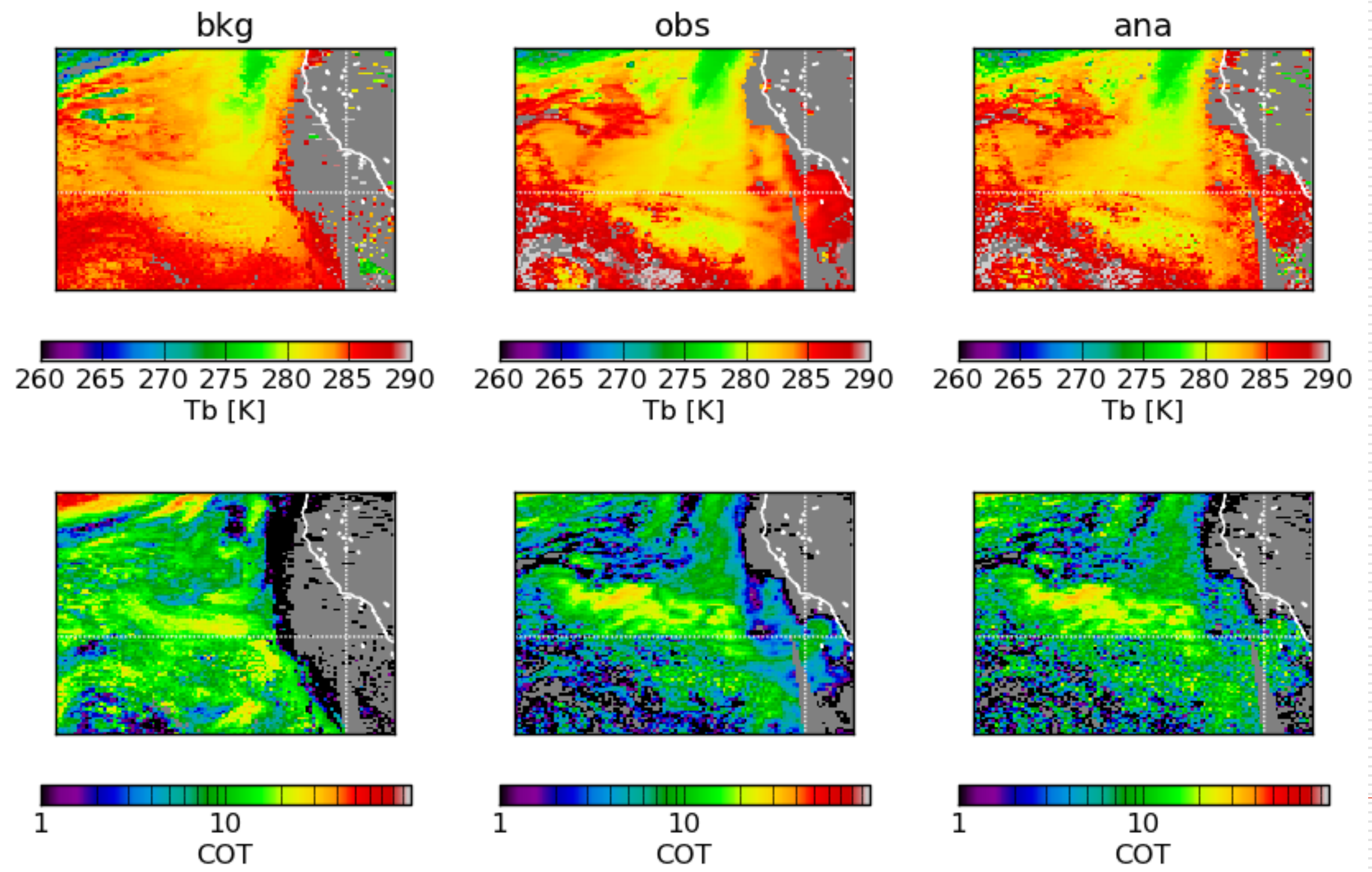
Reduction in Cloud Property Biases



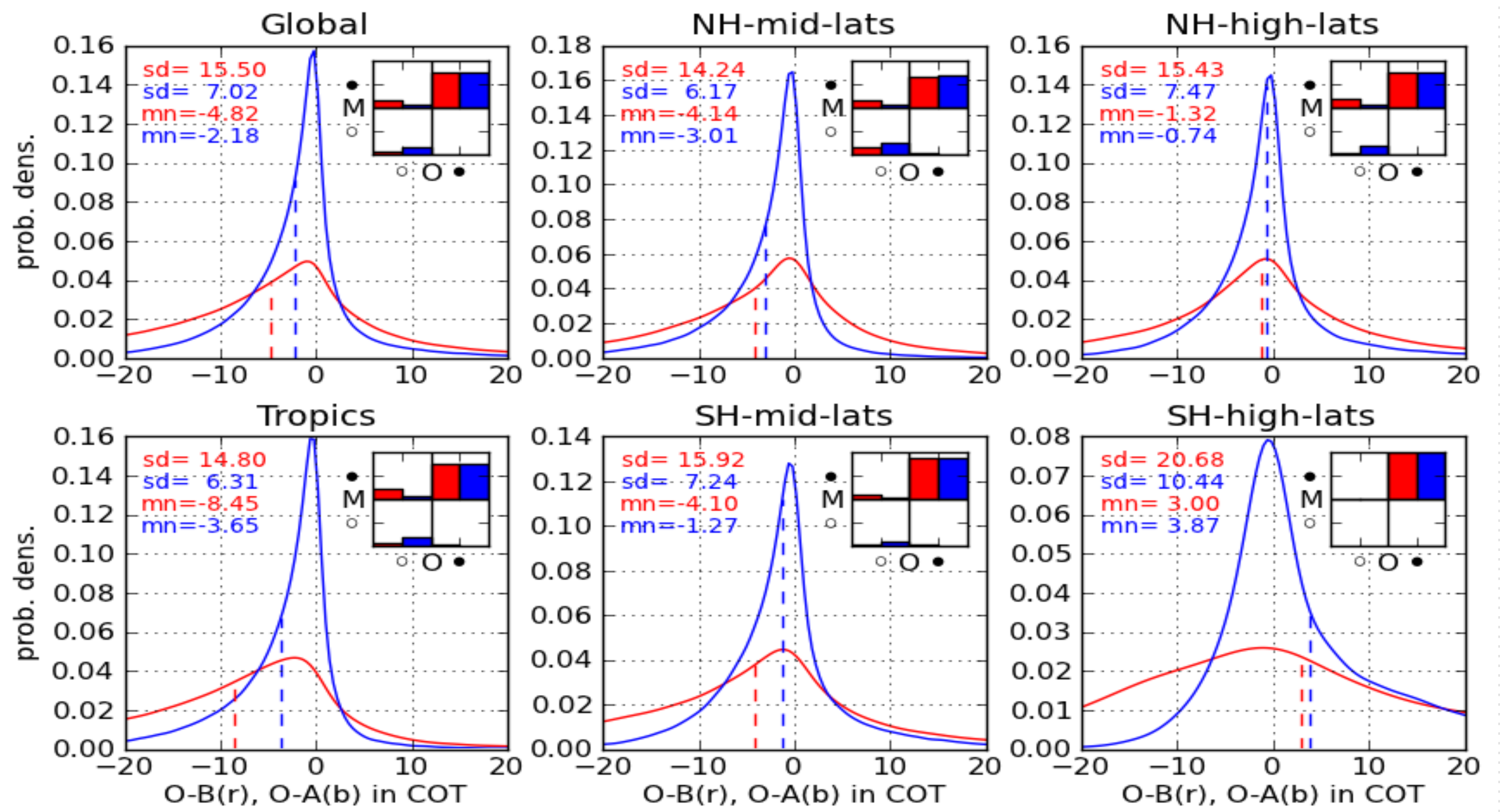
Hurricane Bill



Californian Stratocumulus

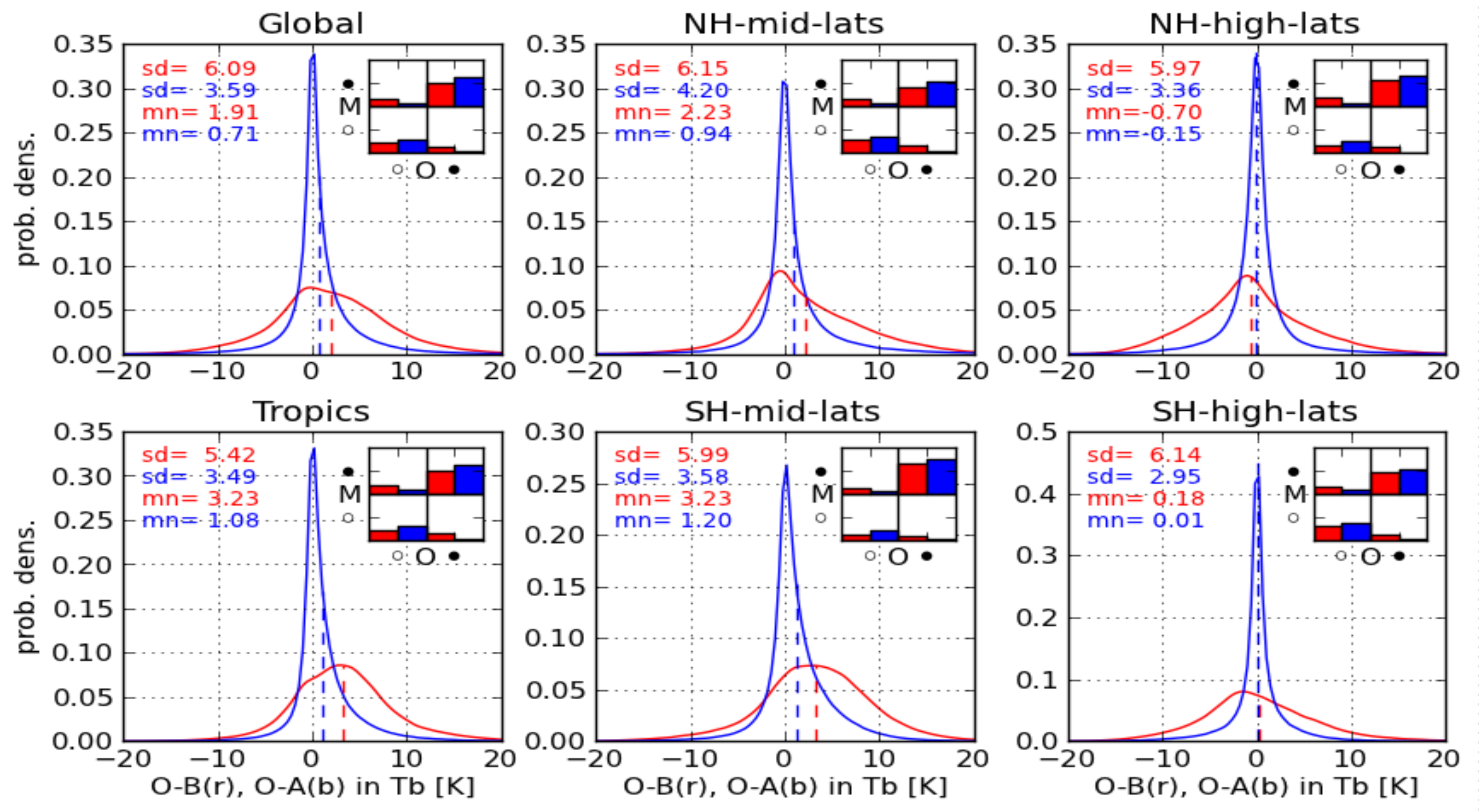


Grid-mean Cloud Optical Depth



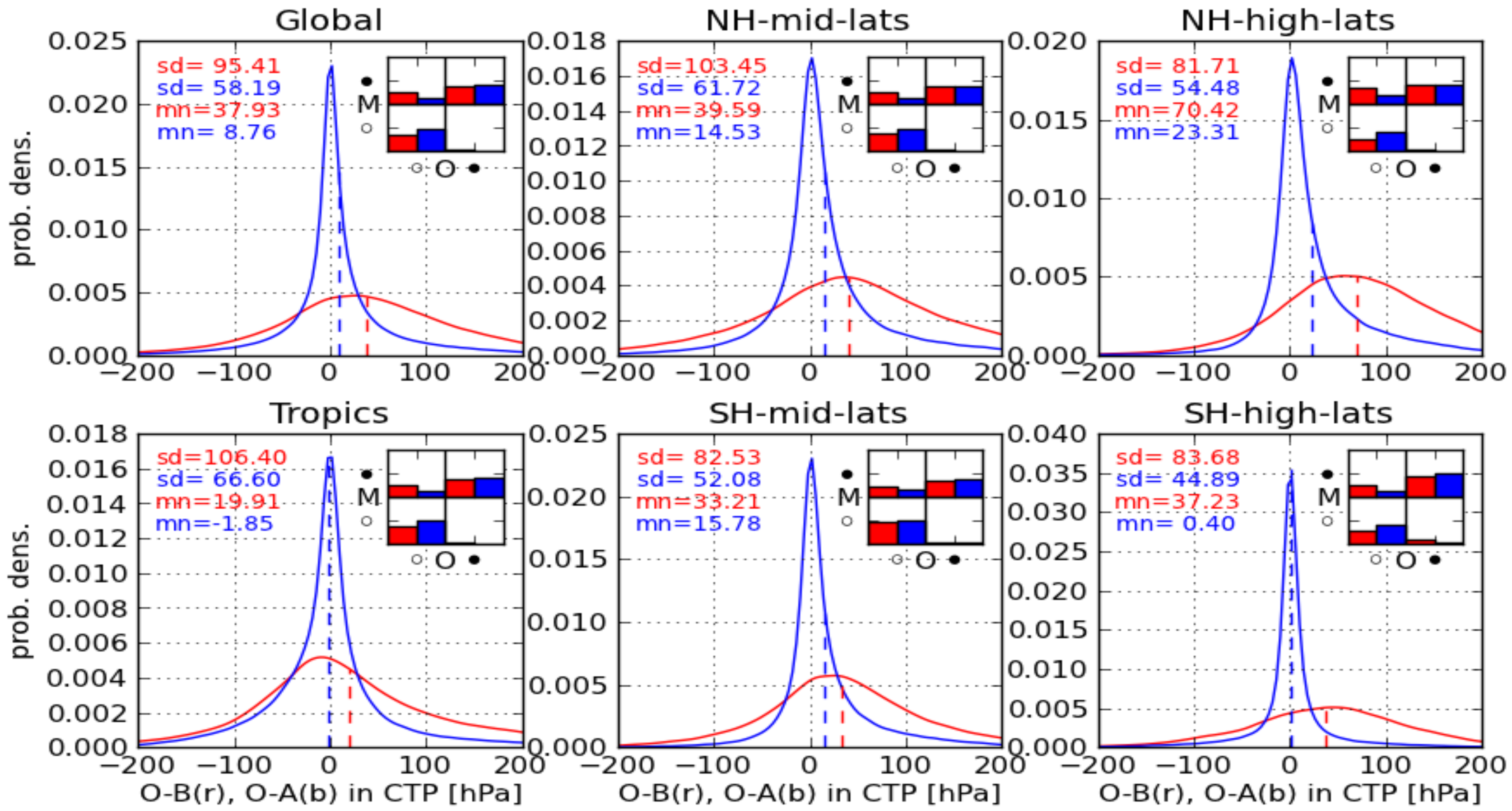
Background, Analysis

Brightness temperature



Background, Analysis
(lower clouds, p > 550 hPa)

Cloud-Top Pressure



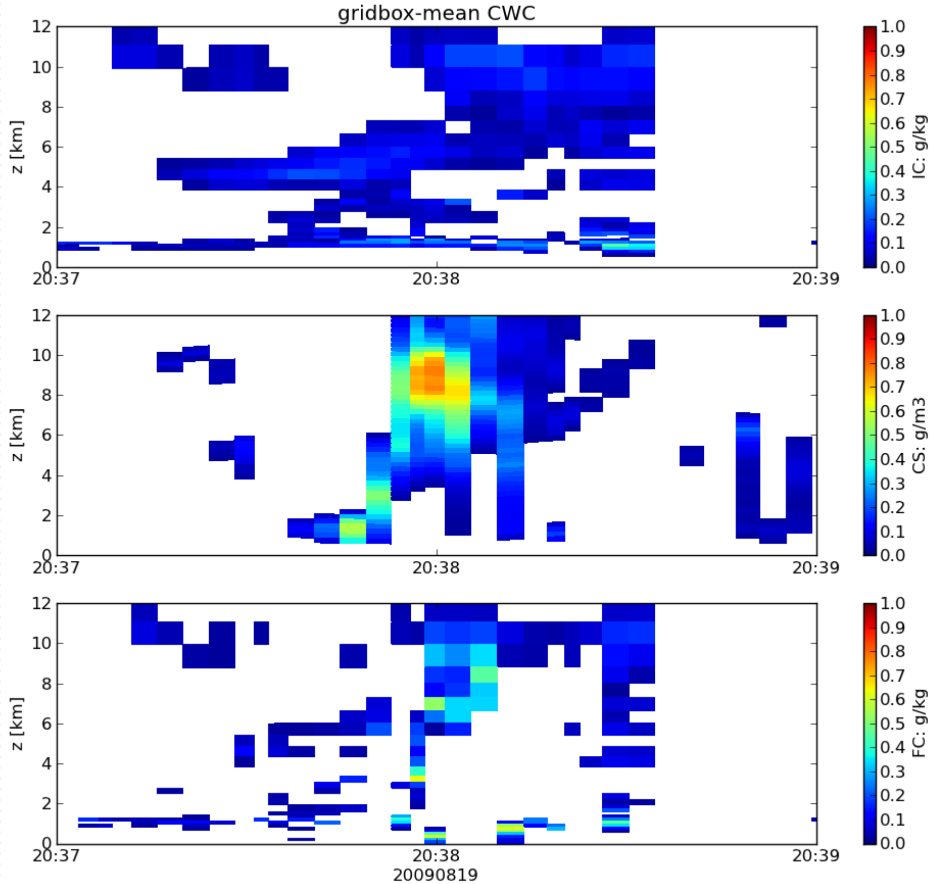
Background, Analysis
(upper clouds, p < 550 hPa)

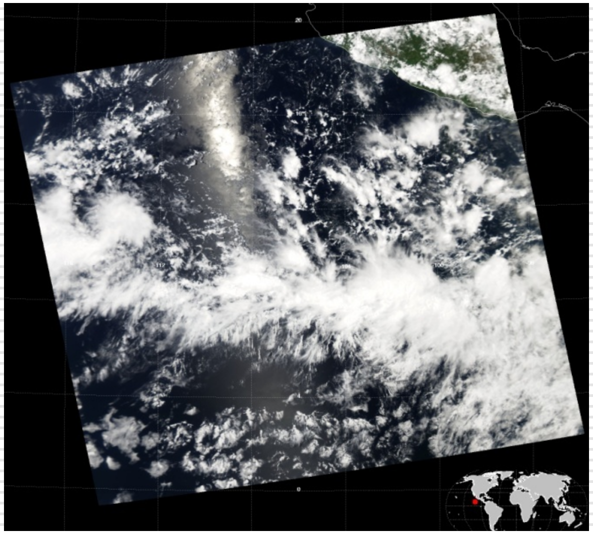
Vertical Cloud Structure -- ITCZ

bkg

Cloud Sat Valid-ation

ana





Concluding Remarks

- Representation of clouds in GEOS-5 is becoming good enough to warrant assimilation of cloudy data
- High resolution MODIS/Geostationary data useful to constrain total water PDFs in GEOS-5
  - "Cloud Bias Corrector" or "Cloud Relocator"
- Improved cloud background are essential for effective assimilation of cloudy radiances in GSI
- Data retention requires high degree of consistency across GCM and assimilation algorithms.
- Speed: now about 15 d/d for SEVIRI.

The End