Uncertainty in forward microwave satellite radiance calculations and what to do about them

Jeff Steward, Ziad Haddad, Svetla Hristova-Veleva, Tomi Vukicevic and Hui Wang

E-mail: jsteward@jifresse.ucla.edu

UCLA

JIFRESSE
Joint Institute for Regional Earth System Science & Engineering

Satellite data

- Satellite data is plentiful
- Can be used to improve NWP
- However, many challenges with using this data:
  - Measures upwelling radiation
  - Limited spatial resolution
  - Non-linear response
  - Uncertain processes + parameters

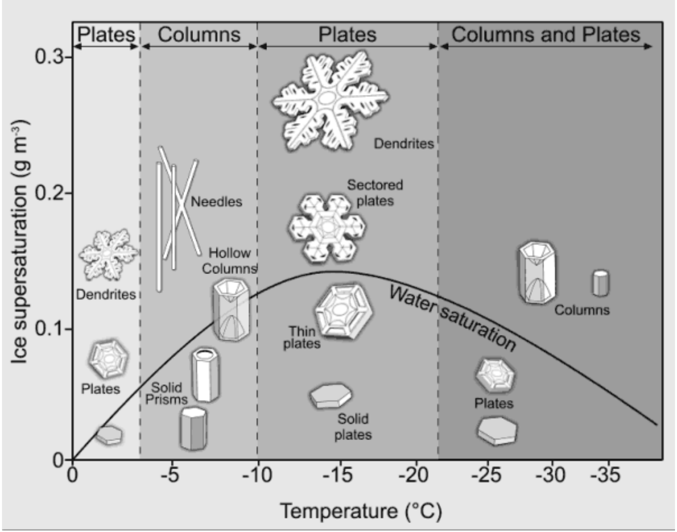


Figure: Ice habit versus temperature, saturation `http://www.cas.manchester.ac.uk/images/photos/themes/600x400/cp_scatt_Fig2_hr.png`

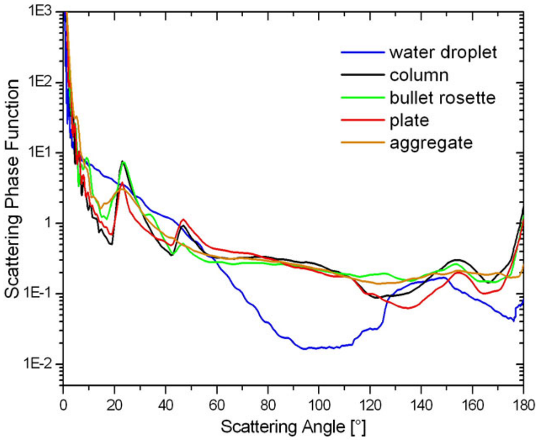
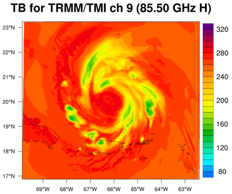
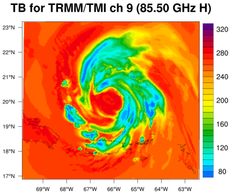


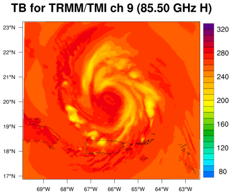
Figure: Scattering phase functions (solar) `http://www.uni-leipzig.de/~strahlen/web/research/Arctic/images/phasefunctionc.jpg`



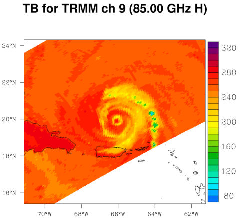
(a) Ice as spheres



(b) Hex columns



(c) Dendrites



(d) Observation

Figure: Ice habits BT response calculated from Liu 2008 ScatDB and RTTOV 11.1

Drop-size distribution

- Drop-size distribution (DSD) important
- Entire DSD spread plays a role
- Two main assumptions:
  - Exponential:

$$N(D) = n_0 \exp(-\Lambda D) \tag{1}$$

  - Gamma:

$$N(D) = n_1 D^{\nu} \exp(-\lambda D) \tag{2}$$

- Also use mass/diameter relationship $m = \alpha D^{\beta}$
- Gives three and four parameters, respectively

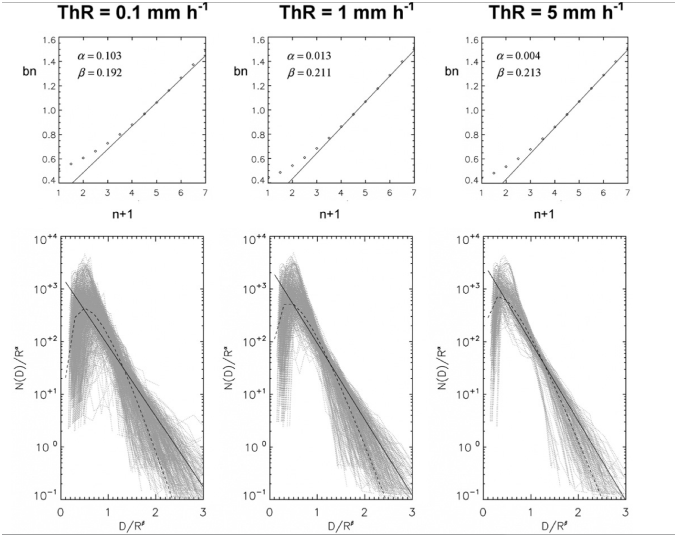
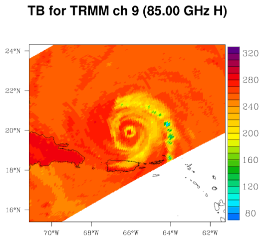
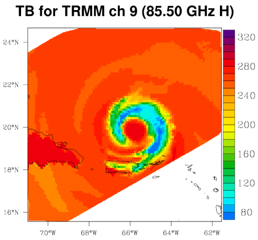


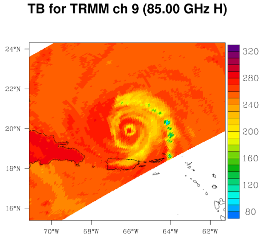
Figure: "Fit" of parameters to observations Benoit Chapon, Guy Delrieu, Marielle Gosset, Brice Boudevillain, Variability of rain drop size distribution and its effect on the Z-R relationship: A case study for intense Mediterranean rainfall, Atmospheric Research, Volume 87, Issue 1, January 2008, Pages 52-65, ISSN 0169-8095, http://dx.doi.org/10.1016/j.atmosres.2007.07.003. (http://www.sciencedirect.com/science/article/pii/S0169809507001226)



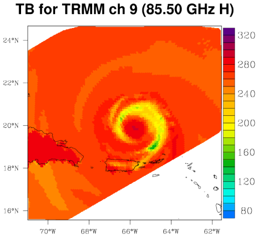
(a) Observation



(b) DSD 1



(c) Observation



(d) DSD 2

What's wrong with constant DSD parameters?

- Usually assume a constant n_c (#/kg)
- Mixing ratio varies by 3+ orders of magnitude
- DSD needs to represent all situations
- e.g. rain D must range from $0.4 - 4$ mm
- A lot to ask from a DSD parameterization!

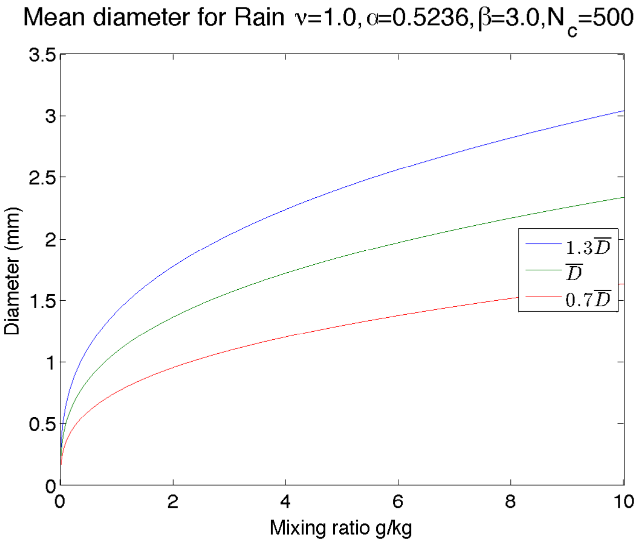


Figure: Rain mean diameter perturbation ranges

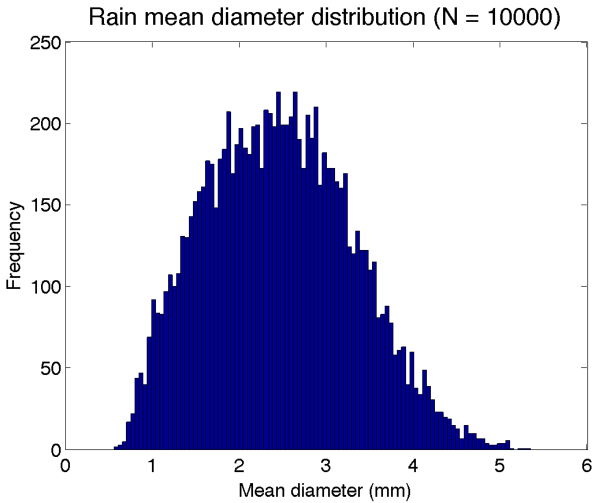


Figure: Histogram of rain mean diameter

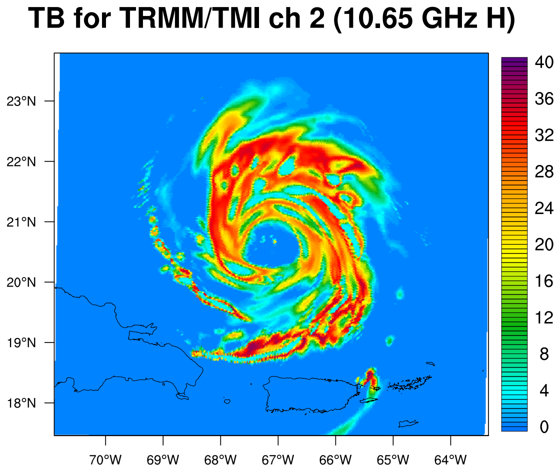


Figure: Standard deviation of brightness temperature with perturbation in gamma parameters. On the order of 30 K stddev with reasonable uncertainty!

What to do about this uncertainty?

- Quite different from traditional obs
- Obs over a distribution and spatial range
- Sweep uncertainty into obs. err. covariance??
- “Bias” is a generous term
- Need to extract the crucial information
- Requires statistical techniques
- Train from range of model, RTM realizations

Methodology

- Take HWRF model columns as X
- Use ψ, χ, P, T, RH, W, QCloud, QRain, QIce, QSnow, QGraup, QHail at 12 levels = 504 variables
- Take simulated TRMM brightness temperatures as $Y = H(X)$
- Used 2010-08-29 12:00 – 2010-09-03 18:00
- Use 12 million land-cleared columns/obs as samples of i.i.d. variables
- Assume clear/cloudy probabilities given by model
- Inflates both B, R vs. clear/cloudy

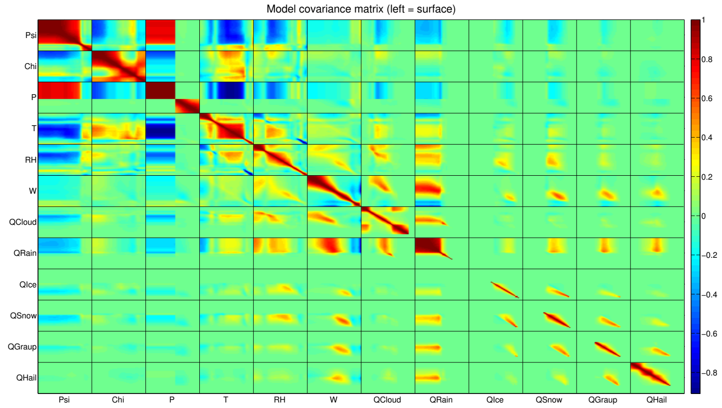


Figure: Model covariance: left is surface

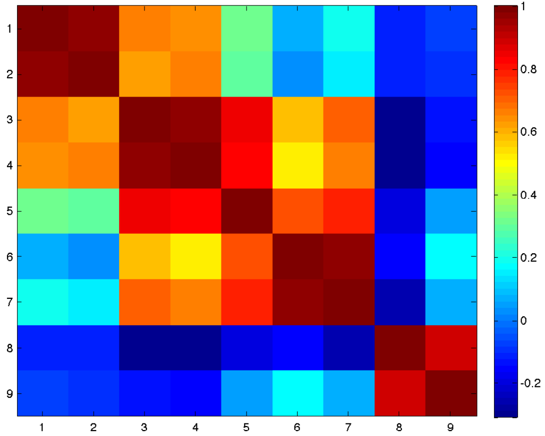


Figure: TRMM observation covariance: V/H order

Principal component analysis

- Collect samples of model columns, observations
- HWRF model columns as X, CRTM TRMM as Y
- Assume clear/cloudy probabilities given by model
- Calculate X and Y covariance matrices C_{xx}, C_{yy}
- Compute the singular value decompositions

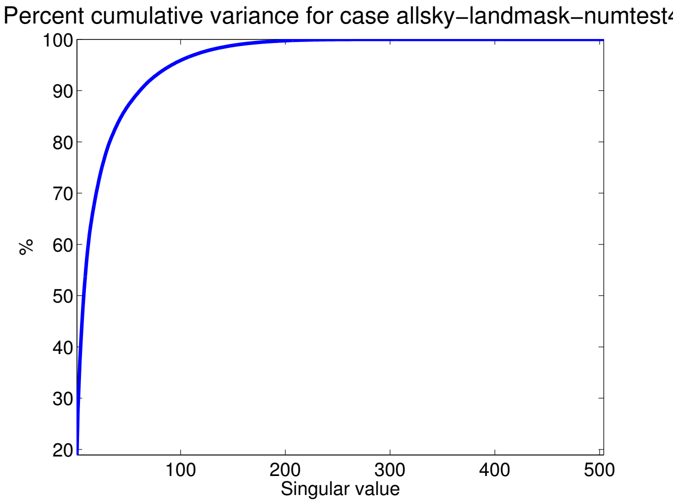


Figure: % cumulative variance of first *n* model PCs. 100 PCs contribute 95%, while 200 contribute 99.9%

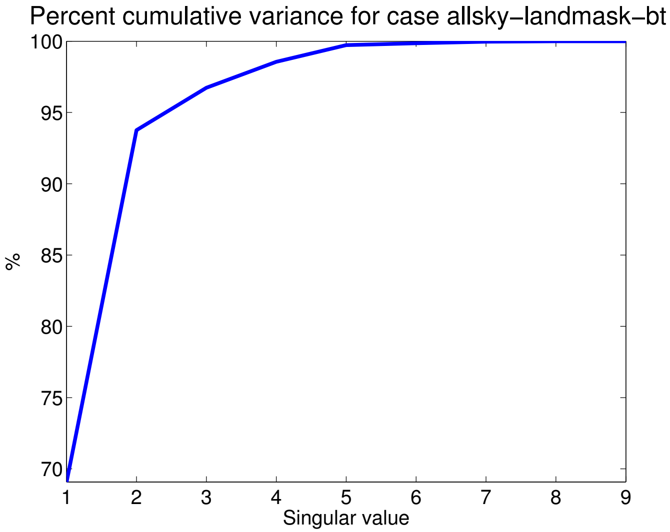


Figure: % cumulative variance of first *n* TRMM/TMI PCs. 2 PCs contribute 94%, while 5 contribute 99%

Methodology

- Standardize X and Y by μ, σ by level/channel
- Compute covariances C_{xx} and C_{yy}
- Provides useful, detailed **correlations**
- R is clearly **non-diagonal**; V/H assumption??
- Compute principal components (PC) w/ SVD
- Neglecting small PCs **regularizes** the problem

Extracting important relationships

- **Problem**: RTM overly-dependent on **uncertain parameters**
- Neglecting small PCs in B, R regularizes
- ...but PCs are unrelated
- **Idea**: find **best questions** to ask model, obs
- As in PCA, can we neglect uncertain relationships?
- What would a “best” relationship look like?

Motivation

- Want a linear relationship with tight correlation
- i.e., find a vector a for the model, b for obs s.t. $a^T X$, $b^T Y$ have the best correlation (scatter)
- Math: find a, b that maximize R^2

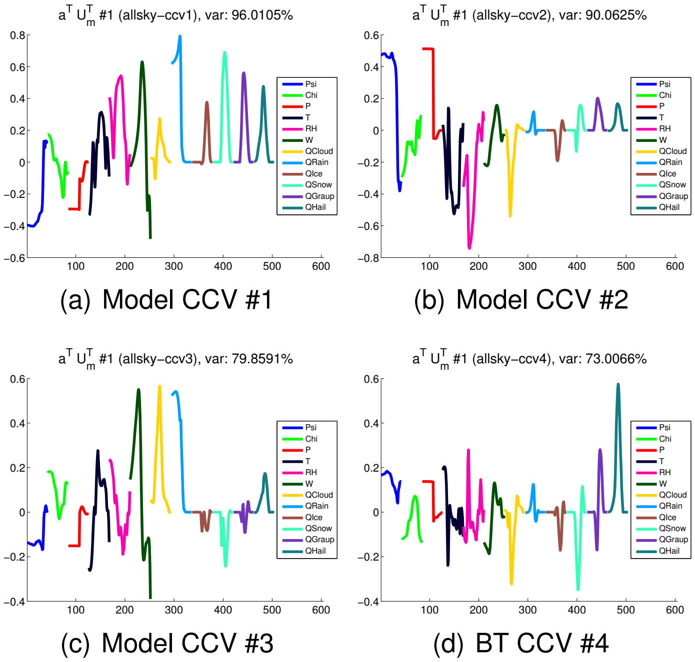
$$J(a,b) = \frac{\mathrm{cov}(a^{\mathrm{T}} X, b^{\mathrm{T}} Y)^2}{\mathrm{var}(a^{\mathrm{T}} X)\,\mathrm{var}(b^{\mathrm{T}} Y)} \tag{3}$$

- The solution is SVD of C_{xy}, where

$$ASB^{\mathrm{T}} = C_{xy} \tag{4}$$

- And the cross-covariance C_{xy} is given by

$$\left(C_{xy}\right)_{i,j} = E\left[\left(X_i - \mu_i\right)\left(Y_j - \mu_j\right)^{\mathrm{T}}\right] \tag{5}$$

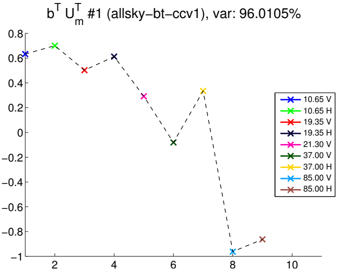


(a) Model CCV #1

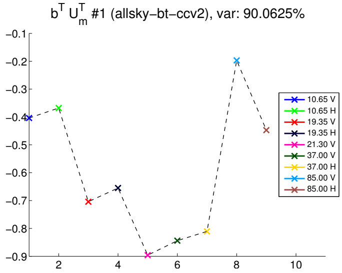
(b) Model CCV #2

(c) Model CCV #3

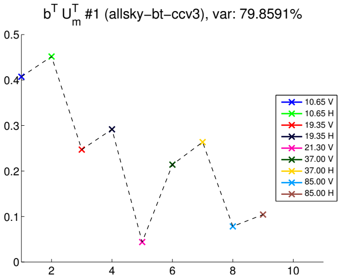
(d) BT CCV #4



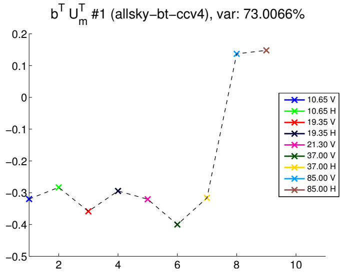
(a) BT CCV #1



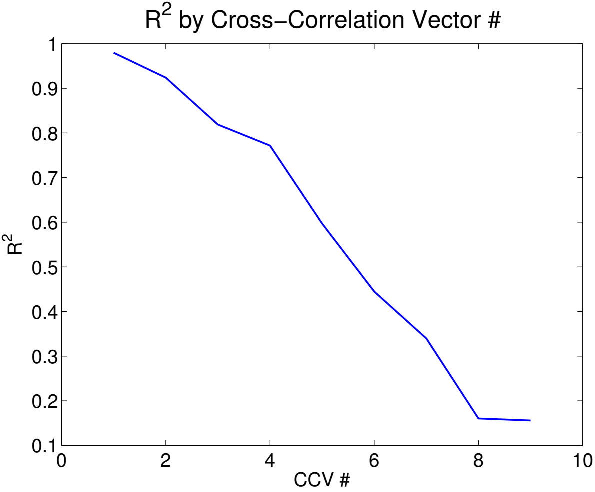
(b) BT CCV #2



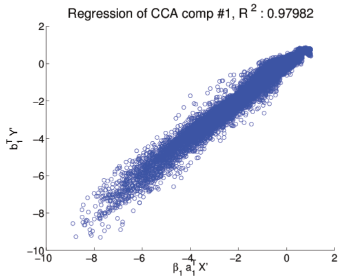
(c) BT CCV #3



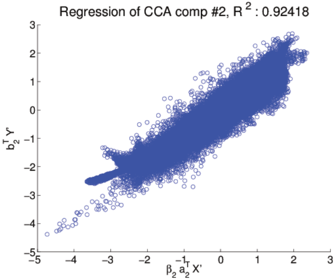
(d) BT CCV #4



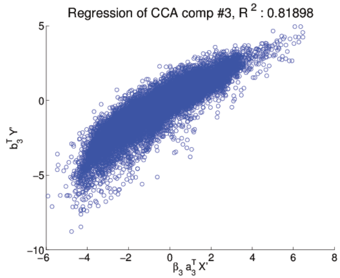
R^2 by Cross−Correlation Vector #
R^2
CCV #
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
4
6
8
10



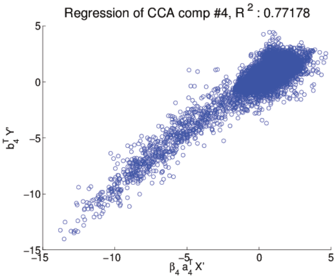
(a) CCV #1



(b) CCV #2



(c) CCV #3



(d) CCV #4

Methodology

- Now do linear regression between $a_i^{\mathrm{T}} X$ and $b_i^{\mathrm{T}} Y$

$$b_i^{\mathrm{T}} Y = \alpha_i + \beta_i a_i^{\mathrm{T}} X \tag{6}$$

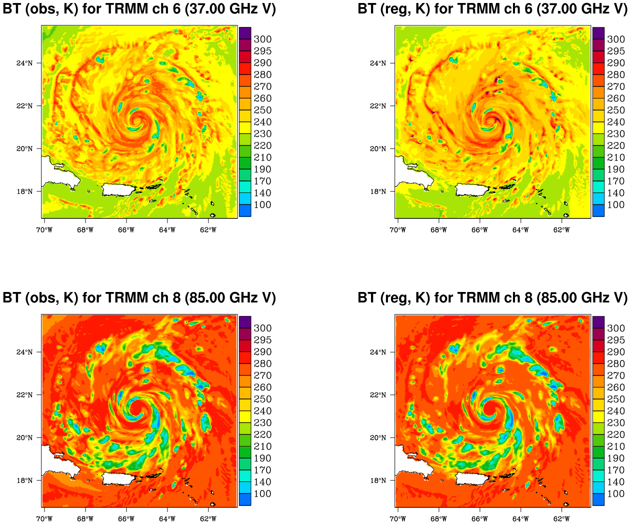
- Analytically, $\alpha_i = 0$ and

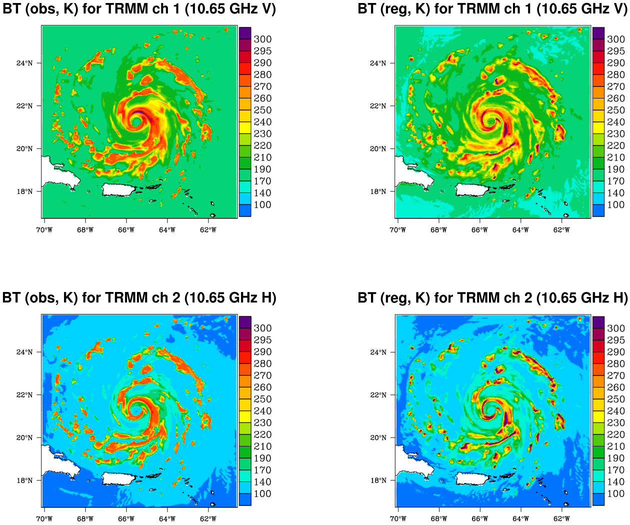
$$\beta_i = a_i^{\mathrm{T}} C_{xy} b_i = S_i \tag{7}$$

- We can use this to find H_r to full obs space:

$$Y_p = H_r(X) = C_{xy_r}^{\mathrm{T}} B_r^{-1} X \tag{8}$$

- Unstandardize Y_p for “best” linear operator (LS)


BT (obs, K) for TRMM ch 6 (37.00 GHz V)
BT (reg, K) for TRMM ch 6 (37.00 GHz V)
BT (obs, K) for TRMM ch 8 (85.00 GHz V)
BT (reg, K) for TRMM ch 8 (85.00 GHz V)
24°N
22°N
20°N
18°N
70°W
68°W
66°W
64°W
62°W
300
295
290
280
270
260
250
240
230
220
210
190
170
140
100

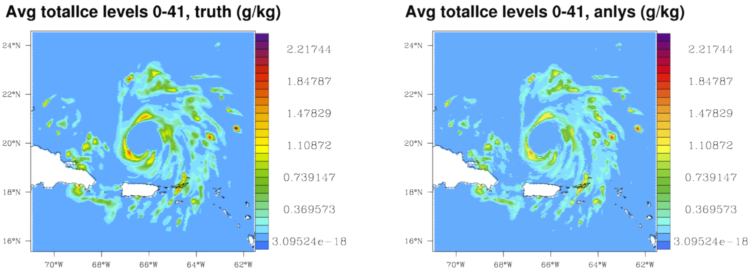

BT (obs, K) for TRMM ch 1 (10.65 GHz V)
BT (reg, K) for TRMM ch 1 (10.65 GHz V)
BT (obs, K) for TRMM ch 2 (10.65 GHz H)
BT (reg, K) for TRMM ch 2 (10.65 GHz H)
24°N
22°N
20°N
18°N
70°W
68°W
66°W
64°W
62°W
300
295
290
280
270
260
250
240
230
220
210
190
170
140
100

Benefits

- We now have a linear (i.e. **fast**) operator based on regression (but can do much better w/ NL fit)
- Can go to full observation space (K) or remain in CCV space (standardized)
- CCV space gives **uncorrelated** observations
- The first-few CCVs have **physical** meaning
- Using only these extracts the **essence** of the data
- The data **speaks for itself** about the relative importance and quality of relationships

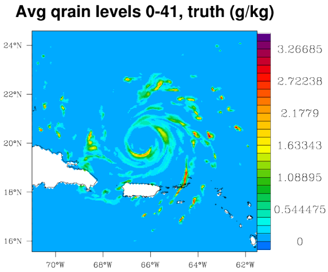
1D-Var OSSE test

- Have CCV operator $\mathcal{H}$ (with improvements, see paper)
- Use in an OSSE w/ CRTM hi-res observations
- Segmented regions with individual statistics
- Take 3 CCVs “obs” for each region
- Run 1D-Var with $\mathcal{H}$, B, R, and mean background

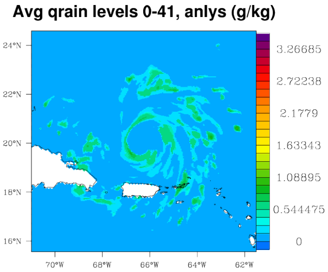


(a) Truth

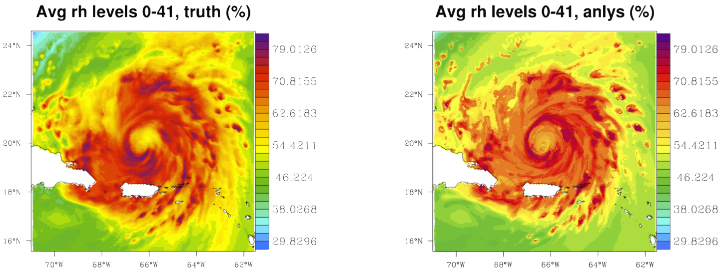
(b) Analysis



(a) Truth

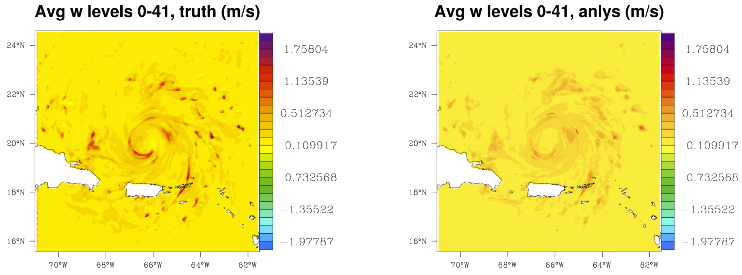


(b) Analysis



(a) Truth

(b) Analysis



(a) Truth

(b) Analysis

Conclusions

- MW radiances are highly sensitive to uncertain parameters
  - Ice habit: 10+ K
  - DSD parameters: 30+ K
  - Surface emissivity (esp. over land): ?
  - Sub-grid scale variability: ?
- Need to **understand** uncertainty, **extract** certainty
- Only then can MW be fully utilized for DA

Conclusions

- Canonical correlation vectors (CCVs) are maximally linearly correlated vectors
- Have “physical” meaning, capture most important relationships
- Neglect uncertain, noisy relationships
- Fast linear reduced-order observation operator
- More complex obs operator, background, first guess, segmenting data...
- Assimilation results very encouraging at this scale
- Can bring in hydrometeors, humidity, along with vertical velocity, hopefully ensuring **consistency**

Next steps

- Have computed w/ antenna **convolution** in place
- Thus will assimilate at observation resolution
- Integrate with the Hurricane Ensemble Data Assimilation System (HEDAS), GSI...
- Extend method to other satellites, frequencies
- Additional ideas? `jsteward@jifresse.ucla.edu`

Questions / discussion