A Deterministic Inverse Method for SST Retrieval from VIIRS: Incorporating Aerosol in the Retrieval Vector

Research: Prabhat Koner

PI: Andy Harris

CICS, ESSIC, University of Maryland

NASA MODIS-VIIRS ST Meeting, June 6 – 10, 2016

Physical Retrieval - Recap

- **Reduces the problem to a local linearization**
  - Dependent on ancillary data (NWP) for an initial guess
  - More compute-intensive than regression – not an issue nowadays
    - Especially with fast RTM (*e.g.* CRTM)
- **Widely used for satellite sounding**
  - More channels, generally fewer (larger) footprints
- **Initially, started with a simple reduced state vector**
  - $\boldsymbol{x} = [\mathrm{SST}, \mathrm{TCWV}]^T$
  - *N.B.* Implicitly assumes NWP profile shape is more or less correct
- **Selection of an appropriate inverse method**
  - Ensure that satellite measurements are contributing to signal
  - Avoid excessive error propagation from measurement space to parameter space
    - If problem is ill-conditioned

History of Inverse Model

- **Forward model:** $\mathbf{Y} = \mathbf{K}\mathbf{X}$
- **Simple Inverse:** $\mathbf{X} = \mathbf{K}^{-1}\mathbf{Y}$ **(measurement error)**

- **Legendre (1805) Least Squares:**

$$\mathbf{X} = \mathbf{X}_{ig} + (\mathbf{K}^{T}\mathbf{K})^{-1}\mathbf{K}^{T}(\mathbf{Y}_{\delta} - \mathbf{Y}_{ig})$$

- **MTLS:** $\mathbf{X} = \mathbf{X}_{ig} + (\mathbf{K}^{T}\mathbf{K} + \lambda\mathbf{R})^{-1}\mathbf{K}^{T}(\mathbf{Y}_{\delta} - \mathbf{Y}_{ig})$

- **OEM:** $\mathbf{X} = \mathbf{X}_{a} + (\mathbf{K}^{T}\mathbf{S}_{e}^{-1}\mathbf{K} + \mathbf{S}_{a}^{-1})^{-1}\mathbf{K}^{T}\mathbf{S}_{e}^{-1}(\mathbf{Y}_{\delta} - \mathbf{Y}_{a})$

Uncertainty Estimation

Physical retrieval

Normal LSQ Eqn: $\Delta x = (K^TK)^{-1}K^T\Delta y \quad [= G\Delta y]$

MTLS modifies gain: $G' = (K^TK + \lambda I)^{-1}K^T$

Regularization strength: $\lambda = (2\log(\kappa)/||\Delta y||)\sigma^2_{end}$

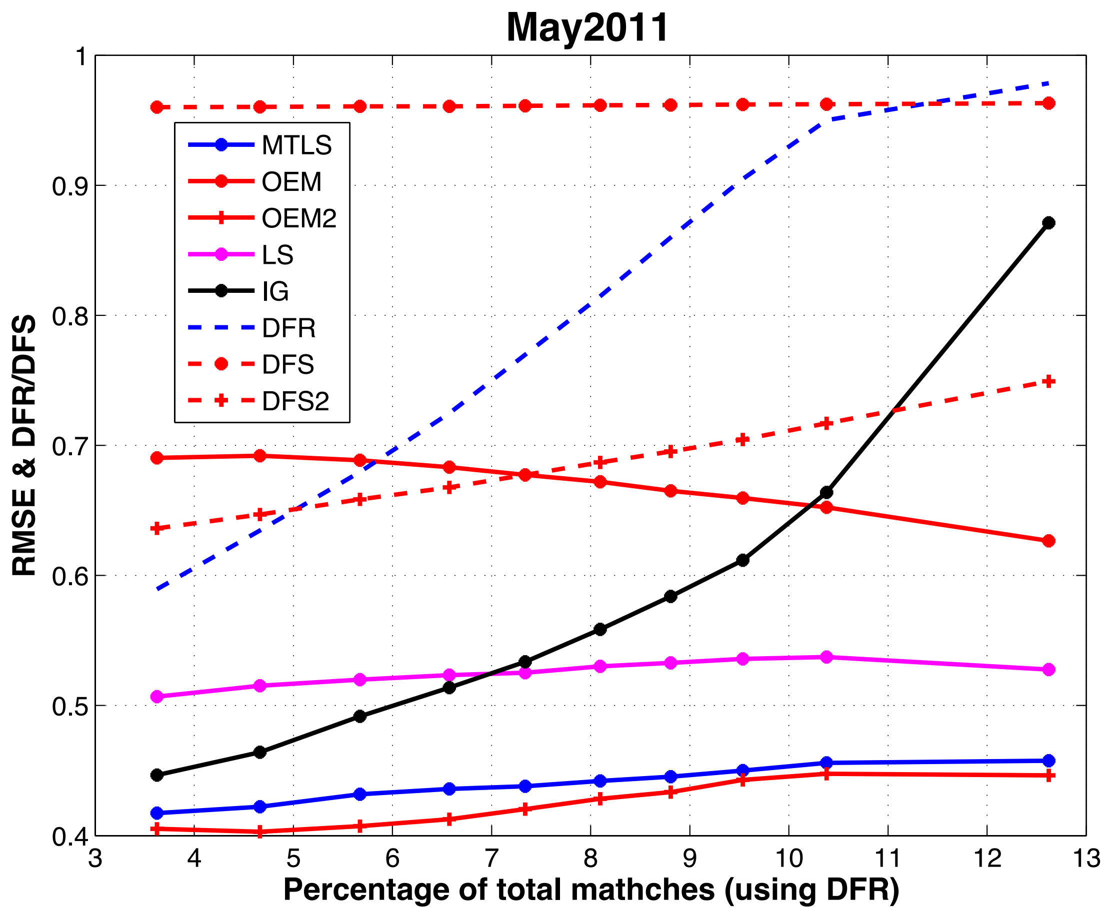
(σ^2_{end} = lowest singular value of $[K\ \Delta y]$)

Total Error

$$||e|| = ||(MRM - I)\Delta x|| + ||G'||\langle||(\Delta y - K\Delta x)||\rangle$$

***N.B.* Includes TCWV as well as SST**

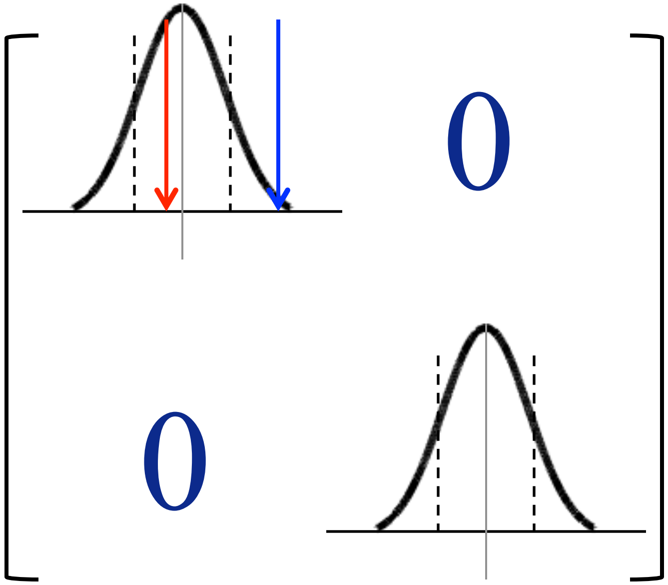
DFS/DFR and Retrieval error



- ❑ Retrieval error of OEM higher than LS
- ❑ More than 75% OEM retrievals are degraded w.r.t. *a priori* error
- ❑ DFR of MTLS is high when *a priori* error is high

“Optimized” OE

- **$[S_e]$, S_a =**

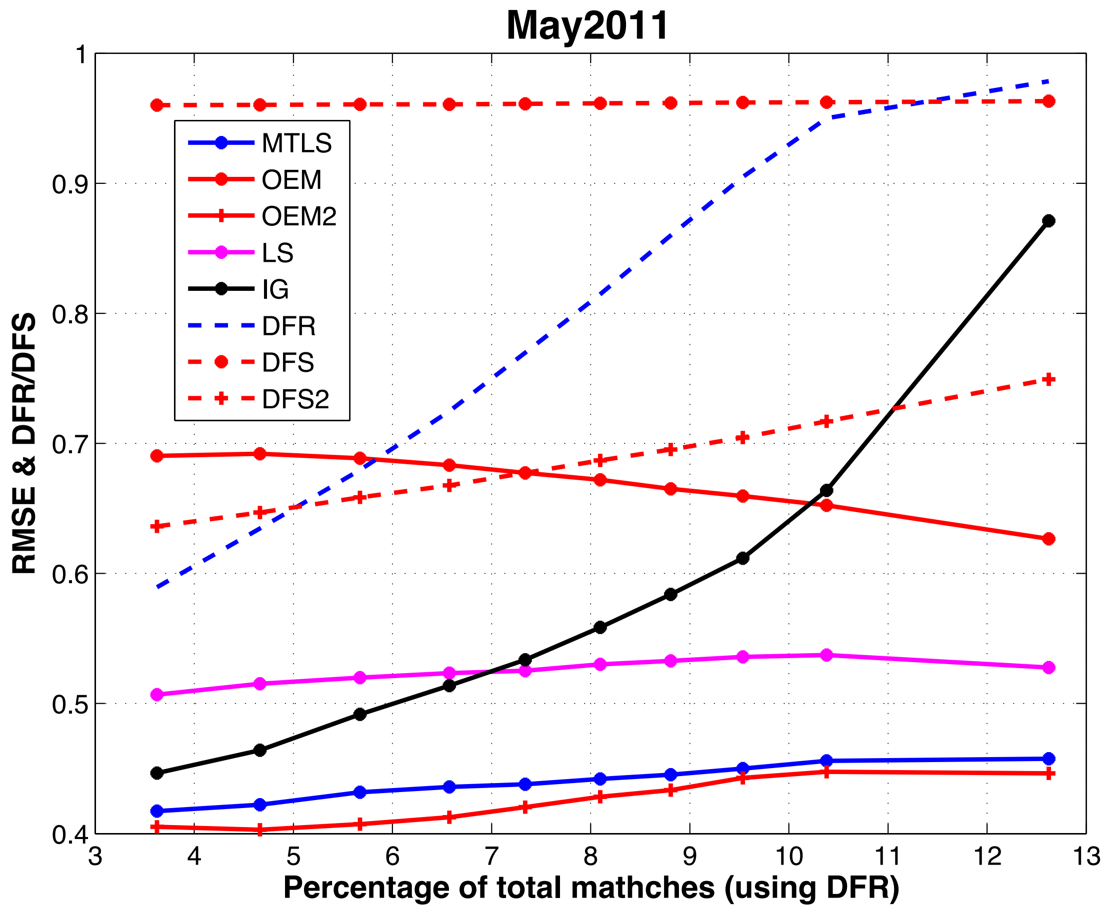


σ^2 is an overestimate…

…or an underestimate

- **Perform experiment – insert “true” SST error into S_a^{-1}**
  - Can only be done when truth is known, *e.g.* with matchup data

DFS/DFR and Retrieval error

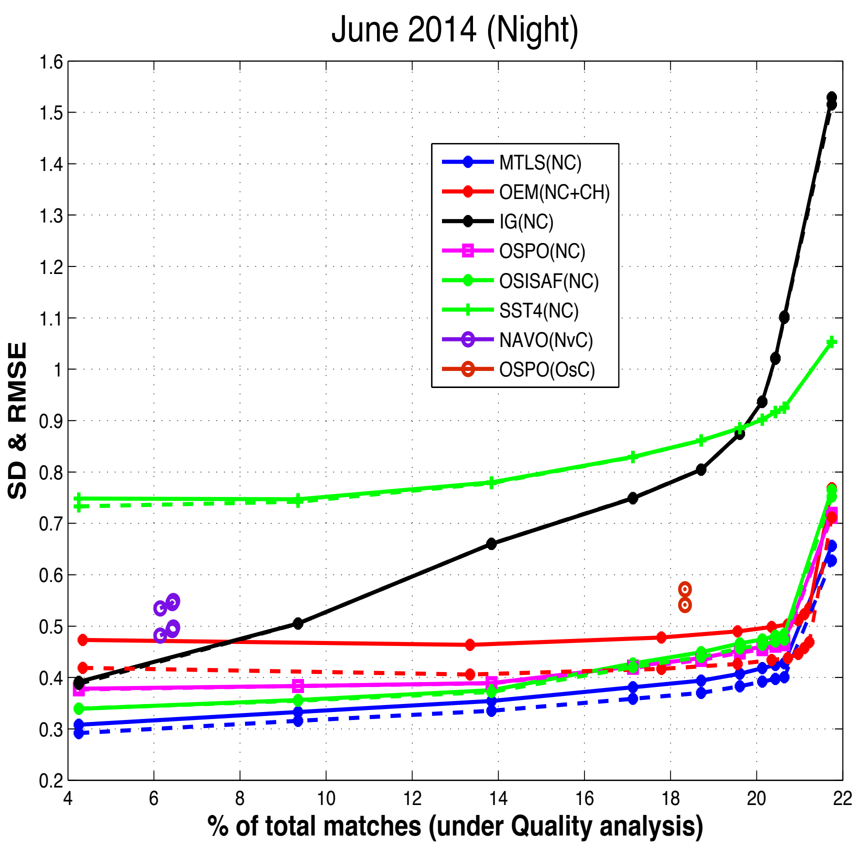


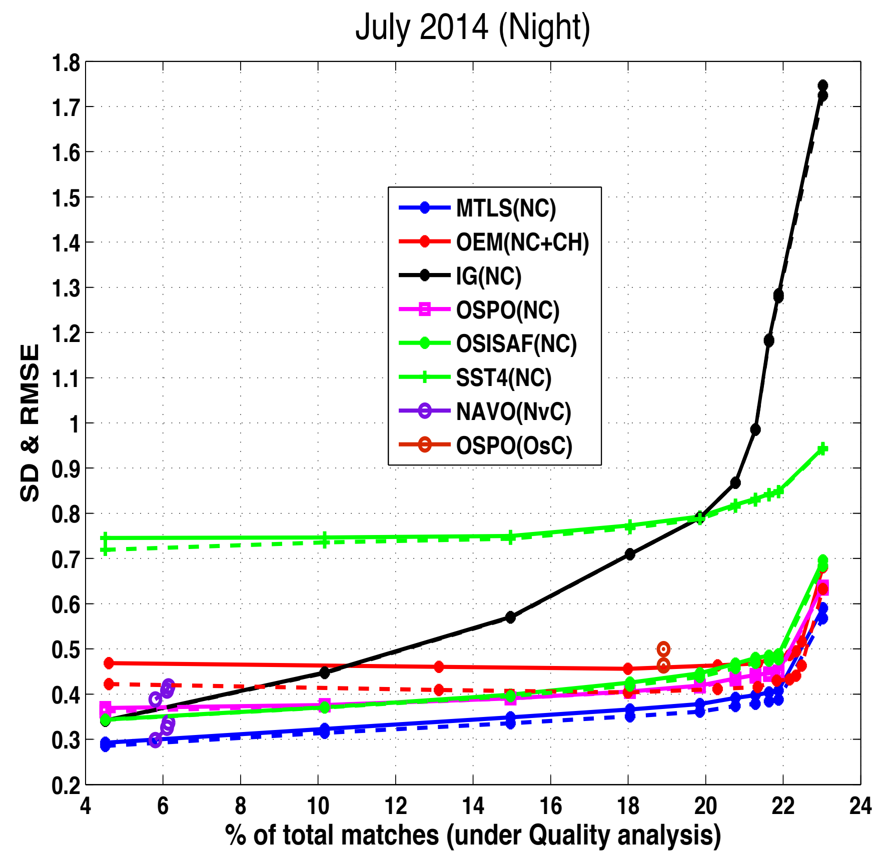
- Retrieval error of OEM higher than LS
- More than 75% OEM retrievals are degraded w.r.t. *a priori* error
- DFR of MTLS is high when *a priori* error is high
- The retrieval error of OEM is good when *a priori* SST is perfectly known, but DFS of OEM is much lower than for MTLS

Improved cloud detection

- **Use a combination of spectral differences and RT**
  - Envelope of physically reasonable clear-sky conditions
- **Spatial coherence (3×3)**
- **Also check consistency of single-channel retrievals**
- **Flag excessive TCWV adjustment & large MTLS error**
- **Increased coverage w.r.t. GHRSST QL3+, but with reduced cloud leakage**
  - **Prabhat's talk in yesterday's Oceans Breakout**
  - **~50% increase in coverage & ~50% reduction in error**

VIIRS Initial Results

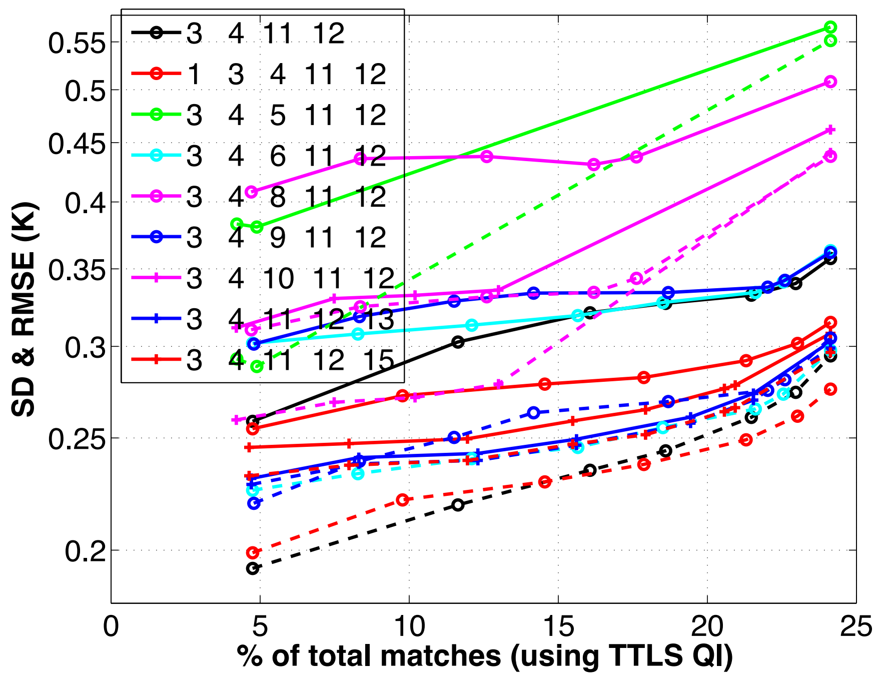


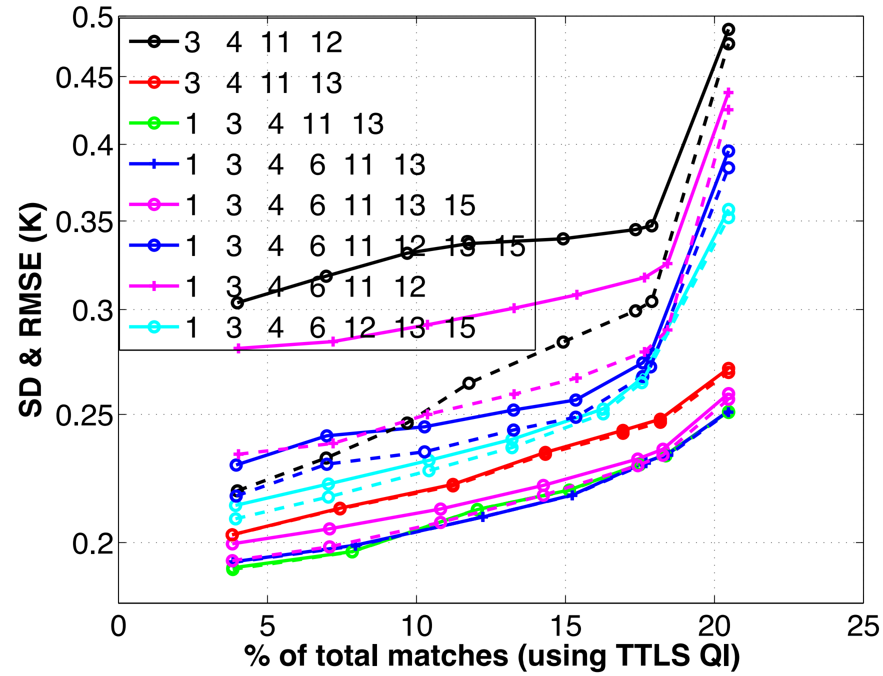


- **Data are ordered according to MTLS error**
  - Reliable guide for regression as well as MTLS
  - Trend of initial guess error is expected

MODIS experiments

- **Channel selection**
  - Test various combinations and look at accuracy of retrieval





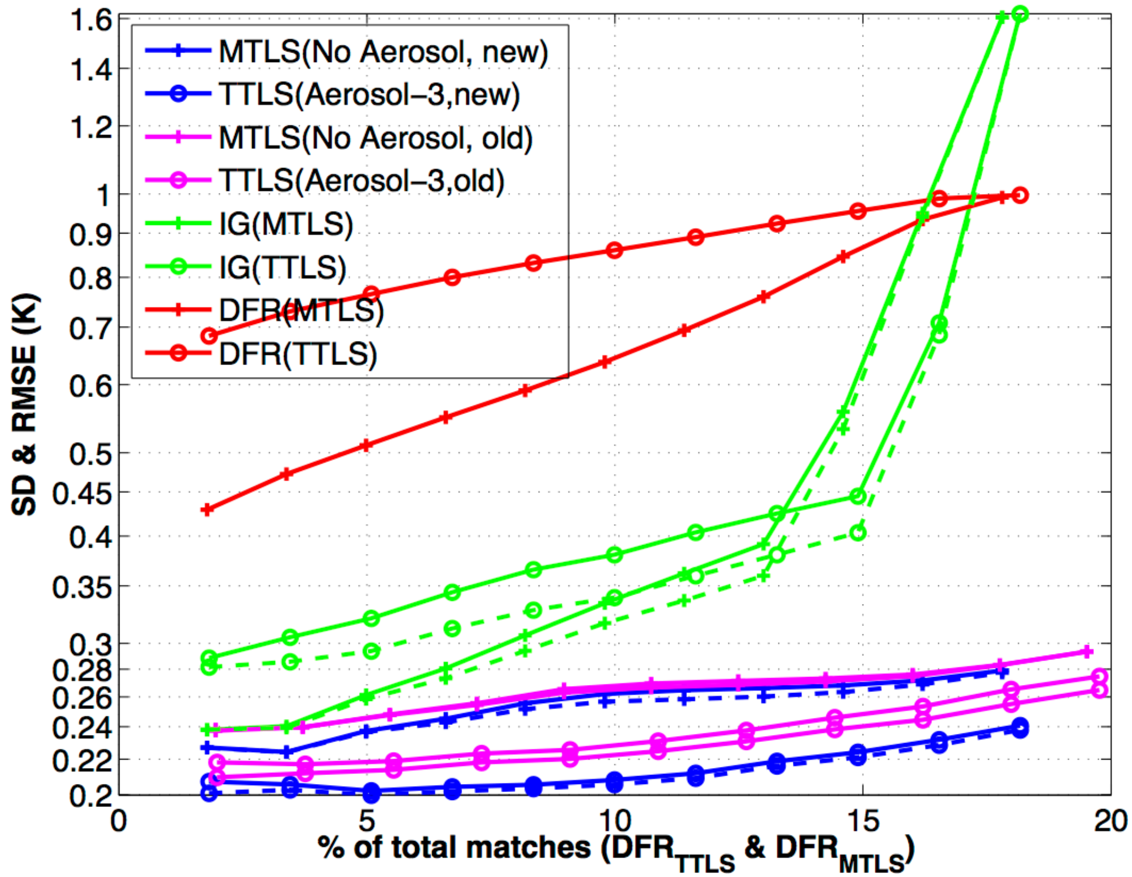
- **RTM may be inadequate for some channels ➔ bias**
- **Channels 1, 3, & 13 are particularly useful**

Addition of aerosol

- **Put aerosol information in the CRTM**
  - NGAC profiles, multiple species (dust, salt, sulfate, soot)
  - Improve match of RTM to observation
  - Does this improve retrieval?
- **Put aerosol in the retrieval vector**
  - Allow Total Column Aerosol to vary
  - $\boldsymbol{x} = [\text{SST}, \text{WV}, \text{TCA}]^T$
  - Jacobian now includes $\partial T/\partial \text{TCA}$ for each channel
  - Does this improve retrieval?
- **MTLS developed for 2-parameter retrieval**
  - Try different regularization operator since problem is now more ill-conditioned: **Truncated Total Least Squares (TTLS)**

$$|\Delta \boldsymbol{y}| \leq 1: \ \lambda = (\sigma_{end-1})^2 \qquad |\Delta \boldsymbol{y}| > 1: \ \lambda = (\sigma_{end-1}/\log(|\Delta \boldsymbol{y}|))^2$$

Inclusion of aerosol



- **Accuracy with TTLS & joint [SST, WV, TCA] ~0.2 K**
- **Algorithm sensitivity is also improved *cf.* MTLS**

Summary

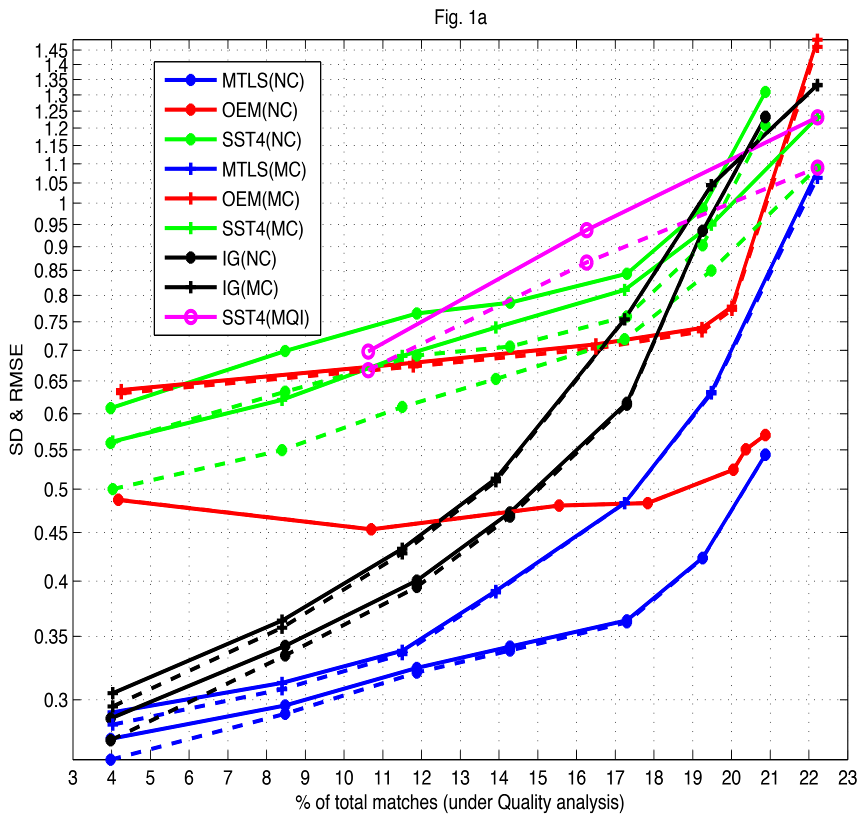
- **Addition of aerosol has significant benefit**
  - Most of all when included in retrieval vector as well as CRTM
- **Better partitioning of brightness temperature residuals**
  - No longer forcing delta-BTs caused by aerosol into the SST and/or WV retrieval space
  - Also improves algorithm sensitivity to SST (better overall fit to model)
- **TTLS better choice for 3-parameter retrieval**
  - Initial “tuning” with MODIS works well
  - Adaptation to VIIRS channels underway
- **Validation results are approaching buoy accuracy limit**
  - Best ~50% of retrievals at 0.2 K
  - Implies actual retrieval accuracy is better than this
- **Need to consider what might be needed @SIPS**
  - Full aerosol profiles as well as NWP

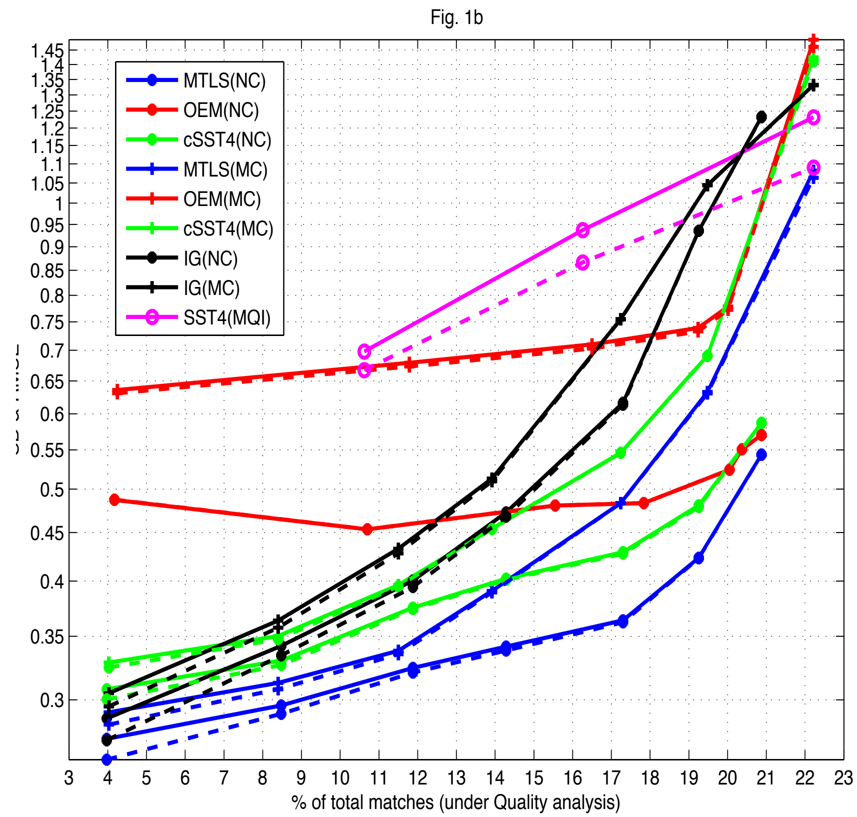
Backup slides

Improvements

- **It seems "obvious" that a sensitivity of 1 is desirable**
  - *E.g.* if there is diurnal warming of 5 K, it will be observed in the data, and strong upwellings will be accurately observed, *etc*.
- **However, there is a penalty to be paid**
  - Ill-conditioned problem ➔ noise propagates from measurement space to parameter space
  - Compromise is usually struck (*e.g.* minimum least squares result for training data in a regression algorithm)
- **Regression algorithms may have sensitivity <1 for large regions**
  - *E.g.* daytime algorithms in the tropics (diurnal warming!)
  - Causes bias if local atmospheric conditions are different from the ensemble mean for the training data

MODIS Initial Results





- **Note improvement from discarding MTLS error "last bin"**
  - Irrespective, MTLS is quite tolerant of cloud scheme
- **Recalculated SST4 coefficients produce quite good results**