A Deterministic Inverse Method for SST Retrieval from VIIRS: Early Results

Andy Harris, Prabhat Koner
CICS, ESSIC, University of Maryland
Motivation

• **Previous generation SST algorithms are regression-based**
  – *E.g.* MCSST, NLSST (Pathfinder)
  – Usually employed direct regression of radiances against *in situ* SSTs
  – Ameliorates issues with instrument calibration/characterization

• **Some success for RT-based regression**
  – Primary example (A)ATSR series
  – Well-calibrated and characterized radiometer
  – Dual-view permitted robust retrieval, but fairly narrow swath

• **Regression-based algorithms could result in regional/seasonal biases**
  – Attempt to characterize global retrieval conditions with only a few coefficients
  – Causes bias if local atmospheric conditions are different from the ensemble mean for the training data
Simulated Pathfinder Retrieval Errors

ERA-40 data

- Atmospheric profiles
- SSTs

ERA-40 “matchup” subset

CRTM

Simulated matchup BTs

SST Retrieval coefficients

Simulated Pathfinder SSTs (+ ERA-40 SSTs)

Pathfinder matchups
- Lat, lon, time, view angle

No Aerosols & ERA-40 data filtered for cloud fraction

N.B. Bias is Pathfinder SST – ERA-40 SST
What happens when we include volcanic aerosol?
Simulated Pathfinder Retrieval Errors

ERA-40 data
- Atmospheric profiles
- SSTs

ERA-40 “matchup” subset

CRTM

Simulated matchup BTs

SST Retrieval coefficients

Pathfinder matchups
- Lat, lon, time, view angle

Simulated global BTs

Insert volcanic aerosol optical depths

Simulated Pathfinder SSTs (+ ERA-40 SSTs)

Insert volcanic aerosol optical depths

N.B. Bias is Pathfinder SST – ERA-40 SST
Include Pinatubo in RTM radiances

- Negative bias is reduced, but positive biases are propagated N & S
- Split-window based algorithm has no skill in compensating for aerosol

NASA MODIS-VIIRS ST Meeting, May 18 – 22, 2015
Pathfinder V5 – Daily OI $\frac{1}{4}$°

- Common features w.r.t. biases induced by Pinatubo aerosol
- Actual seasonal variability is greater than predicted by modeling

NASA MODIS-VIIRS ST Meeting, May 18 – 22, 2015
Physical Retrieval

• **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

• **Start with a simple reduced state vector**
  – \( x = [\text{SST}, \text{TCWV}]^T \)
  – *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:** \( Y = KX \)
- **Simple Inverse:** \( X = K^{-1}Y \) (measurement error)

- **Legendre (1805) Least Squares:**
  \[
  X = X_{ig} + (K^TK)^{-1}K^T(Y_\delta - Y_{ig})
  \]

- **MTLS:**
  \[
  X = X_{ig} + (K^TK + \lambda R)^{-1}K^T(Y_\delta - Y_{ig})
  \]

- **OEM:**
  \[
  X = X_a + (K^T S_e^{-1}K + S_a^{-1})^{-1}K^T S_e^{-1} (Y_\delta - Y_a)
  \]
Physical retrieval

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

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

Regularization strength:
\[ \lambda = \frac{2 \log(\kappa)}{||\Delta y||} \sigma^2_{\text{end}} \]
(\(\sigma^2_{\text{end}} = \text{lowest singular value of } [K \Delta y]\))

Total Error
\[ ||e|| = ||(MRM - I)\Delta x|| + ||G'||\langle||\Delta y - K\Delta x||\rangle \]

N.B. Includes TCWV as well as SST
- 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

- \([S_e], S_a = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\) → \(\sigma^2\) is an overestimate…
  …or an underestimate

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

- Almost as many as GHRSSST QL3+, but with greatly reduced leakage
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 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
Things to consider

• 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
Things to consider cont’d

- Physical retrieval methods locally linearizes the retrieval
  - Ameliorate regional bias issues
- Physical retrievals still ill-conditioned
  - Least-Squares generally considered to have unacceptable noise
- Optimal Estimation can have sensitivity ~1
  - Requires somewhat inflated SST error covariance
  - Leads to relatively poor noise performance
  - Using “true” SST error greatly improves retrieval accuracy
  - However, SST sensitivity is substantially reduced
- MTLS algorithm adjusts its sensitivity
  - Sensitivity <1 when initial guess is close to truth
  - Sensitivity ≥ 1 when initial guess is far from truth
  - Retrieval accuracy approaches “optimized” OEM
  - May still be an issue for fine structure
Summary

• **MTLS seems applicable to VIIRS**
  - Well-calibrated instrument, with reliable fast RTM available
  - Error calculation useful quality indicator

• **MODIS offers even more possibilities**
  - “Sounding” channels permit inclusion of basic profile shape information in the state vector
  - See Prabhat’s presentation at the Oceans Breakout

• **Cloud detection can be aided by RTM**
  - “Single-channel” retrieval consistency, MTLS error calculation

• **Options for improvement**
  - Close to validation limit for conventional *in situ*
  - Take advantage of differing length scales to reduce atmospheric noise
  - Perhaps combine with sounder for more local atmospheric information
  - Refine fast RTM, iteration
  - Tropospheric aerosols…
Backup slides
### Deterministic & Stochastic

#### Deterministic

**MTLS/RTLS/Tikhonov**: Single pixel

\[ dX = K^{-1} dY \]

measurement error

Lengendre (1805)

Least Squares:

\[ X_{ls} = X_{ig} + (K^T K)^{-1} K^T dY_\delta; \quad dY_\delta = Y_\delta - Y_{ig} \]

Last 30~40 years: \( \delta X \leq \kappa \delta E \); \( \kappa = \text{cond}(K) \)

\[ X_{rg} = X_{ig} + (K^T K + \lambda R)^{-1} K^T dY_\delta \]

\[ \lambda = (2 \log(\kappa) / \| dY_\delta \|^2) \sigma_{end}^2 \]

MTLS:

\[
[u \ \sigma \ v] = [K \ dY_\delta]; \quad R = I
\]

Total Error:

\[
\| X_{true} - X_{mls} \| + \| K_{ps}^{inv} (dY_\delta - Kx_{mls}) \|
\]

#### Stochastic/Probabilistic

**OEM**: A set of measurement

\[ X_{oem} = X_{ig} + (K^T S_e^{-1} K + S_a^{-1})^{-1} K^T dY_\delta \]

Low confidence for pixel retrieval

Chi-Square test:

\[ \chi_{resd} = K X_{oem} - dY_\delta \]

\[ \chi = \chi_{resd}^T (S_e (K^T S_a^{-1} K + S_e^{-1}) S_e)^{-1} \chi_{resd} \]

**Regression**: A set of measurement

Historical heritage in SST retrieval using Window channels.

Coefficient Vector/matrix: \( C \quad X_{reg} = C Y_\delta \)

Main concerns: Correlation & Causation
Recent update to Geo-SST

- Physical retrieval based on Modified Total Least Squares
- Improved bias and scatter *cf.* previous regression-based SST retrieval

GOES-15

**Daytime**

**Nighttime**
How sensitive is retrieved SST to true SST?

- If SST changes by 1 K, does retrieved SST change by 1 K?
- CRTM provides tangent-linear derivatives $\frac{\partial T_{11}}{\partial SST_{true}}$ and $\frac{\partial T_{12}}{\partial SST_{true}}$

Response of **NLSST algorithm** to a change in **true SST** is...

$$\frac{\partial NLSST}{\partial SST_{true}} = \left( a_1 + a_2 \times SST_{bg} + a_3 \times \{\sec(ZA) - 1\} \right) \times \frac{\partial T_{11}}{\partial SST_{true}}$$

$$- \left( a_2 \times SST_{bg} + a_3 \times \{\sec(ZA) - 1\} \right) \times \frac{\partial T_{12}}{\partial SST_{true}}$$

Sensitivity to true SST

Sensitivity often <1 and changes with season

NASA MODIS-VIIRS ST Meeting, May 18 – 22, 2015
Sensitivity to true SST
Air – Sea Temperature Difference

Jul

(1 - Sensitivity) x ASTD (K)

NASA MODIS-VIIRS ST Meeting, May 18 – 22, 2015
Seasonal Geographic Distribution of Bias
### Characteristics of different cloud detections

**Ospo Cloud (June 2014 Day)**

**New Cloud (Day June 2014)**

- The data coverage of new cloud (NC) 50% more than OSPO
- # cloud free pixels for high SZA is sparse – maybe OSPO & OSI-SAF regression form are not working for this regime

- There is no physical meaning from RT for a regression variable of SSTg multiplied with (T11-T12).