Accounting for Correlated Satellite Observation Error in Variational Data Assimilation

Bill Campbell, Liz Satterfield, and Ben Ruston
U. S. Naval Research Laboratory
Monterey, CA, USA

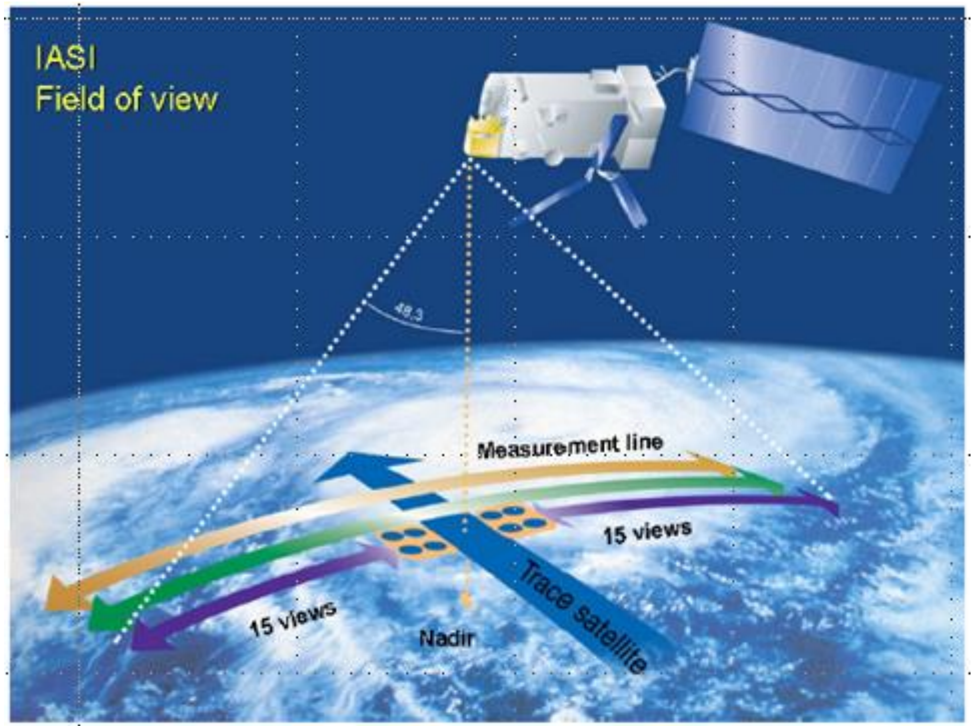
5th International Symposium on Data Assimilation
University of Reading, Reading, UK
21 July 2016

Outline

- High-impact Satellite Observations
- 4D Variational Data Assimilation Systems
- Observation Error and its Covariance Estimation
- Satellite Interchannel Error Characteristics and Practical Implementation
- Experiments and Forecast Results
- CrIS and Posterior Channel Selection
- Conclusions and Future Work

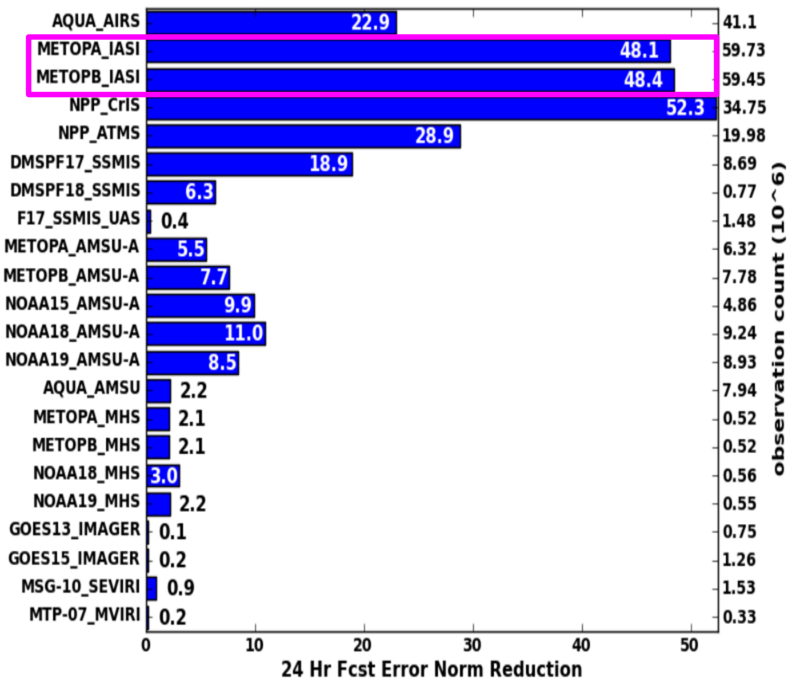
Infrared Atmospheric Sounder/Interferometer

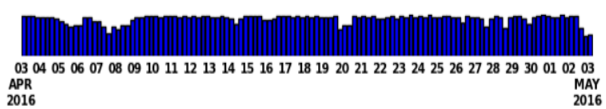
- 8641 channels
  - NRL receives **500** channels of Level 1c IASI data
  - We assimilate **69** of these channels, and monitor **251** more
- Metop-A & -B satellites (launched 19 Oct 2006 & 17 Sep 2012)
- 2200 km wide swath
  - **30** FOR, **4** FOV
  - +/- 48.33 scan angle
  - **12 km** at NADIR
  - **20 x 39 km** at edge of swath
- Three spectral bands:
  - long-wave band (LWIR):
    - spectral range: **8.26 - 15.50 µm**
  - mid-wave band (MWIR):
    - spectral range: **5.00 - 8.26 µm**
  - short-wave band (SWIR):
    - spectral range: **3.62 – 5.00 µm**



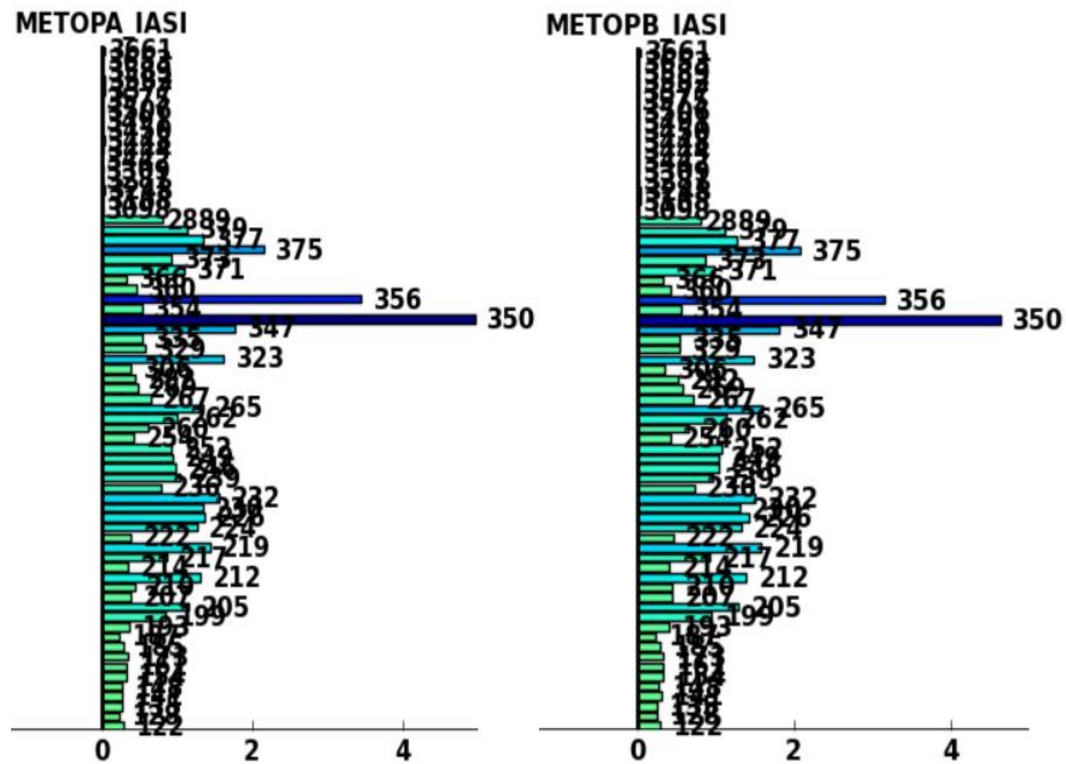
IASI Impact

NAVDAS-AR Observation Sensitivity





April 3 – May 3, 2016

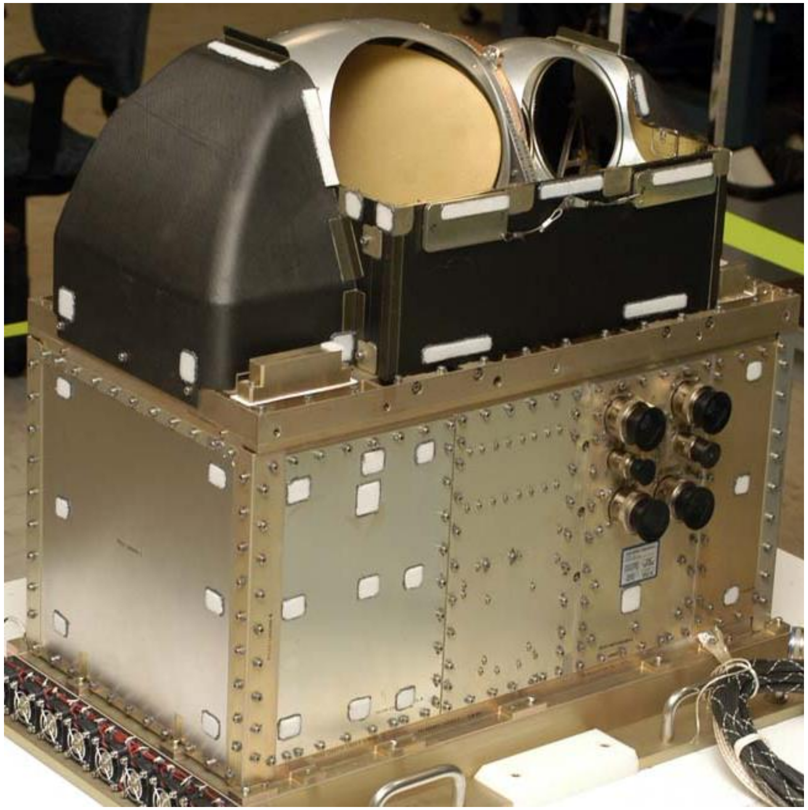


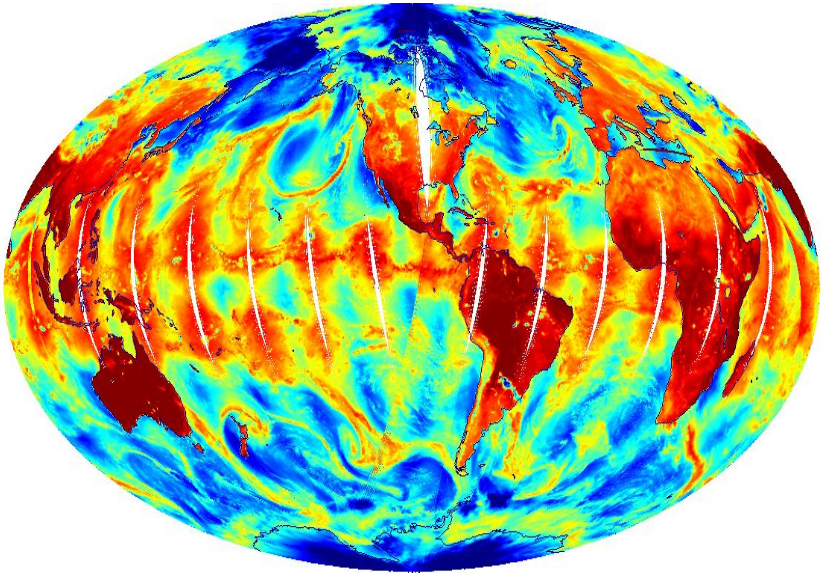
24 hr. Forecast Error Norm Reduction (J/kg)

Advanced Technology Microwave Sounder

High Precision Cross-Track MW Sounder

- **Suomi NPP (launched 28 Oct 2011)**
- **Heritage AMSU-A/MHS**
- **Additional O_2 channel at 51.76 GHz vs. AMSU-A**
- **Additional H_20 vapor channels at 166 and 5 about the 183 GHz line vs. MHS**



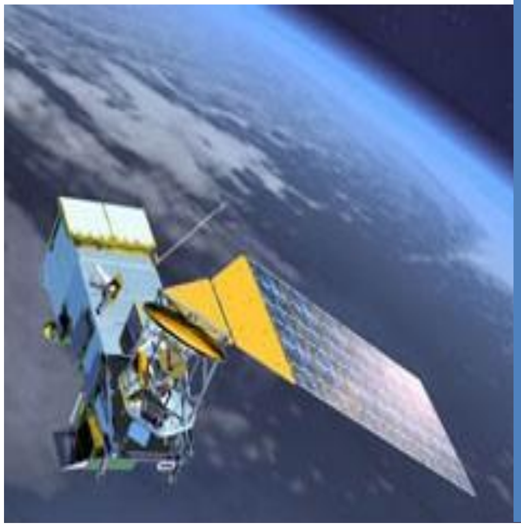


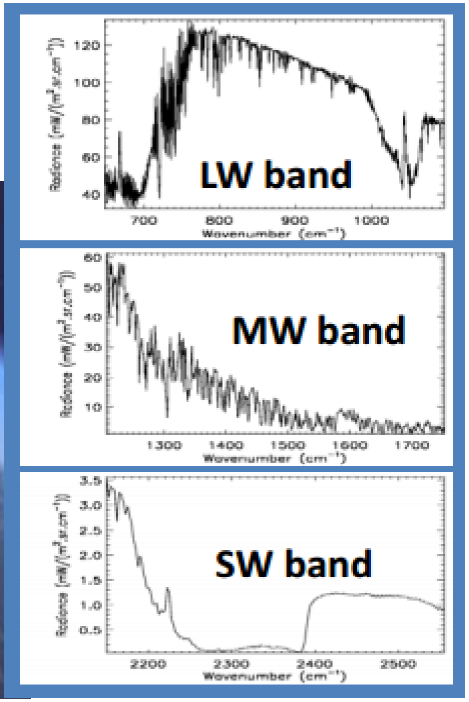
ATMS   166 GHz Window Channel

U.S. NAVAL RESEARCH LABORATORY

Cross-track Infrared Sounder

- 1305 channels
  - Data received is **1317** unapodized channels (1305 apodized)
  - Data sampled to **399** NESDIS subset for CRTM
- Suomi-NPP satellite (launched 28 Oct 2011)
- 2200 km wide swath
  - **30** FOR, **9** FOV
  - +/- 48.33 scan angle
  - **14km** at NADIR
  - **~24 x 46km** at edge of swath
- Update to full resolution
  - **2211** apodized channels
  - Only affected mid- and short-wave
    - long-wave band (LWIR):
      - spectral range: **9.14 - 15.38µm**
    - mid-wave band (MWIR):
      - spectral range: **5.71 - 8.26µm**
    - short-wave band (SWIR):
      - spectral range: **3.92 - 4.64 µm**

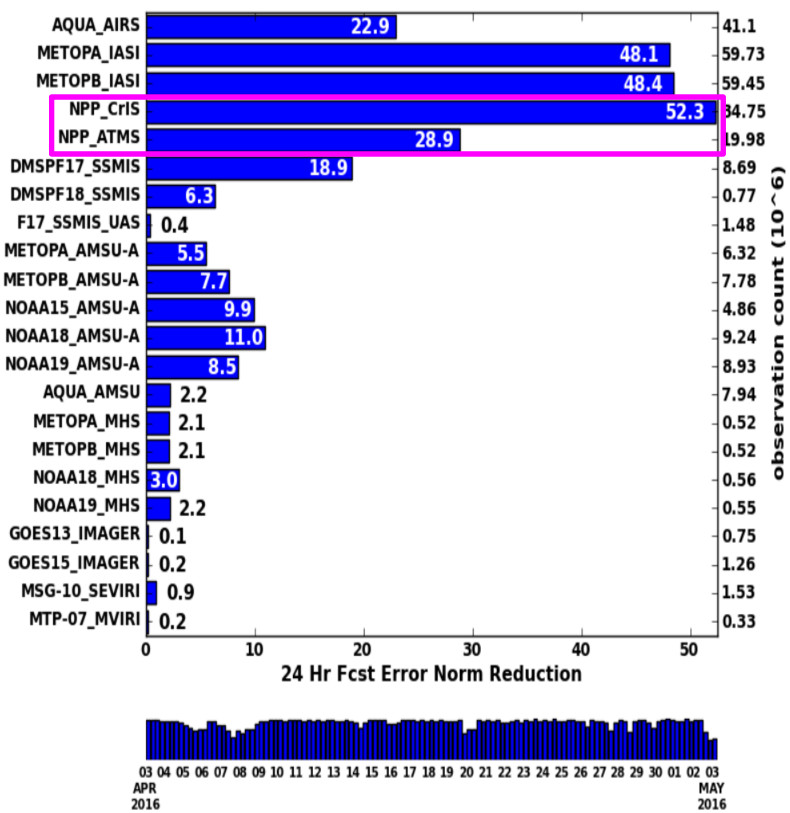


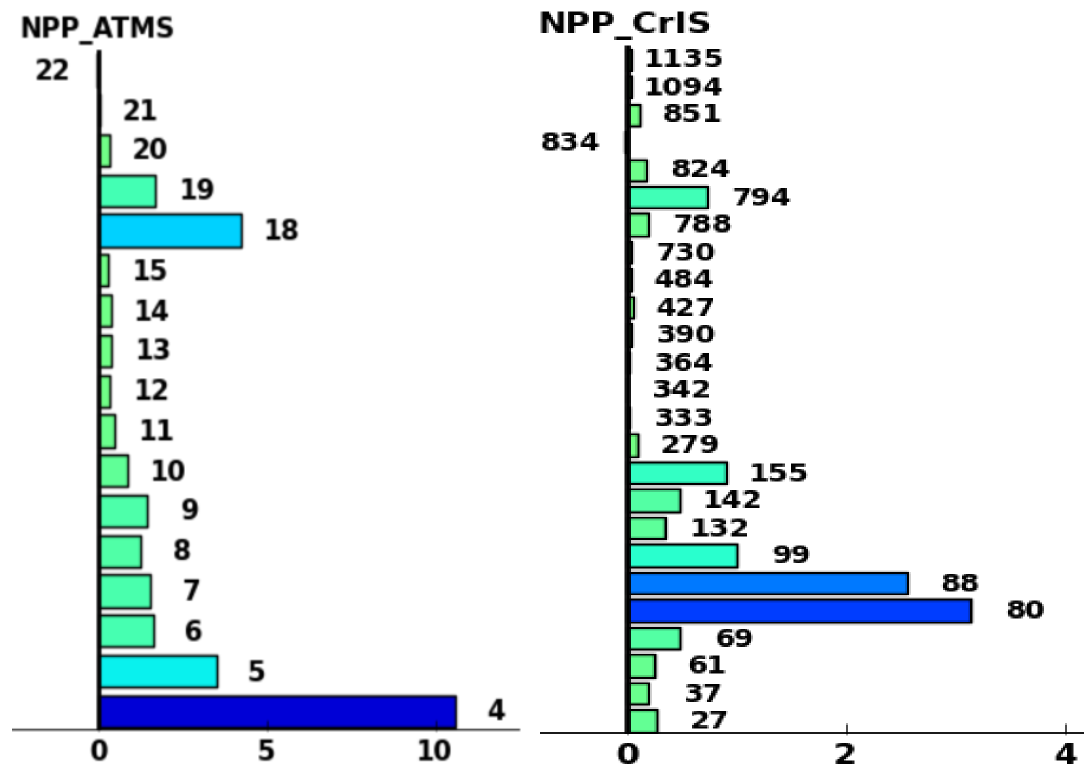


U.S. NAVAL RESEARCH LABORATORY

ATMS and CrIS Impact

NAVDAS-AR Observation Sensitivity





24 hr. Forecast Error Norm Reduction (J/kg)

Outline

- High-impact Satellite Observations
- 4D Variational Data Assimilation Systems
- Observation Error and its Covariance Estimation
- Satellite Interchannel Error Characteristics and Practical Implementation
- Experiments and Forecast Results
- CrIS and Posterior Channel Selection
- Conclusions and Future Work

4DVar Primal Formulation

$$\underline{w} \equiv \underline{x} - \underline{x}_f = BH^T \boxed{\left(HBH^T + R\right)^{-1}\left(\underline{y} - H\underline{x}_f\right) = \underline{z}}$$

$$\left(B^{-1} + H^T R^{-1} H\right)\underline{w} = \left(B^{-1} + H^T R^{-1} H\right) BH^T \left(HBH^T + R\right)^{-1}\left(\underline{y} - H\underline{x}_f\right)$$

$$\left(B^{-1} + H^T R^{-1} H\right)\underline{w} = H^T R^{-1}\left(\underline{y} - H\underline{x}_f\right)$$

4DVar Dual Formulation

$$\boxed{(HBH^T + \tilde{R})\underline{z} = (\underline{y} - H\underline{x_b})}$$

$$\tilde{R}^{-1/2}(HBH^T + \tilde{R})\underline{z} = \tilde{R}^{-1/2}(\underline{y} - H\underline{x_b})$$

Outline

- High-impact Satellite Observations
- 4D Variational Data Assimilation Systems
- Observation Error and its Covariance Estimation
- Satellite Interchannel Error Characteristics and Practical Implementation
- Experiments and Forecast Results
- CrIS and Posterior Channel Selection
- Conclusions and Future Work

Current Practice

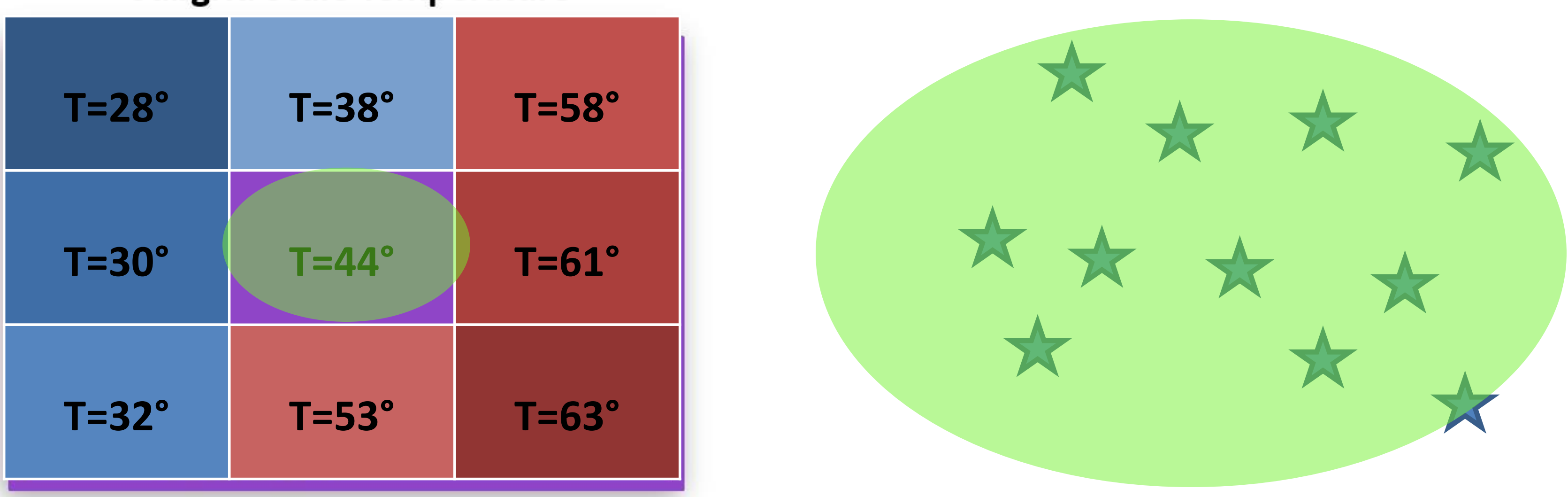
- **Until recently, most operational DA systems assumed no correlations between observation errors (i.e., a diagonal R), which is wrong**
- **To compensate for observation errors that *are* actually correlated, one or more of the following is typically done:**
  - **Discard ("thin") observations until the remaining ones are uncorrelated (Bergman and Bonner (1976), Liu and Rabier (2003))**
  - **Local averaging ("superobbing") (Berger and Forsythe (2004))**
  - **Inflate the observation error variances (Stewart et al. (2008, 2013)**
- **Theoretical studies (e.g. Stewart et al., 2009) indicate that including even *approximate* correlation structures outperforms diagonal R with variance inflation**
- ***In January, 2013, the Met Office went operational with a vertical observation error covariance submatrix for the IASI instrument, which showed forecast benefit in seasonal testing in both hemispheres (Weston et al. (2014))**
- **The current version of NAVGEM (v1.4) has the capability to account for interchannel correlated observation error for the ATMS, CrIS, and IASI instruments**

Sources of Observation Error

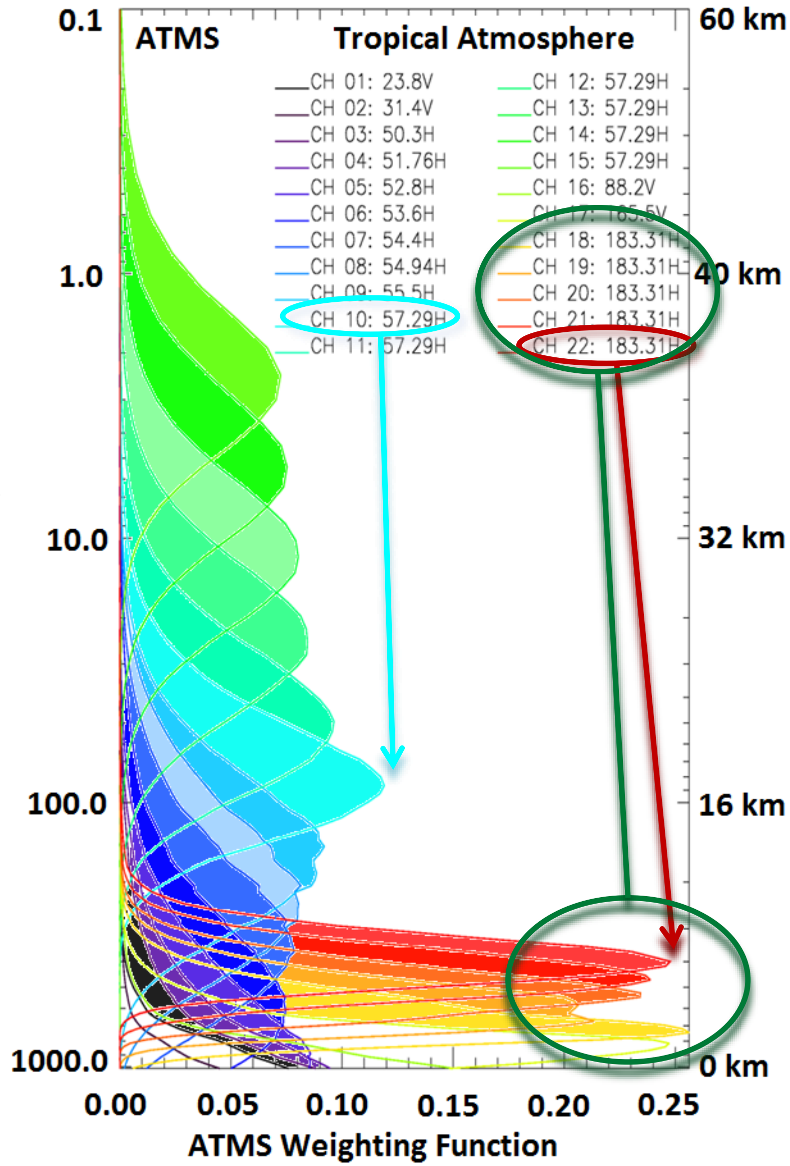
1) Instrument error (usually, but not always, uncorrelated)
2) Mapping operator (**H**) error (interpolation, radiative transfer)
3) Pre-processing, quality control, and bias correction errors
4) Error of representation (sampling or scaling error), which can lead to correlated error:

True Temperature in Model Space

Subgrid Scale Temperature



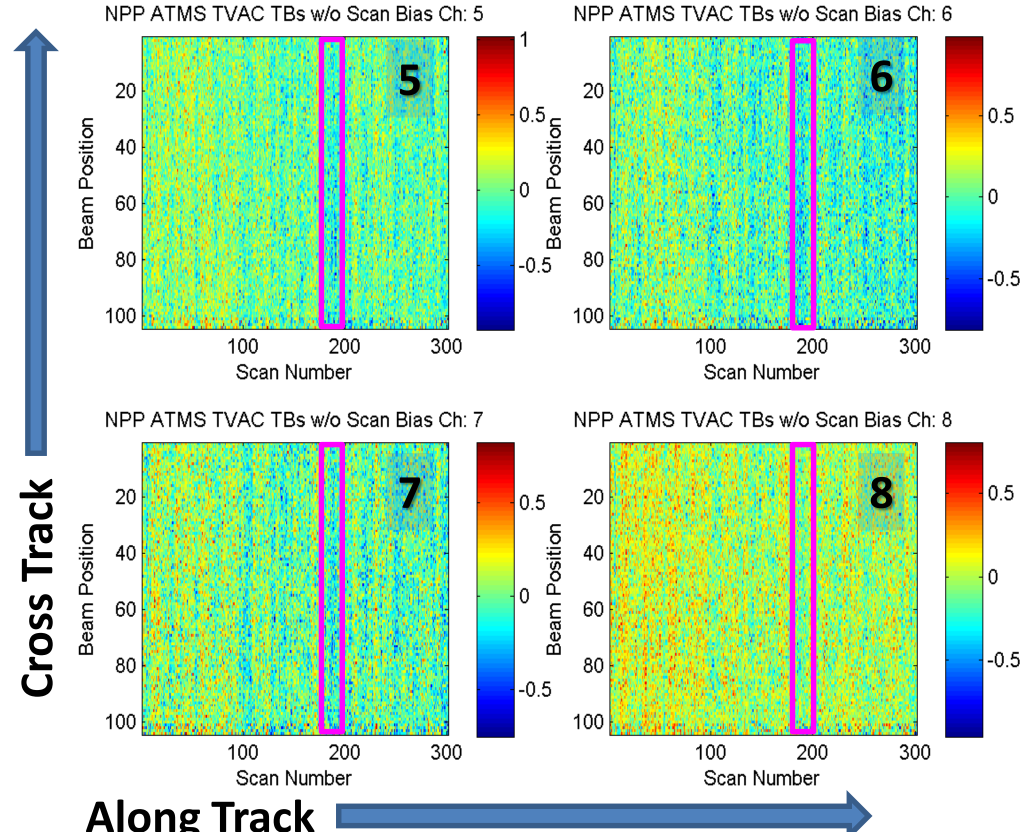
Correlated Instrument Error



Advanced Technology Microwave Sounder (ATMS)

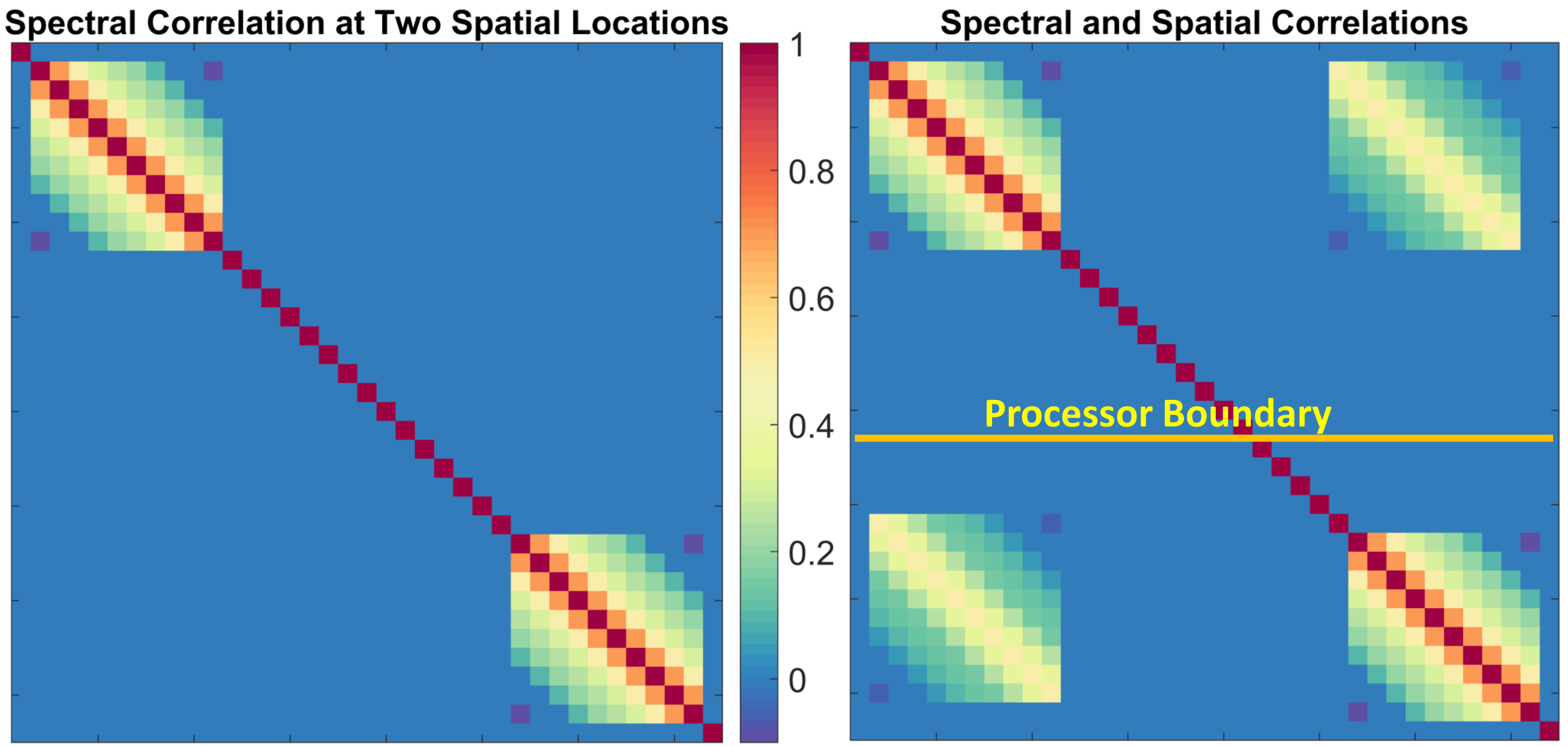
13 temperature channels

9 moisture channels



Spectral and Spatial Correlation

- The Desroziers method can estimate both spectral (interchannel) and spatial correlations; here we are only concerned with interchannel correlations
- Accounting for spatial correlations is significantly more complex
- There are important scaling considerations when dealing with 2D correlations



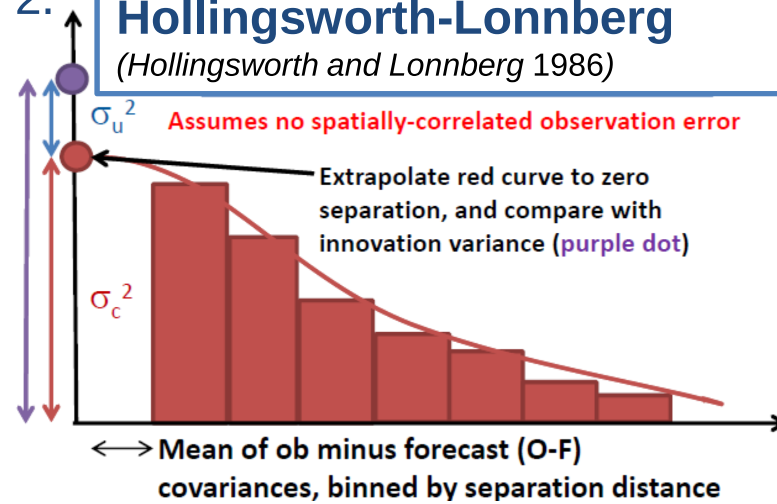
Covariance Estimation Methods

- **Several methods exist which can inform estimates of the background and/or observation error covariance matrices**
- **All methods have free parameters and make different assumptions; none are clearly superior to the others.**
- **Knowledge of when and how each method may produce sub-optimal results is the subject of current research.**

2.

Hollingsworth-Lonnberg

(Hollingsworth and Lonnberg 1986)



1.

Desroziers' Method

(Desroziers *et al.* 2005)

$$\left\langle (\vec{O}-\vec{F})(\vec{O}-\vec{A})^T \right\rangle = \mathbf{R}$$

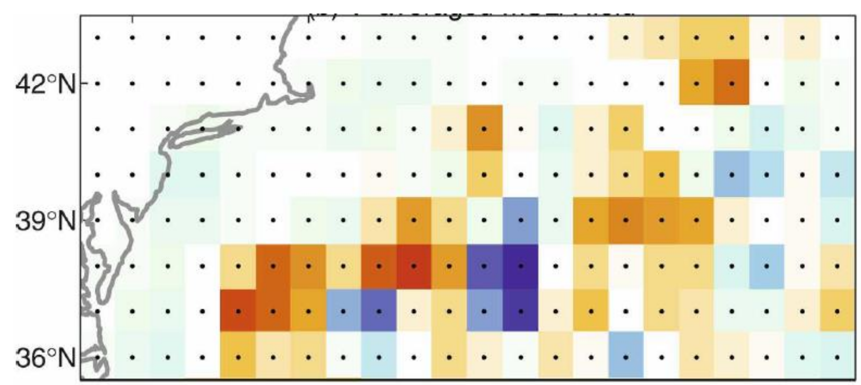
$$\left\langle (\vec{A}-\vec{F})(\vec{O}-\vec{F})^T \right\rangle = \mathbf{HBH}^T$$

$$\left\langle (\vec{O}-\vec{F})(\vec{O}-\vec{F})^T \right\rangle = \mathbf{R} + \mathbf{HBH}^T$$

3.

Observation Based Methods

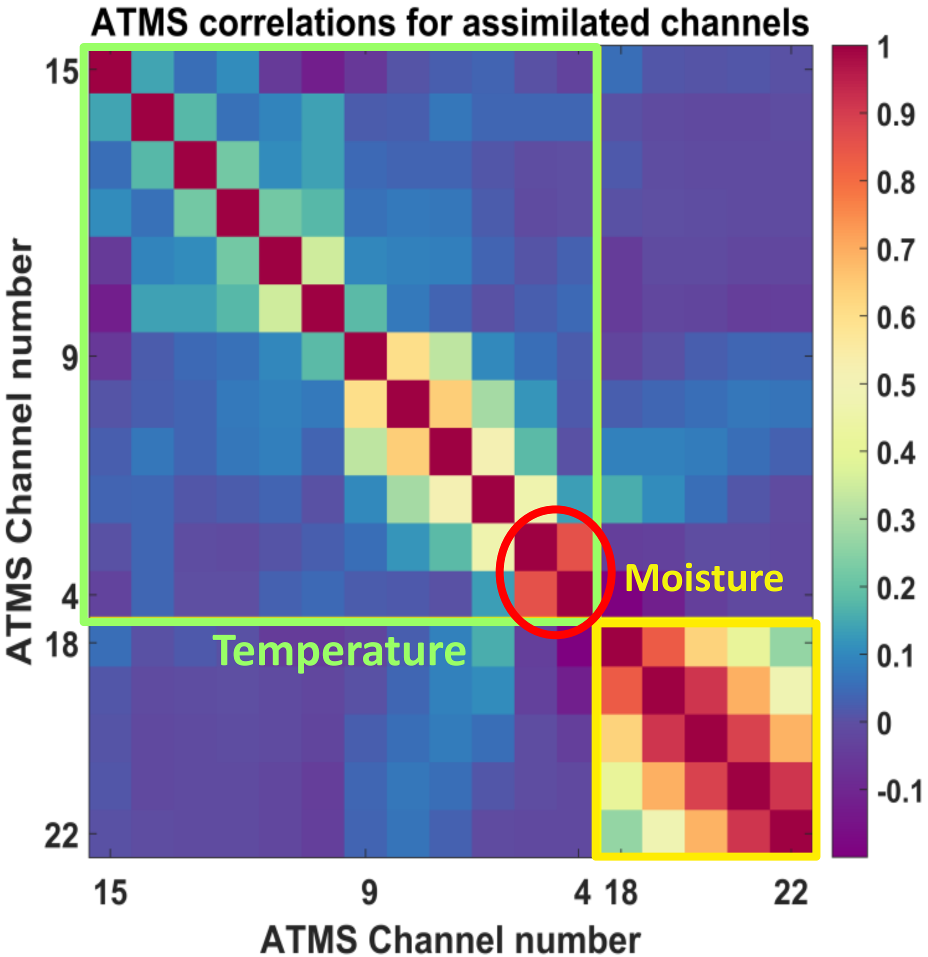
(e.g. Oke and Sakov 2007)

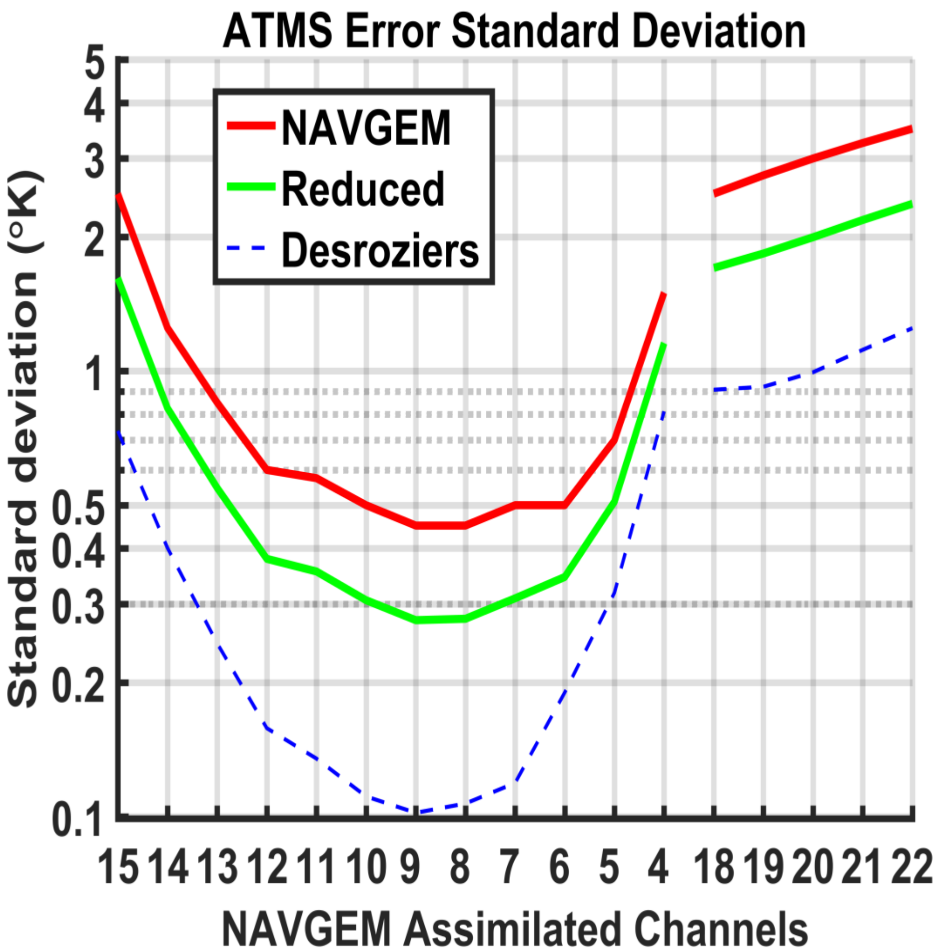


Outline

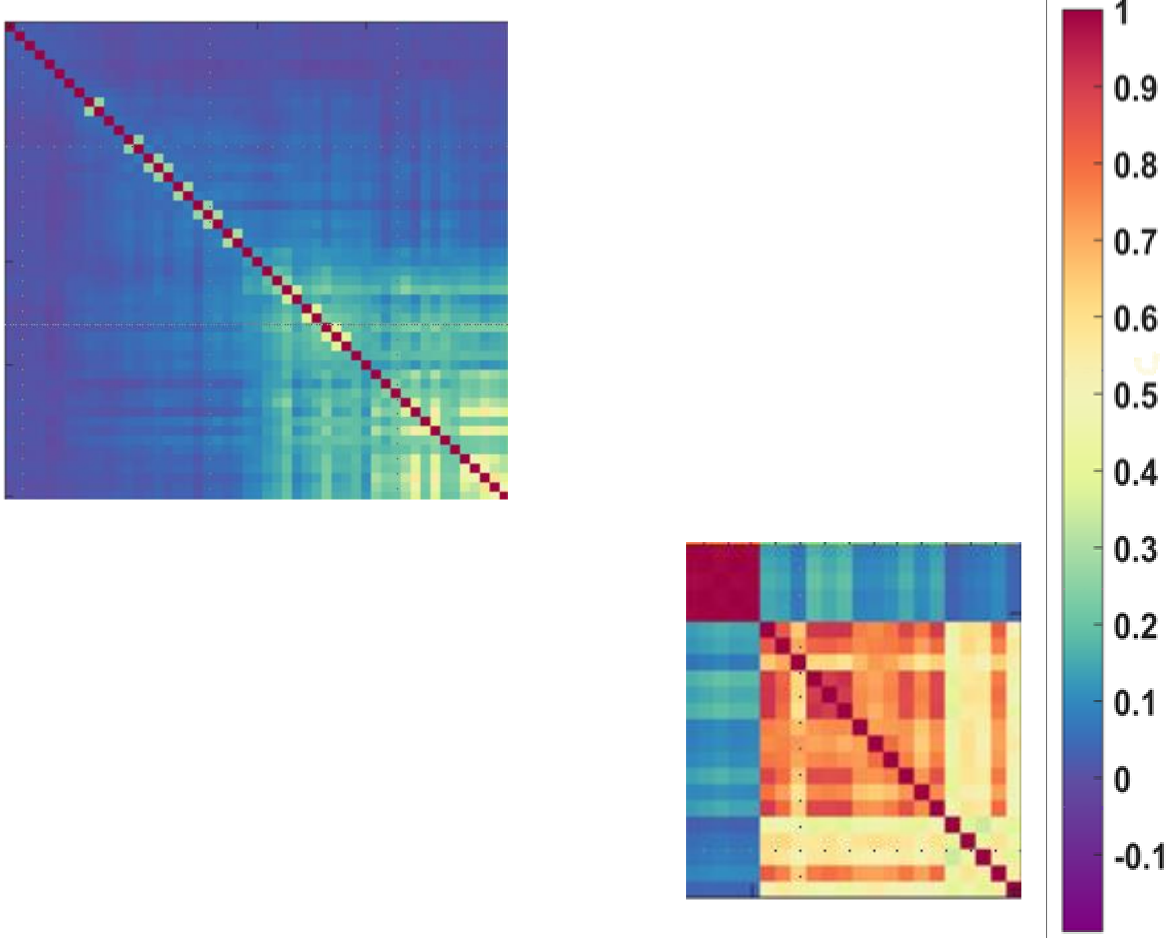
- High-impact Satellite Observations
- 4D Variational Data Assimilation Systems
- Observation Error and its Covariance Estimation
- Satellite Interchannel Error Characteristics and Practical Implementation
- Experiments and Forecast Results
- CrIS and Posterior Channel Selection
- Conclusions and Future Work

ATMS Error Characteristics





IASI Error Characteristics



Condition Number and Solver Convergence

- The condition number of a matrix X is defined by $\sigma_{max}(X)/\sigma_{min}(X)$, which is the ratio of the maximum singular value of X to the minimum one. (Singular value == eigenvalue for symmetric, positive definite X)
- Adding correlated error increases the condition number, slowing down convergence of the solver
- We can control how long the solver takes by constructing an approximate matrix with *any condition number we choose*

Reconditioning Methods

1. **Multiplicative preconditioning** by diagonal scaling matrices
2. **Increase the diagonal** values (additively) of the matrix (e.g. Weston et al. (2014))
3. **Find the optimal linear combination** of the sample covariance matrix and the scaled identity matrix (optimal Steinian linear shrinkage –similar to 2.)
4. **Find a positive definite approximation** to the matrix by altering the eigenvalue spectrum (constrained minimization of the Ky-Fan p-k norm)
5. **Channel selection** to eliminate channels with highly correlated error

Outline

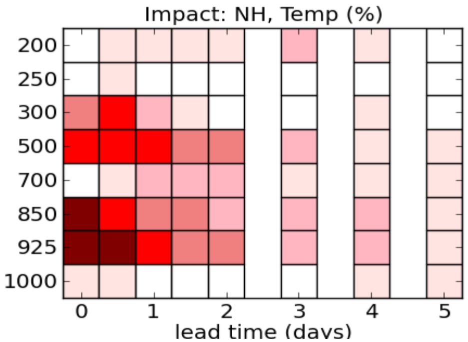
- High-impact Satellite Observations
- 4D Variational Data Assimilation Systems
- Observation Error and its Covariance Estimation
- Satellite Interchannel Error Characteristics and Practical Implementation
- Experiments and Forecast Results
- CrIS and Posterior Channel Selection
- Conclusions and Future Work

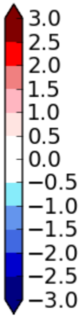
Experimental Design

Experiment	ATMS cn	IASI cn	Mean Iter	Description
atid	1	1	56	Control run, no correlated error for ATMS or IASI, default NAVGEM diag(R)
atmsiasi	18	169	78	Ky-Fan reconditioned Desroziers correlation, default NAVGEM diag(R)
Drzatmsiasi	18	169	104	Ky-Fan reconditioned Desroziers correlation, compromise diag(R)
Wesboth	18	169	87	Additive reconditioned Desroziers correlation, compromise diag(R)

Two months cycling data assimilation runs (Jun-Jul 2013) with T425L60 NAVGEM v1.3. Model top at 0.04 hPa, variational bias correction, CRTM v2.2, IASI data thinned to 110km, ATMS to 48km at nadir(every 3rd beam position and scanline), moisture channels used only over ocean, Watts-McNally type cloud detection (NWP-SAF package) for AIRS and IASI.

Correlated ATMS & IASI Assimilation Results

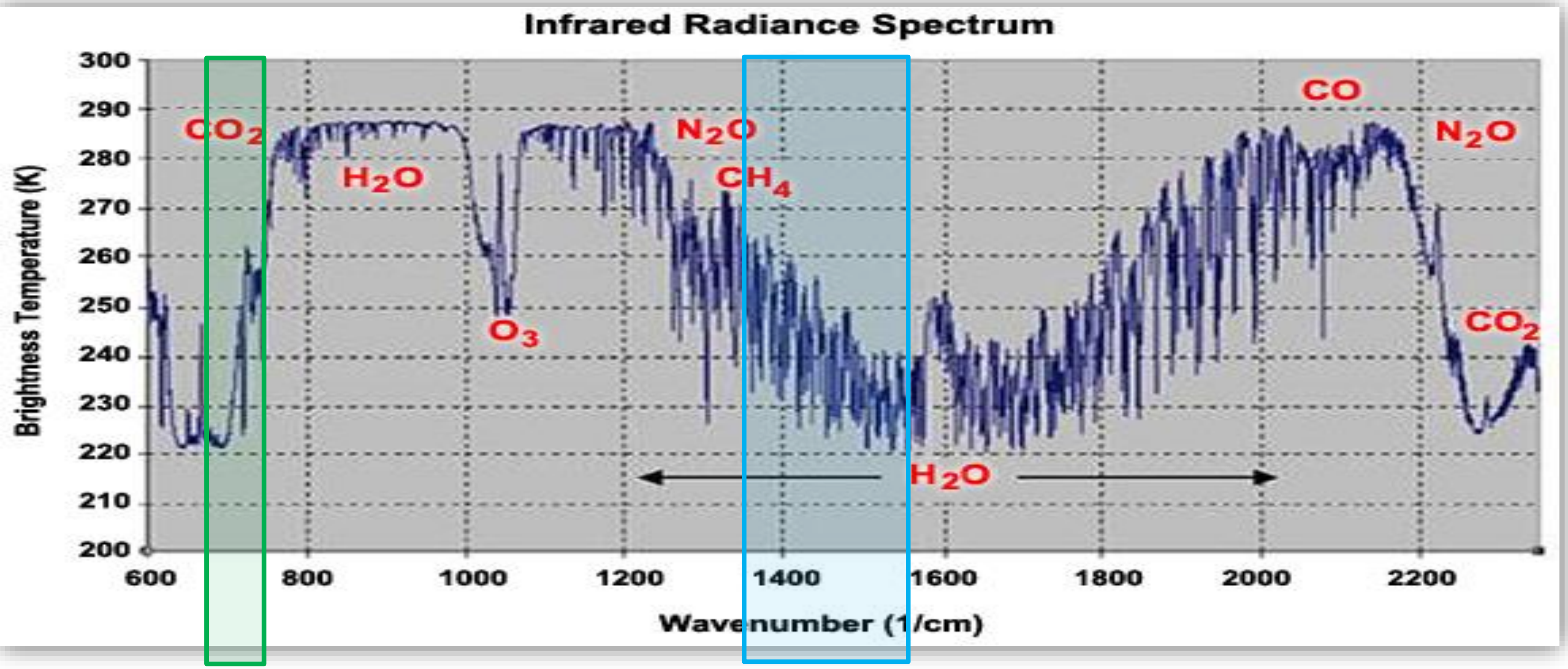
	T	u	Z	q
NH				
Tropics				
SH				



Outline

- High-impact Satellite Observations
- 4D Variational Data Assimilation Systems
- Observation Error and its Covariance Estimation
- Satellite Interchannel Error Characteristics and Practical Implementation
- Experiments and Forecast Results
- CrIS and Posterior Channel Selection
- Conclusions and Future Work

Spectrum of Assimilated IASI Channels

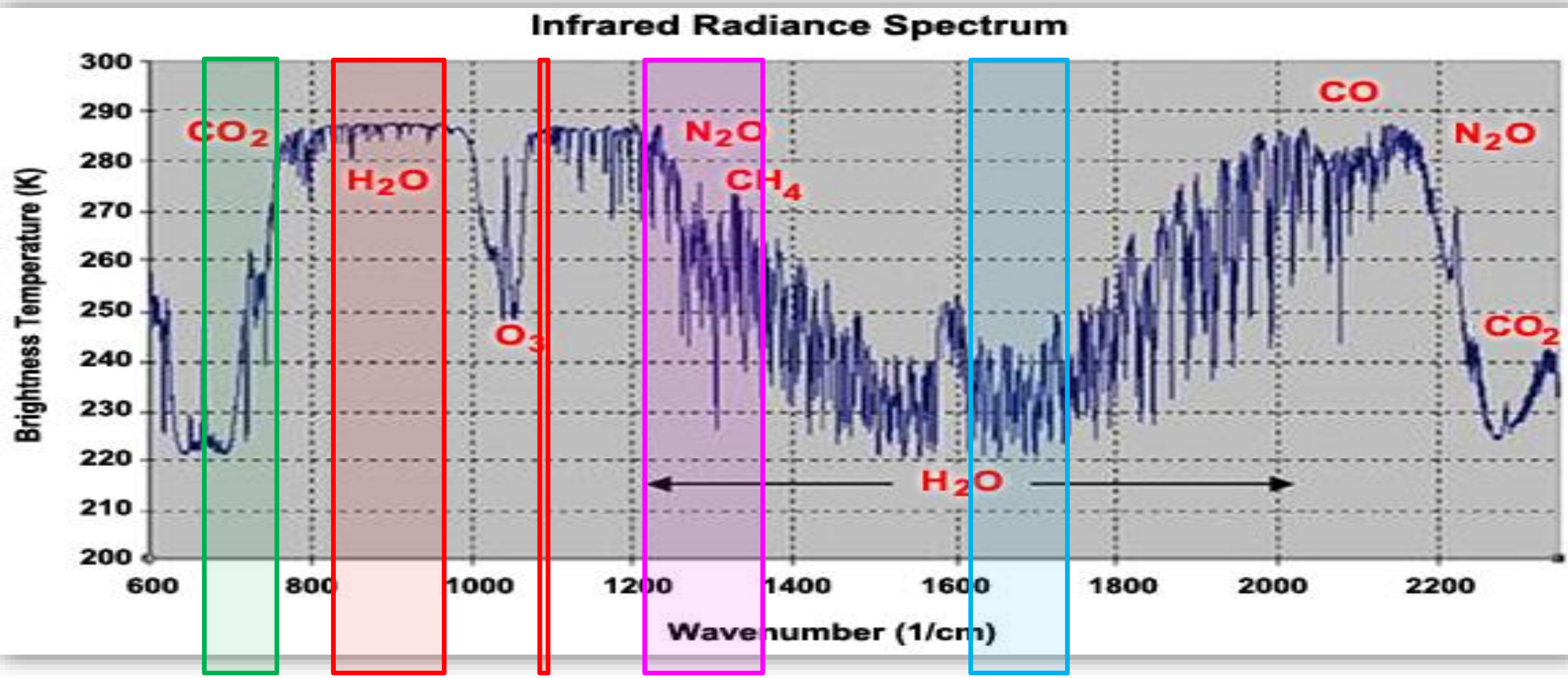


675.25 - 739.50 cm^{-1}

1367.00 - 1560.00 cm^{-1}

69 IASI Channels (55 CO_2, 14 H_2O) assimilated in NAVGEM

Spectrum of Assimilated CrIS Channels



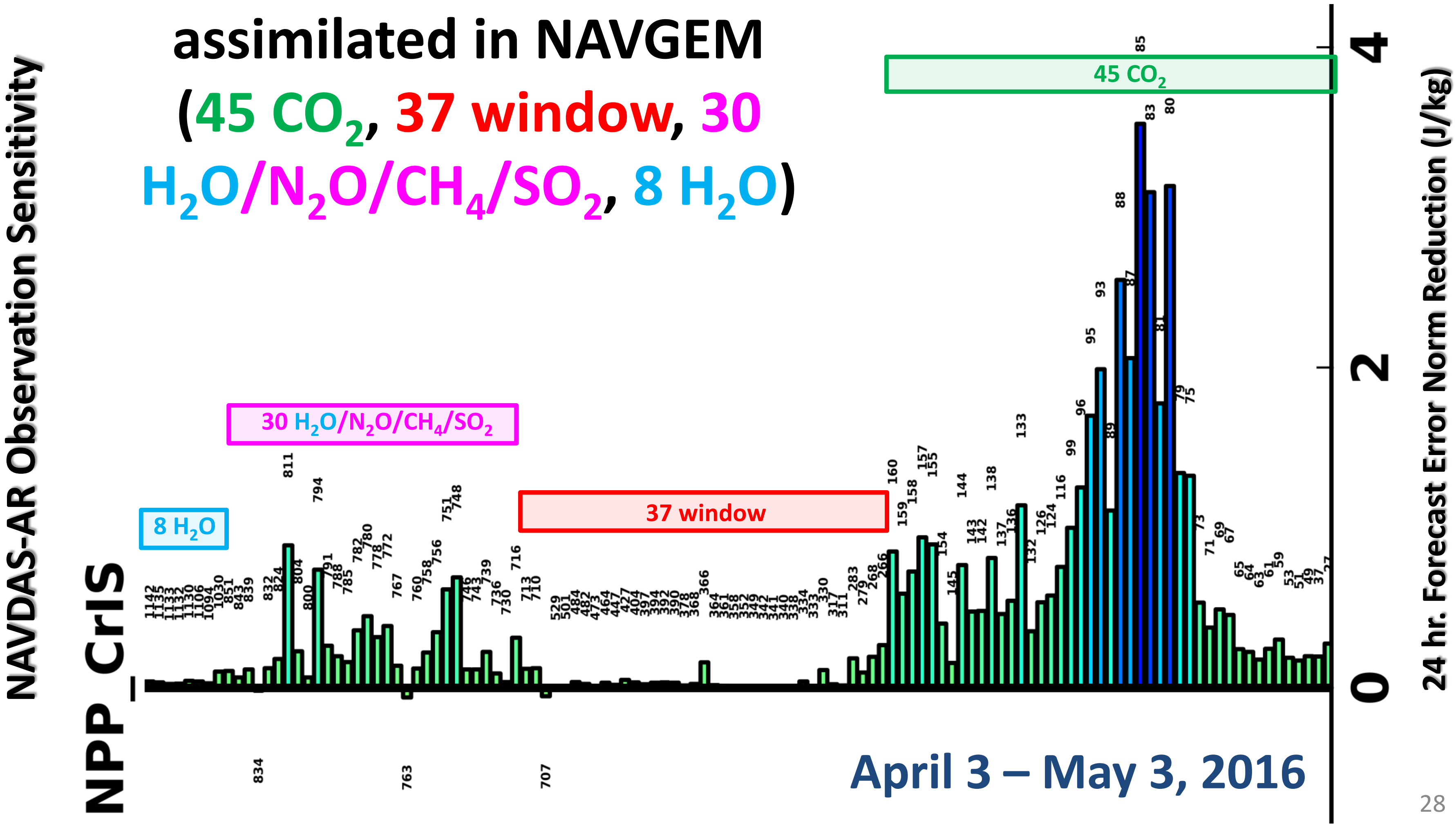
666.25 - 749.38 cm^{-1} 815.63–980, 1091-1095.00 cm^{-1} 1212.00 - 1381.00 cm^{-1} 1605.00 - 1745.00 cm^{-1}

120 CrIS Channels assimilated in NAVGEM

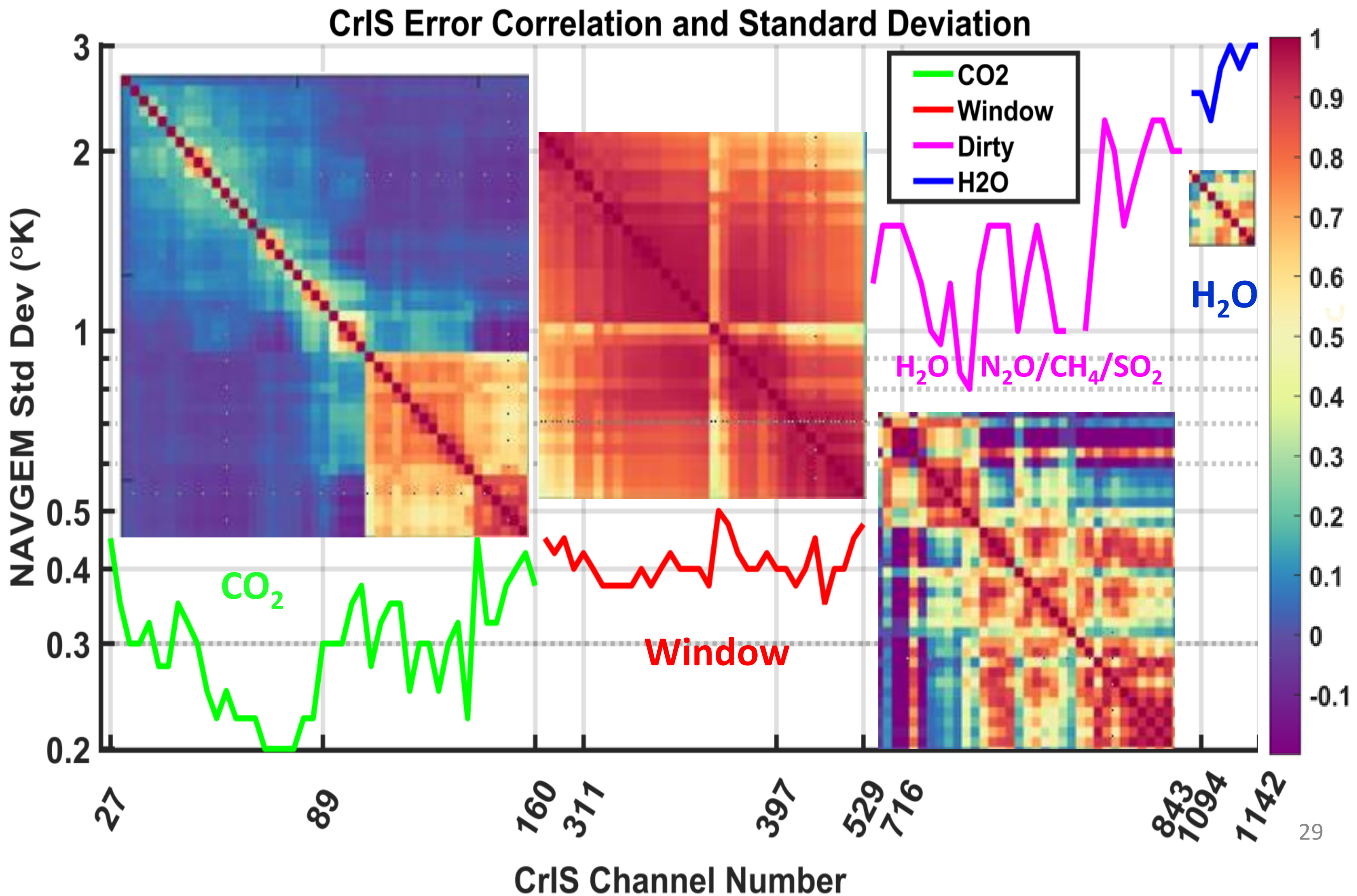
(45 CO_2, 37 window, 30 $H_2O/N_2O/CH_4/SO_2$, 8 H_2O)

CrIS Details by Channel

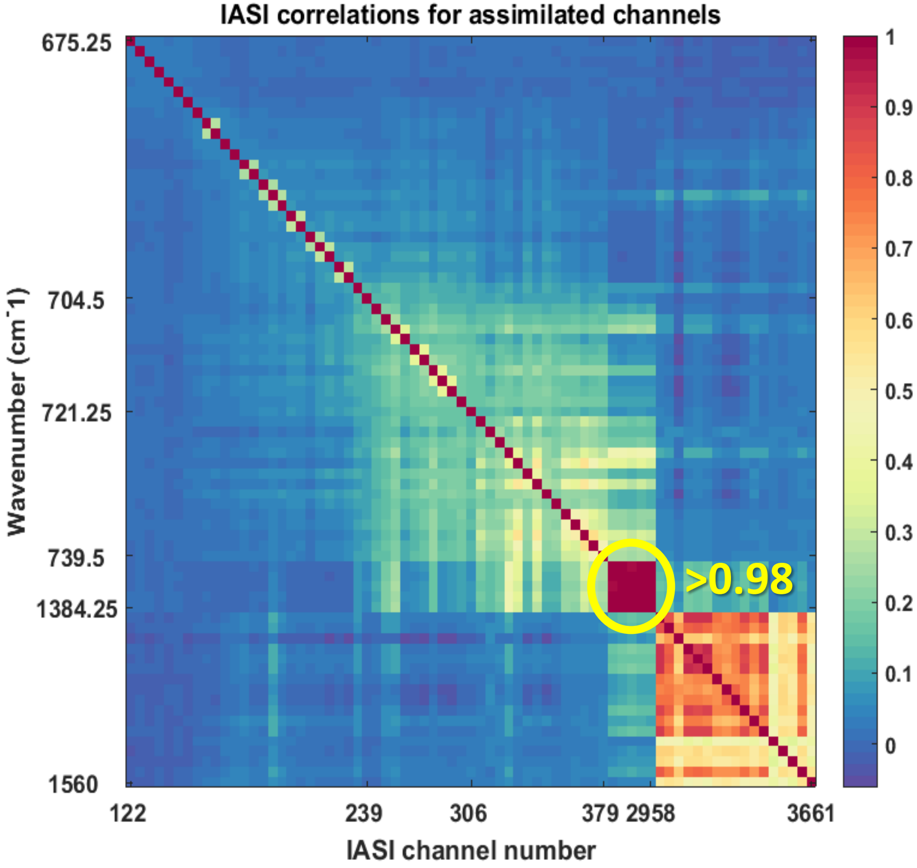
120 CrIS Channels assimilated in NAVGEM ($45\ CO_2$, 37 window, 30 $H_2O/N_2O/CH_4/SO_2$, $8\ H_2O$)



CrIS Error Characteristics



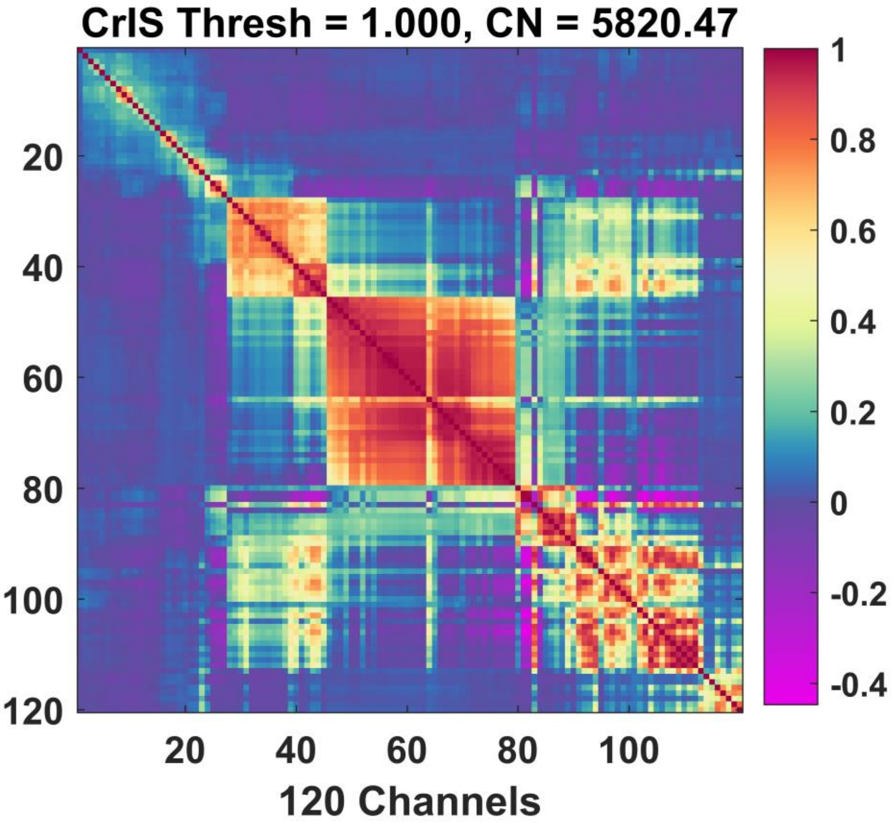
Posterior Channel Selection



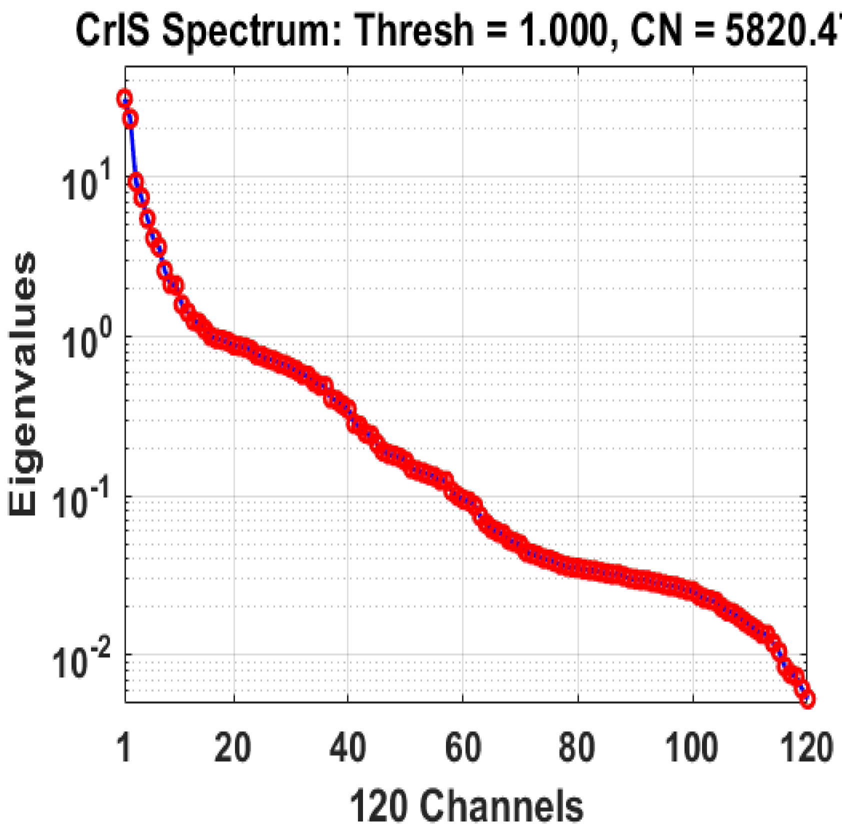
CrIS Channel Selection

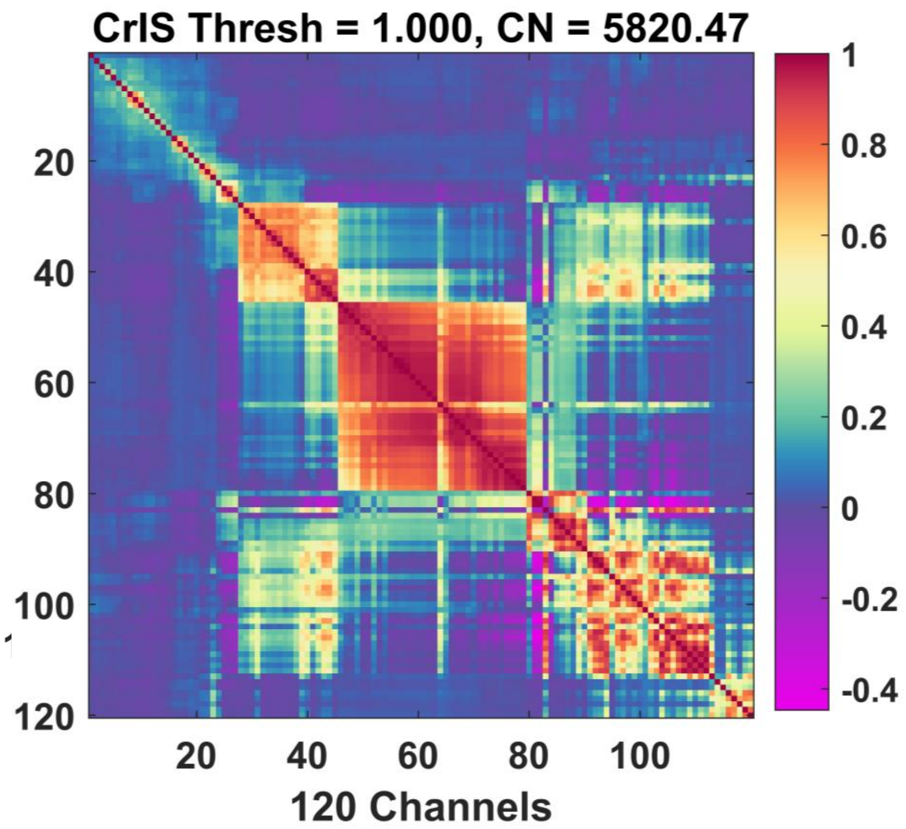
We may benefit from a similar procedure for CrIS channel selection with specified maximum error correlation:

- Compute the 2-norm of each row of the correlation matrix that has at least one value exceeding a set threshold
- Eliminate the channel with the largest norm
- Repeat until no correlations exceed the threshold



CrIS Channel Selection and Eigenspectrum Change





Outline

- High-impact Satellite Observations
- 4D Variational Data Assimilation Systems
- Observation Error and its Covariance Estimation
- Satellite Interchannel Error Characteristics and Practical Implementation
- Experiments and Forecast Results
- CrIS and Posterior Channel Selection
- Conclusions and Future Work

Conclusions

- The Desroziers error covariance estimation method can quantify correlated observation error
- Minimal changes can be made to the estimated error correlations to fit operational time constraints, and may improve the solution quality
- Accounting for correlated error for ATMS and IASI yields an unexpectedly large forecast benefit for a relatively small computational cost
- Because true error covariances have been taken into account, prescribed error variances can be reduced, which lead to further gains in forecast skill
- When a set of channels have highly correlated error, we may see no loss in skill by excluding some members of that set

Future Plans

- By accounting for correlated error for CrIS, additional forecast skill gains are expected
- We plan to use the Desroziers diagnostic to alter our channel selection and simultaneously decrease the condition number of the CrIS observation error correlation matrix
- ATMS, IASI, and CrIS error correlation matrices must be recomputed for the hybrid 4DVar (NAVGEM v1.4), retested, and retuned for the new system to regain the benefits seen in NAVGEM v1.3
- Correlated error in space or time?

Questions?

Correlated ATMS & IASI Assimilation Results

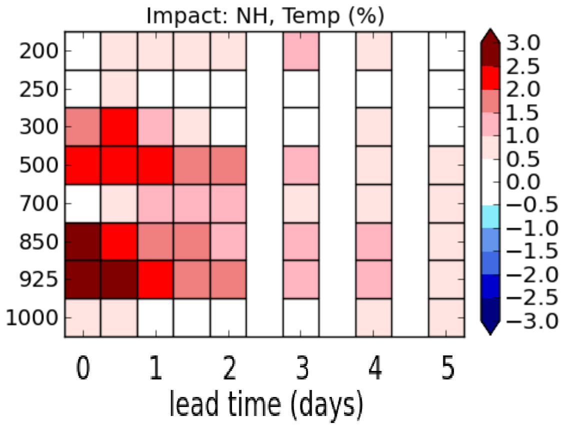
T

u

Z

q

NH



Tropics

SH