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
• 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
April 3 – May 3, 2016

24 hr. Forecast Error Norm Reduction (J/kg)
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 $H_2O$ vapor channels at 166 and 5 about the 183 GHz line vs. MHS
• 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):
    • mid-wave band (MWIR):
      – spectral range: 5.71 - 8.26µm
    • short-wave band (SWIR):
      – spectral range: 3.92 - 4.64 µm
ATMS and CrIS Impact

NAVDAS-AR Observation Sensitivity

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

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

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

\[ \left( \underline{B}^{-1} + H^T R^{-1} H \right) \underline{w} = H^T R^{-1} \left( \underline{y} - H \underline{x}_f \right) \]
Change of variables

The correlation matrix can be used directly – no inverse required.

This problem is solved iteratively.

\[
(HBH^T + \tilde{R})\tilde{z} = (y - Hx_b)
\]

\[
\tilde{R}^{-1/2} (HBH^T + \tilde{R})\tilde{z} = \tilde{R}^{-1/2} (y - Hx_b)
\]
• 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
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**

<table>
<thead>
<tr>
<th>Subgrid Scale Temperature</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>T=28°</td>
<td>T=38°</td>
</tr>
<tr>
<td>T=30°</td>
<td>T=44°</td>
</tr>
<tr>
<td>T=32°</td>
<td>T=53°</td>
</tr>
</tbody>
</table>
Correlated Instrument Error

Advanced Technology Microwave Sounder (ATMS)

13 temperature channels
9 moisture channels
• 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
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.

1. Desrozières’ Method  
(Desrozières et al. 2005)

\[
\langle (\mathbf{O} - \bar{F})(\mathbf{O} - \bar{A})^T \rangle = \mathbf{R} \\
\langle (\mathbf{A} - \bar{F})(\mathbf{O} - \bar{F})^T \rangle = \mathbf{HBH}^T \\
\langle (\bar{O} - \bar{F})(\bar{O} - \bar{F})^T \rangle = \mathbf{R} + \mathbf{HBH}^T
\]

2. Hollingsworth-Lonnberg  
(Hollingsworth and Lonnberg 1986)

Assumes no spatially-correlated observation error
Extrapolate red curve to zero separation, and compare with innovation variance (purple dot)

Mean of ob minus forecast (O-F) covariances, binned by separation distance

3. Observation Based Methods  
(e.g. Oke and Sakov 2007)
• 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

ATMS correlations for assimilated channels

ATMS Error Standard Deviation

- NAVGEM
- Reduced
- Desrozieras
IASI Error Characteristics
The condition number of a matrix $X$ is defined by $\sigma_{\text{max}}(X)/\sigma_{\text{min}}(X)$, which is the ratio of the maximum singular value of $X$ to the minimum one. (Singular value $\approx$ 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.
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
• 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
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.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>ATMS cn</th>
<th>IASI cn</th>
<th>Mean Iter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>atid</td>
<td>1</td>
<td>1</td>
<td>56</td>
<td>Control run, no correlated error for ATMS or IASI, default NAVGEM diag(R)</td>
</tr>
<tr>
<td>atmsiasi</td>
<td>18</td>
<td>169</td>
<td>78</td>
<td>Ky-Fan reconditioned Desroziers correlation, default NAVGEM diag(R)</td>
</tr>
<tr>
<td>Drzatmsiasi</td>
<td>18</td>
<td>169</td>
<td>104</td>
<td>Ky-Fan reconditioned Desroziers correlation, compromise diag(R)</td>
</tr>
<tr>
<td>Wesboth</td>
<td>18</td>
<td>169</td>
<td>87</td>
<td>Additive reconditioned Desroziers correlation, compromise diag(R)</td>
</tr>
</tbody>
</table>
Correlated ATMS & IASI Assimilation Results

Impact: NH, Temp (%)
• 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
69 IASI Channels (55 \( \text{CO}_2 \), 14 \( \text{H}_2\text{O} \)) assimilated in NAVGEM
Spectrum of Assimilated CrIS Channels

120 CrIS Channels assimilated in NAVGEM

(45 CO₂, 37 window, 30 H₂O/N₂O/CH₄/SO₂, 8 H₂O)
120 CrIS Channels
assimilated in NAVGEM
(45 CO$_2$, 37 window, 30 H$_2$O/N$_2$O/CH$_4$/SO$_2$, 8 H$_2$O)

April 3 – May 3, 2016
CrIS Error Characteristics

CrIS Error Correlation and Standard Deviation

- CO2
- Window
- Dirty
- H2O

NAVGEM Std Dev (°K)

CrIS Channel Number

H2O, N2O/CH4/SO2
Water vapor channels 2889, 2944, 2948, 2951, and 2958 have very high error correlation (>0.98).

The eigenvectors corresponding to the 5 smallest eigenvalues project only on to these 5 channels.

It makes sense to use the Desroziers diagnostic to do an a posteriori channel selection, to improve the condition number while discarding redundant information.

NAVGEM tests show the retaining only channel 2889 helps solver convergence with no discernable negative impact.

We now assimilate 69 IASI channels instead of 73.
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
• 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
• 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

Impact: NH, Temp (%)