Quantifying observation error correlations in remotely sensed data

Conference or Workshop Item

Published Version

Presentation slides

Optimising the use of satellite data in NWP through the inclusion of error correlation structures

L.M. Stewart, University of Reading

Co-authors: S.L. Dance and N.K. Nichols (UoR); J.Cameron, S.English and J.Eyre (Met Office UK)

Funding: NERC and Met Office UK

Adjoint Model Applications In Dynamic Meteorology Workshop 20 May 2009

Talk structure

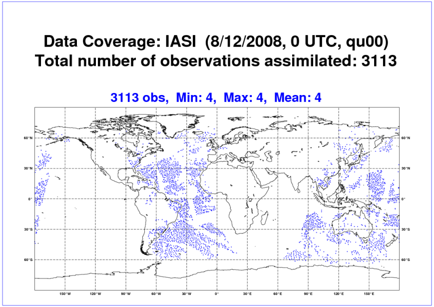
- ★ Introduction and theory
  - Motivation
  - Variational data assimilation
  - Observation error covariance matrices
- ★ Quantifying observation error correlations
  - Desroziers' method of statistical approximation
  - Application to IASI data
  - Results
- ★ Modelling observation error correlation structure
  - Approximate structures for R

What are observation error correlations?

Every observation *y* of a atmospheric variable *x* has an associated error ϵ: $y = Hx + \epsilon$

$\rightarrow$ observation error correlations are present when components of the error vector ϵ are related

$\rightarrow$ measurement errors are attributed to 3 sources: instrument noise, forward model error and representativity error



Error sources

- Instrument noise
  - temperature converted neδt value
  - regular calibrations ensure noise is uncorrelated between channels
- Forward model error
  - errors in discretisation of radiative transfer equation
  - errors in mis-representation of gaseous contributors
  - errors from undetected cloud
- Representativity error
  - contrasting model and observation resolutions
  - observations resolve spatial scales or features that the model cannot
  - contributes to cross channel observation error correlations

Why are correlations important?

Problems

- -ve magnitude and behaviour relatively unknown
- -ve reduce weighting of observations in analysis
- -ve for an observation vector of size 10^6, difficult to store and invert observation error matrix if correlations are included

Benefits

- +ve increase accuracy of gradients of the observed field represented in the analysis
- +ve works with the prior error covariance to specify how observation features should be smoothed
- +ve more information available from observations

Variational data assimilation

Assimilation objective

Model forecast + Observation data → State of atmosphere

Assimilation method Minimise a cost function which measures distance of a solution state x from the observations $y^o \in \mathbb{R}^m$ and the background field $x^b \in \mathbb{R}^n$

Cost Function

$$J(x) = \frac{1}{2}(x - x^b)^T \mathbf{B}^{-1}(x - x^b) + \frac{1}{2}(y^o - H(x))^T \mathbf{R}^{-1}(y^o - H(x))$$

where **B** and **R** are the background and observation error covariance matrices respectively

An error covariance matrix structure

The observation error covariance matrix takes the form:

$$\mathbf{R} = \mathbf{D}^{1/2}\mathbf{C}\mathbf{D}^{1/2}$$

where **C** is the error correlation matrix

$$\mathbf{C} = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1m} \\ \rho_{12} & 1 & \cdots & \rho_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ \rho_{1m} & \rho_{2m} & \cdots & 1 \end{pmatrix}$$

and **D** is the error variance matrix

$$\mathbf{D} = \begin{pmatrix} {\sigma^2}_1 & 0 & \cdots & 0 \\ 0 & {\sigma^2}_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & {\sigma^2}_m \end{pmatrix}$$

Quantifying cross-channel correlations: a study

Objective

Generate the true observation error correlation structure for a sample set of remotely sensed data typical of NWP

Data type

- IASI (infrared atmospheric sounding interferometer) observations
- measurements of the infrared radiation emitted by the earth's surface and atmosphere at different wavelengths

Method

We use a post analysis diagnostic derived from variational data assimilation theory [Desroziers, 2005]

Desroziers' method of statistical approximation

Recall the background state, x_b, and observation vector, y, are approximations to the true state of the atmosphere, x_t,

$$\begin{aligned} y &= Hx_t + \epsilon^o \\ x_t &= x_b + \epsilon^b \end{aligned}$$

where ϵ^o and ϵ^b are the observation and background errors respectively.

The best linear unbiased estimate of the true state, x_a, is given by

$$\begin{aligned} x_a &= x_b + \mathbf{K}(y - Hx_b) = x_b + \mathbf{K}d_b^o \\ \mathbf{K} &= \mathbf{BH}^T(\mathbf{HBH}^T + \mathbf{R})^{-1} \end{aligned}$$

Desroziers' method of statistical approximation

Innovation vector

$$\begin{aligned} d^o_b = y - Hx_b &= Hx_t + \epsilon^o - Hx_b \\ &\approx \epsilon^o + \mathbf{H}\epsilon^b \end{aligned}$$

Analysis innovation vector

$$\begin{aligned} d^o_a = y - Hx_a &= y - H(x_b + Kd^o_b) \\ &\approx (\mathbf{I} - \mathbf{HK})d^o_b \\ &\approx \mathbf{R}(\mathbf{HBH}^T + \mathbf{R})^{-1}d^o_b \end{aligned}$$

Desroziers' method of statistical approximation

Taking the expectation of the cross product of d_a^o and d_b^o, and assuming

$$\mathbb{E}[\epsilon^o(\epsilon^b)^T] = \mathbb{E}[\epsilon^b(\epsilon^o)^T] = 0,$$

we find a statistical approximation for the observation error covariances

$$\begin{aligned}
\mathbb{E}\left[d_a^o(d_b^o)^T\right] &\approx \mathbb{E}\left[\mathbf{R}(\mathbf{HBH}^T + \mathbf{R})^{-1} d_b^o(d_b^o)^T\right] \\
&\approx \mathbf{R}(\mathbf{HBH}^T + \mathbf{R})^{-1}\mathbb{E}\left[(\epsilon^o + H\epsilon^b)(\epsilon^o + H\epsilon^b)^T\right] \\
&\approx \mathbf{R}(\mathbf{HBH}^T + \mathbf{R})^{-1}(\mathbf{R} + \mathbf{HBH}^T) \\
&\approx \mathbf{R}
\end{aligned}$$

Application to IASI data

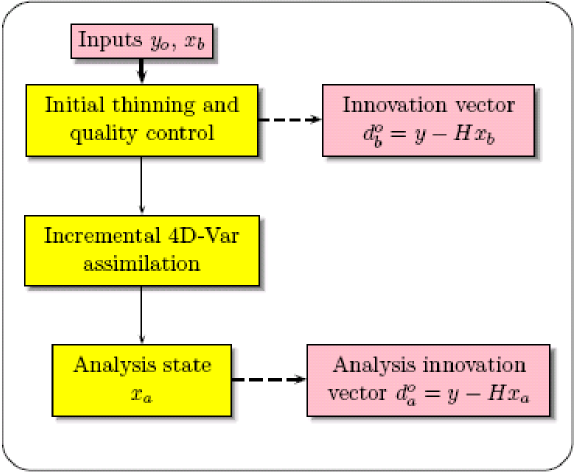


Figure: Assimilation process

Observation error correlation matrix

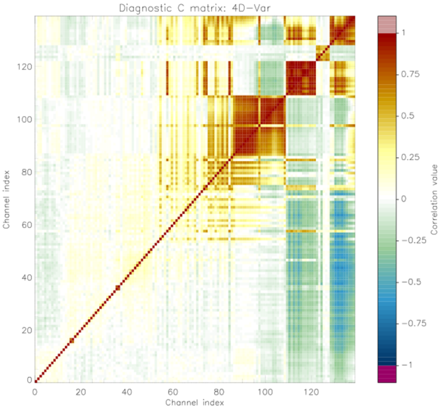
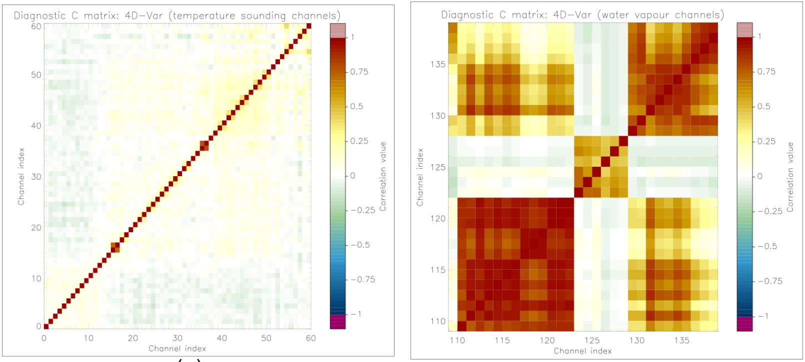


Figure: Error correlation matrix for 139 channels used in Var

Observation error correlation matrix



(a) (b)

Figure: Error correlation matrix for (a) temperature sounding channels; (b) water vapour channels

Operational and diagnosed error variances

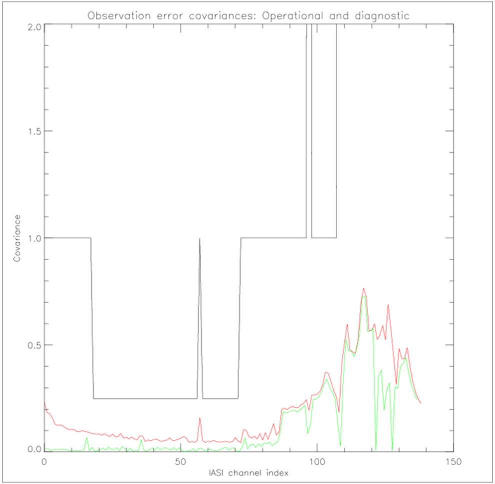


Figure: Operational error variances (black line), diagnosed error variances (red line), and first off-diagonal error covariance (green line)

Diagnosed error variances: comparison with Hollingsworth-Lonnberg (H-L) method

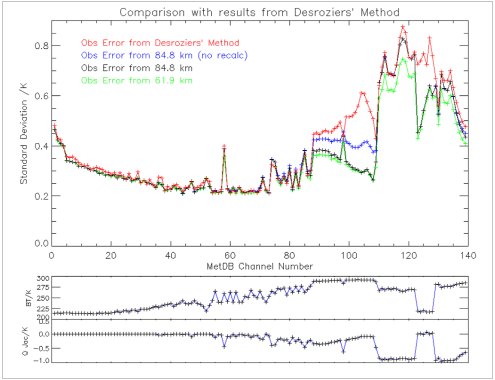


Figure: Diagnosed error variances (red line), H-L diagnosed error variances for 84.8km (blue and black line) and 61.9km (green line) separation. Plot provided by James Cameron, UK Met Office.

Quantifying cross-channel correlations: a summary

- ★ Strong off-diagonal correlations are present between channels with similar spectral properties
- ★ Channels highly sensitive to water vapour have large observation error variances and covariances
- ★ The observation error variance is being overestimated in current asimilation algorithms
- ★ Diagnosed error variances are comparable with those using the H-L diagnostic
- ★ Non-symmetric matrices! → future work

Modelling error correlation structure

What next?

Investigate how to approximate the true error correlation structure within operational assimilation methods...

Current approaches

- a diagonal matrix approximation
- diagonal variance inflation

Alternative approaches

- a Markov error covariance approximation
- a truncated eigendecomposition approximation [Fisher, 2005]
- a Toeplitz to circulant matrix approximation [Healy, 2005]

A desirable error covariance matrix

Main issue in observation error correlation modelling

- ★ need to calculate matrix-vector product $\mathbf{R}^{-1}(y^o - Hx)$ every time we calculate cost function J
- ★ relatively easy if $\mathbf{R} = \mathbf{D} \equiv m$ scalar multiplications
- ★ **BUT** $y \in \mathbb{R}^{10^6}$ and so $\mathbf{R} \in \mathbb{R}^{10^6 \times 10^6}$ which, if dense, is impossible to store and invert

The perfect partner: what do we want from $\mathbf{R} \neq \mathbf{D}$?

- ♥ structure resulting in an $\mathbf{R}^{-1}$ suitable for storage / can be used cheaply in a matrix-vector product
- ♥ representative of the true error correlation structure
- ♥ greater access to information from the observations and improved analysis accuracy

A Markov error covariance approximation

Consider a Markov covariance matrix of the form

$$R_{ij} = \sigma^2 \rho^{|i-j|}, \quad \rho = \exp\left(-\frac{\delta z}{h}\right)$$

where σ^2 is the error variance, δz is the level spacing, and h is the length scale

This is equivalent to a correlation matrix of the form

$$C = \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^n \\ \rho & 1 & \rho & \dots & \rho^{n-1} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-2} \\ \ddots & \ddots & \ddots & & \ddots \\ \rho^n & \dots & \rho^2 & \rho & 1 \end{pmatrix}$$

A Markov error covariance approximation

The benefit of this choice is that C has a tri-diagonal inverse

$$C^{-1} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho & 0 & \dots & 0 \\ -\rho & 1+\rho^2 & -\rho & \dots & 0 \\ 0 & -\rho & 1+\rho^2 & \dots & 0 \\ \ddots & \ddots & \ddots & & \vdots \\ 0 & \dots & -\rho & 1+\rho^2 & -\rho \\ 0 & \dots & 0 & -\rho & 1 \end{pmatrix}$$

and therefore as does R: $R^{-1} = \frac{1}{\sigma^2} I \times C^{-1}$

No need to store and invert R!

An eigendecomposition approximation

Describe C by a truncated eigendecomposition using its leading eigenpairs

$$\tilde{R} = D^{1/2}\tilde{C}D^{1/2} = D^{1/2}\left(\alpha I + \sum_{k=1}^{K}(\lambda_k - \alpha)v_k v_k^T\right)D^{1/2}$$

where (λ_k, v_k) is an eigenvalue, eigenvector pair of C, K is the number of eigenpairs used, and α is chosen such that trace($\tilde{R}$)=trace(D) [Fisher, 2005]

This matrix also has an easily attainable inverse

$$\tilde{R}^{-1} = D^{-1/2}\left(\alpha^{-1} I + \sum_{k=1}^{K}(\lambda_k^{-1} - \alpha^{-1})v_k v_k^T\right)D^{-1/2}$$

No need to store and invert R!

Summary

- ★ Observation error correlations are often created because of contrasting model and observation resolutions
- ★ Including observation error correlation structure can increase analysis accuracy and information content
- ★ In IASI data, observation error correlations are strongest bewteen channels with similar spectral properties
- ★ In IASI data, the largest observation error covariances are between channels highly sensitive to water vapour
- ★ In order to include observation error correlation structure in data assimilation algorithms, the **R** matrix must be suitably structured

References

- G. Desroziers and L. Berre and B. Chapnik and P. Poli. *Diagnosis of observation, background and analysis-error statistics in observation space*. Q.J.R.Meteorol.Soc., 131, 2005.
- M. Fisher. *Accounting for Correlated Observation Error in the ECMWF Analysis*. ECMWF Technical Memoranda, MF/05106, 2005.
- S.B. Healy and A.A. White. *Use of discrete Fourier transforms in the 1D-Var retrieval problem*. Q.J.R.Meteorol.Soc., 131, 2005.