Performance of IASI T and q retrievals from simulated spectra for PC-compressed data and 4D-Var channel selection

Plus... MetOp-B IASI

Fiona Smith, John Eyre and John Remedios* and James Cameron

*University of Leicester

Outline

- Motivation
  - Improving the information content of the assimilation
- Principal Components
- 1D-Var simulation experiments – overall performance
  - Operational 4D-Var channel selection
  - PC-RTTOV 500 Predictor channel selection
  - PC-RTTOV 200 PC Scores
  - 1D-Var simulation experiments – overall performance
- 1D-Var simulation experiments – minimisation

- MetOp-B
- SO_2

Motivation:
Increasing the information content of the 4D-Var assimilation

Hopes

- Before launch, IASI was predicted to give measurements accurate to
  - 1 K per km for temperature
  - 10% per km for humidity
- Pre-launch information content studies (Prunet; Collard) predicted approximately 20 pieces of independent information for NWP per observation
- This would be a vast increase in the level of information available from a non-hyperspectral satellite sounder. (c.f. 4 or so for HIRS)
- What is the information content of the current system?
- How do we access the rest of the information?

Variational analysis

- Variatio… co…

Component from *a priori*

Component from observation

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}[H(\mathbf{x})-\mathbf{y}_o]^T\mathbf{R}^{-1}[H(\mathbf{x})-\mathbf{y}_o]$$

- x is a model state vector
- x_b is the background st…
- y_o is the observation ve…
- H is the observation op…
- **H** is the linearised observation operator
- **B** is the background error covariance matrix
- **R** is the observation error covariance matrix

$$\mathbf{A} = (\mathbf{I} - \mathbf{KH})\,\mathbf{B}$$

- The analysis, x_a is given by

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}[\mathbf{y}_o - H(\mathbf{x}_b)]$$

$$\mathbf{K} = (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}$$

background profile

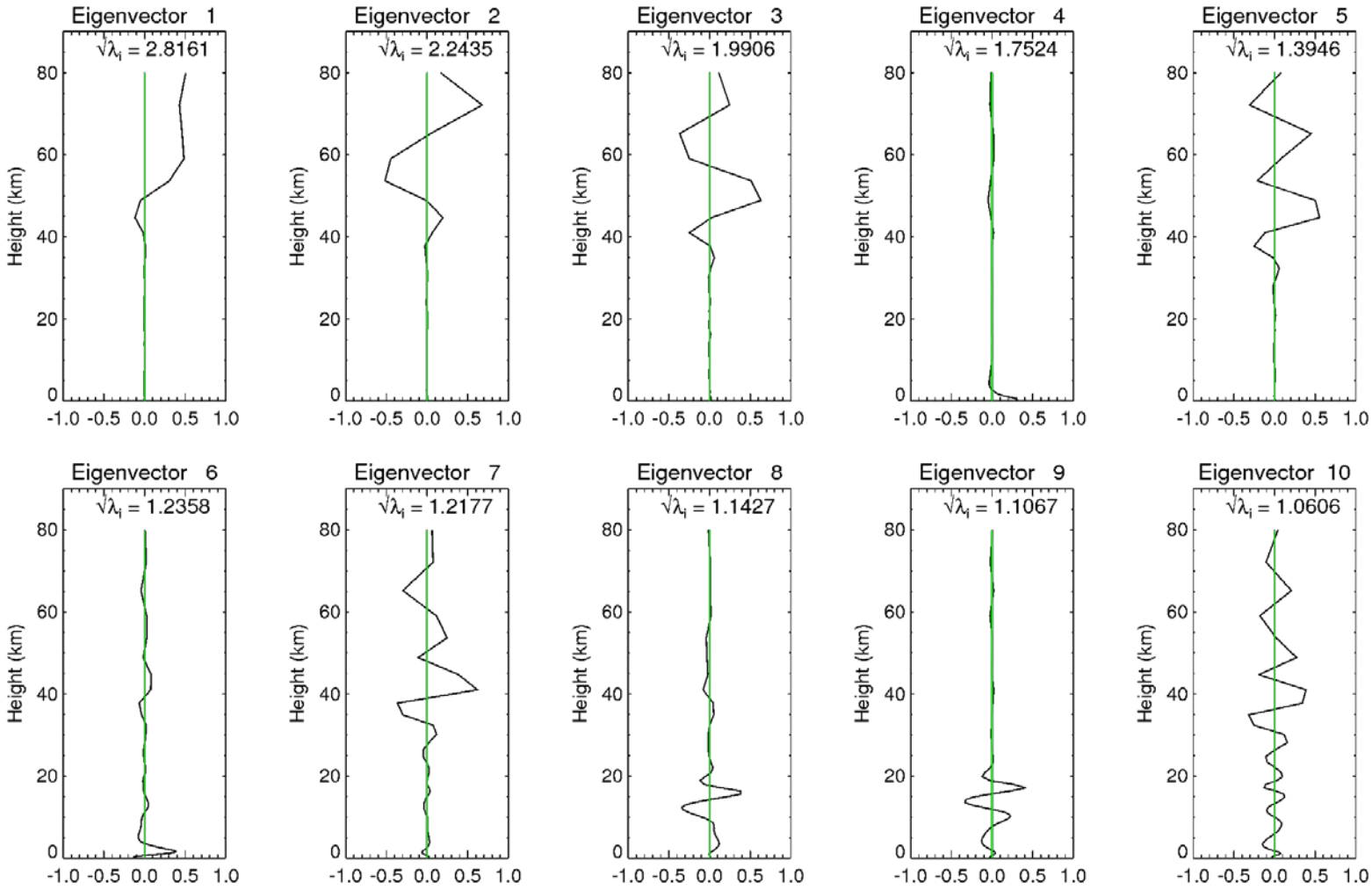
optimum weight or Kalman Gain

observation increment

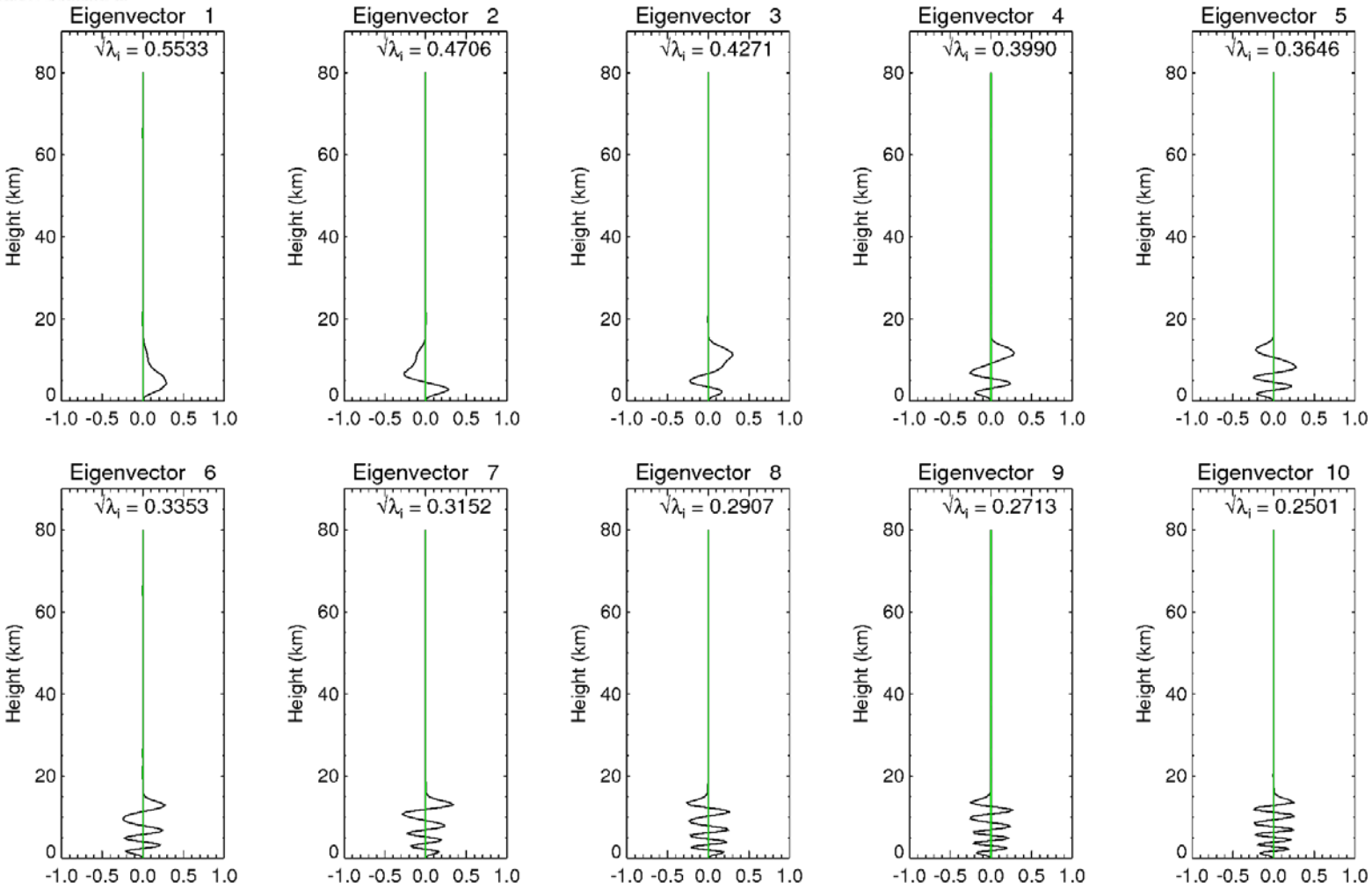
Eigenvectors of **B**

- Are one way of looking at vertical scales in the analysis
- The analysis is more sensitive to eigenvectors with large eigenvalues
- Modes with small eigenvalues are assumed to be “well-known” in the background, or to have small errors in reality (null-space)
  - No ability to influence these scales
  - Tends to be the case that very high vertical resolution modes have small eigenvalues
- They have no true physical meaning, but are broadly consistent with the fact that 4D-Var splits B into vertical modes
- IASI can only influence first 30-35 vectors for T and q

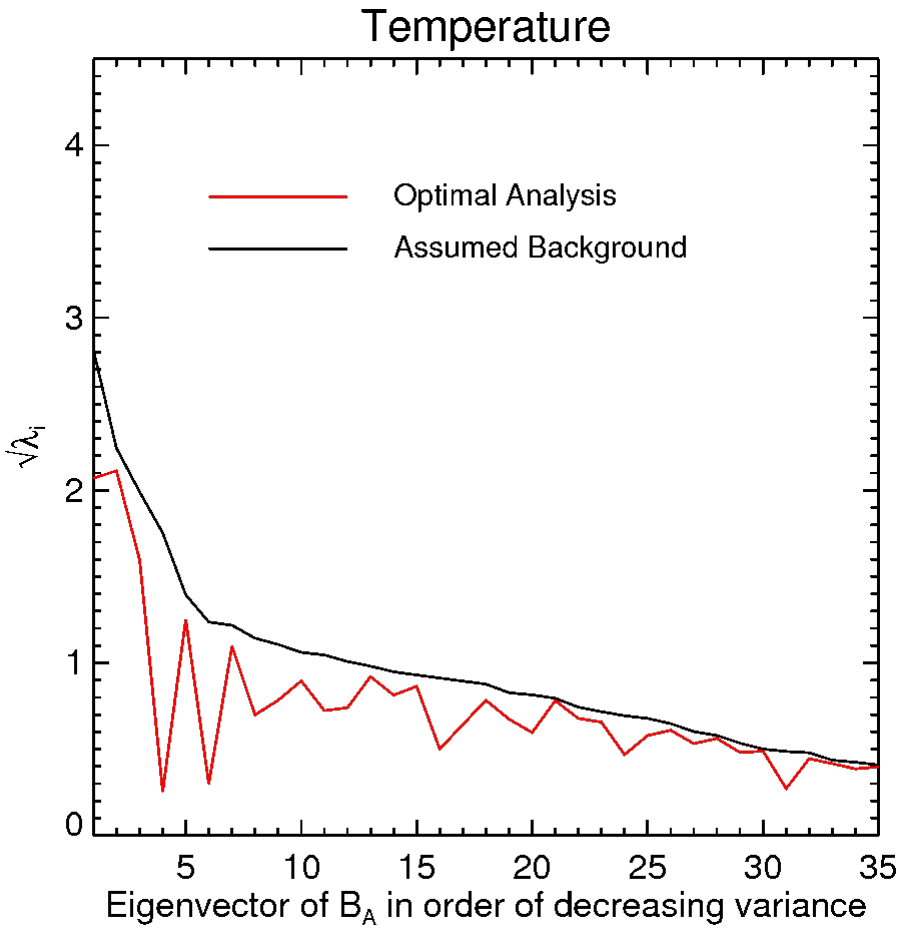
Leading eigenvectors of MoistCov B_A Temperature (vectors 1-10)

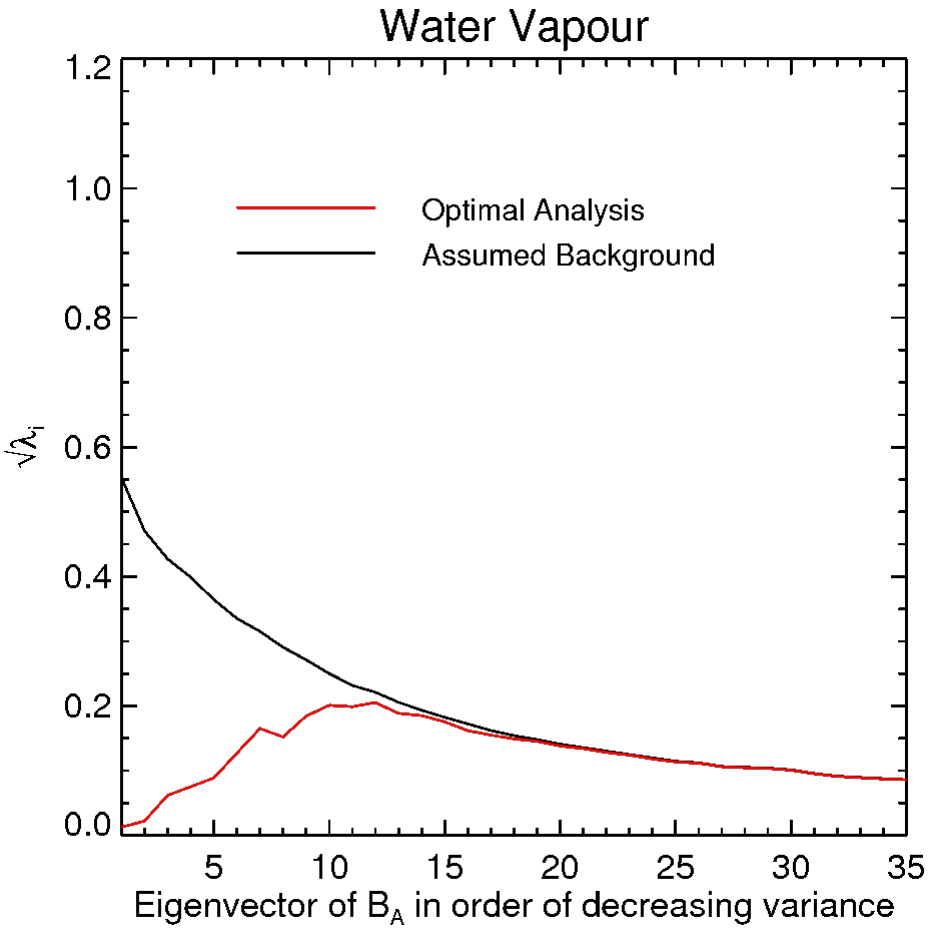


Leading eigenvectors of Moist Cov B_A ln(q) (vectors 1-10)

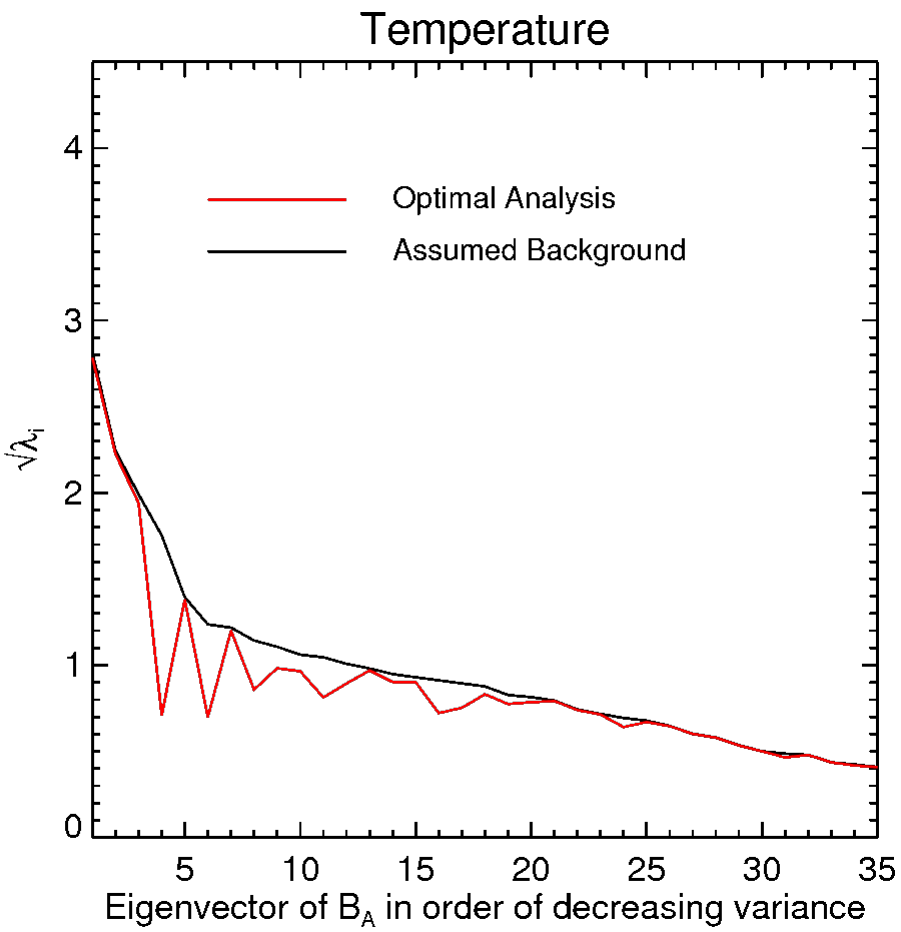


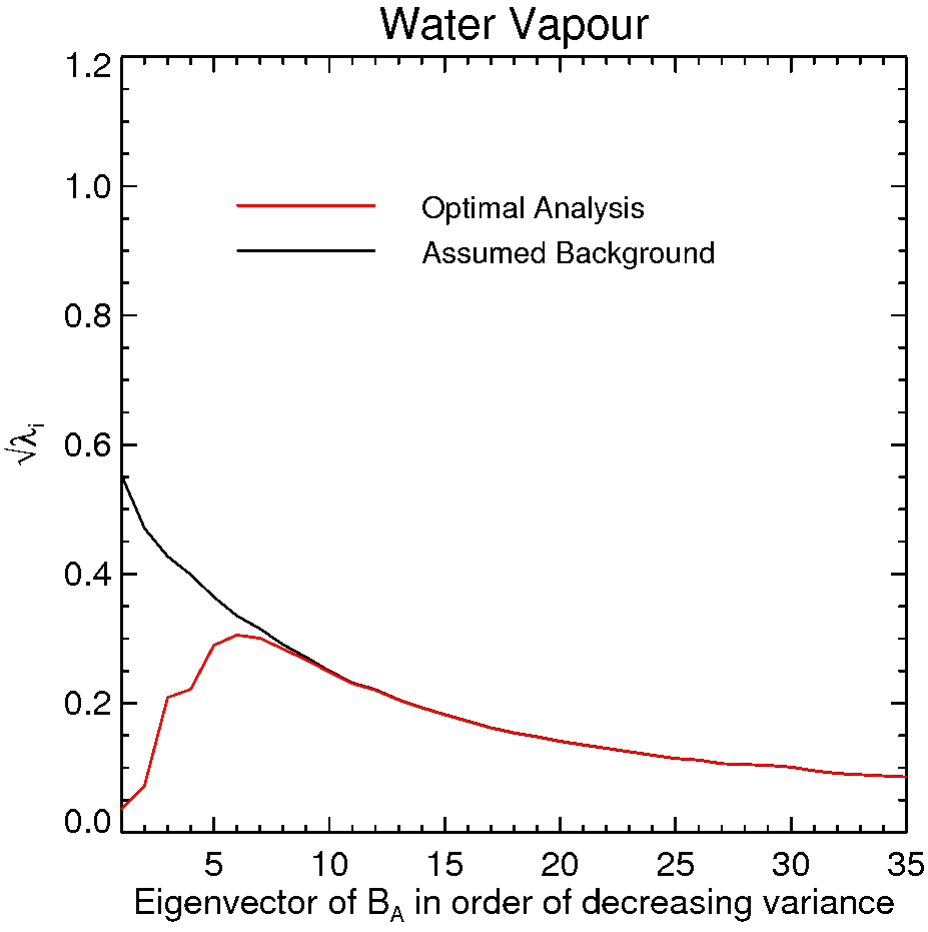
Optimal Analysis Full Spectrum R=instrument noise



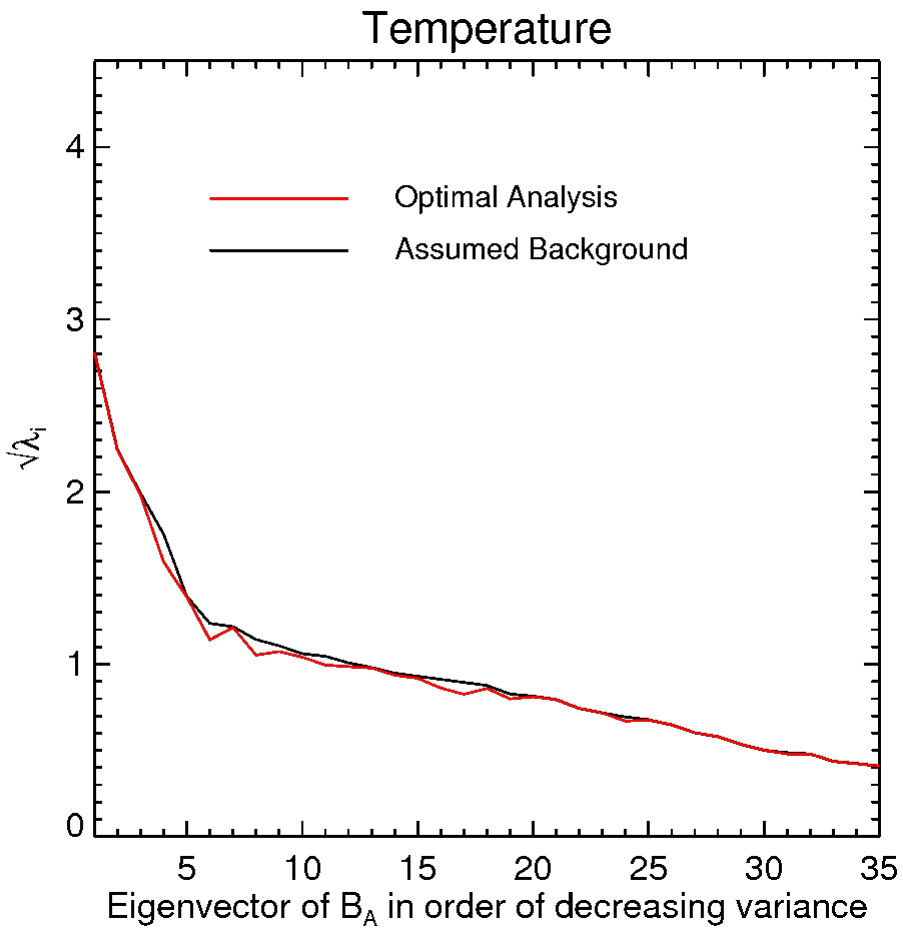


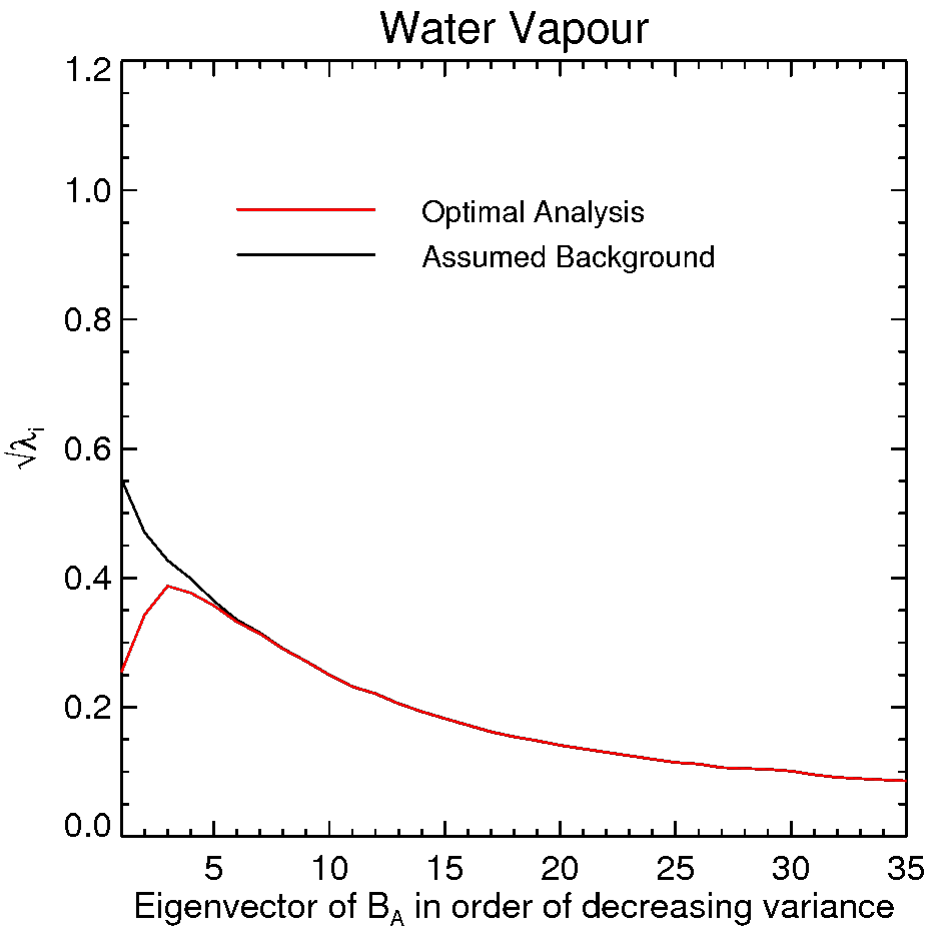
Optimal Analysis 4D-VAR chans
R = instrument noise





Optimal Analysis 4D-VAR chans
R = VAR Correlated Error Matrix



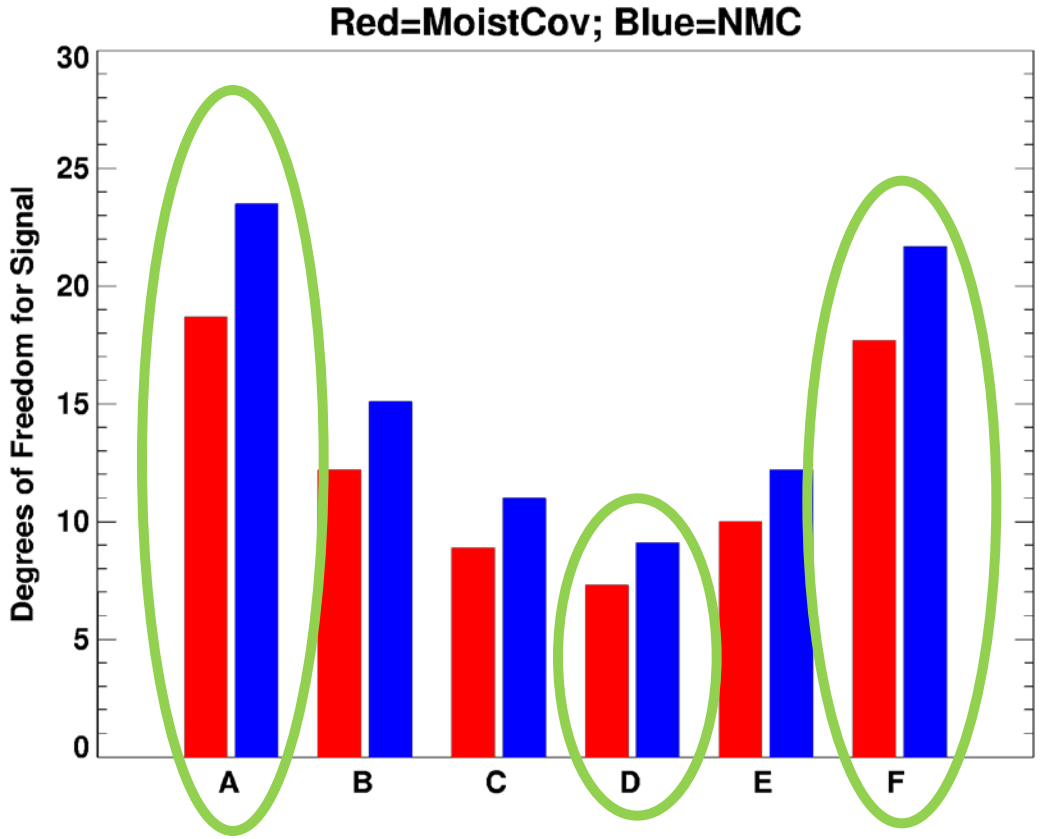


Filtering of IASI information

- This shows that IASI is influencing humidity information out to only about 4 vectors,
  - *Wavelengths more like 7 km than 1 km!*
- For temperature, almost nothing!
- This could indicate over-weighting of background on small scales
- Definite under-use of the spectral information also: under ideal circumstances we could
  - Double the DFS information content
  - Influence humidity out to 12 vectors (2.5 km wavelength)
  - Influence temperature out to about 25 vectors (4 km wavelength)

DFS for optimal analysis
Varying the channel selection

- R = Instrument noise
- A=Full Spec B=314 Chans C=OPS D=VAR
  E=Band 1 F=PCS



Principal Components

Principal Component Compression (1)

- IASI has 8461 channels but not 8461 pieces of information
- PCA is a linear transform from channel space so that the information is represented by orthogonal components.
- Usually rank these components in such a way that the first PC explains most of the variance in the dataset, the second contains the next most etc
  - You can throw away most of the components and still retain nearly all the atmospheric signal
  - For IASI it seems that ~ 300 PCs contain the atmospheric signal
- What you throw away is mostly random noise

Principal Component Compression (2)

- Usually noise-normalise the spectra first, though other norms are used.

$$y_{pc} = \mathbf{L}^{T}_{npc}\, \mathbf{R}^{-1}\, y_{chan}$$

- y_{chan} is the observation in channel space
- y_{pc} is the observation in PC space
- npc is the numper of retained PCs (~300)
- **L** is the transformation matrix (size nchan × npc)
- **R** is the noise covariance matrix (or other norm)

Principal Component Compression (3)

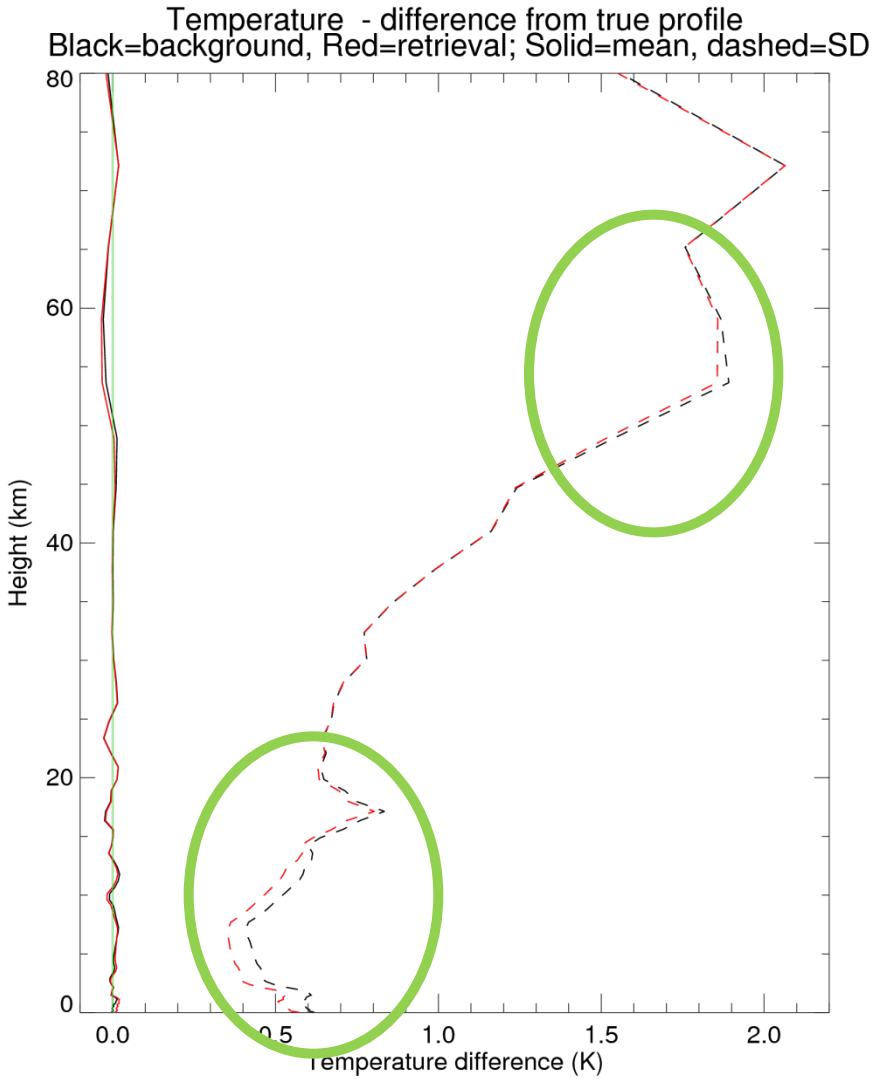
- PCs can be used to compress spectra for dissemination purposes
  - In that case the training data should be real spectra
  - That's not what I am discussing today!
- PCs can also be used for radiative transfer modelling and assimilation
  - The training data may be spectra simulated by a LBL model
  - This *is* what I am discussing today!
- The results presented use PC-RTTOV (Matricardi) to simulate IASI spectra

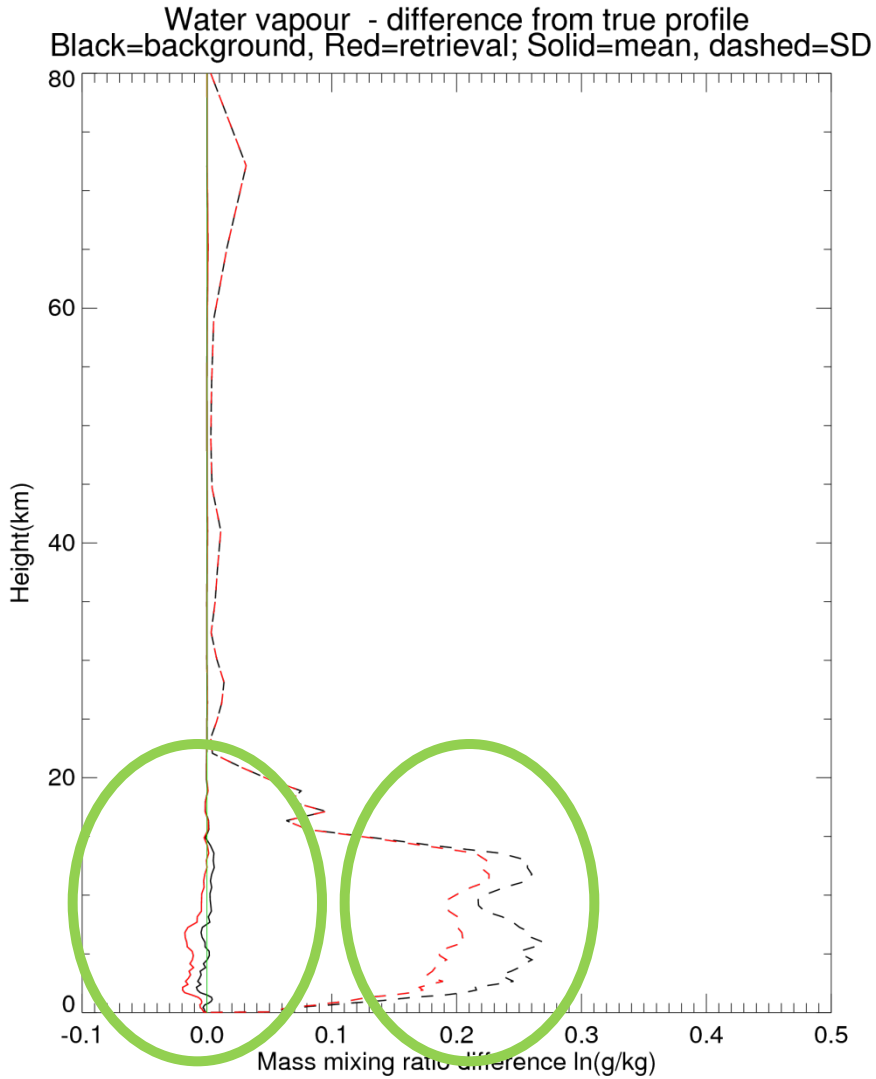
1D-Var simulated assimilation experiments

1D-Var Set-Up

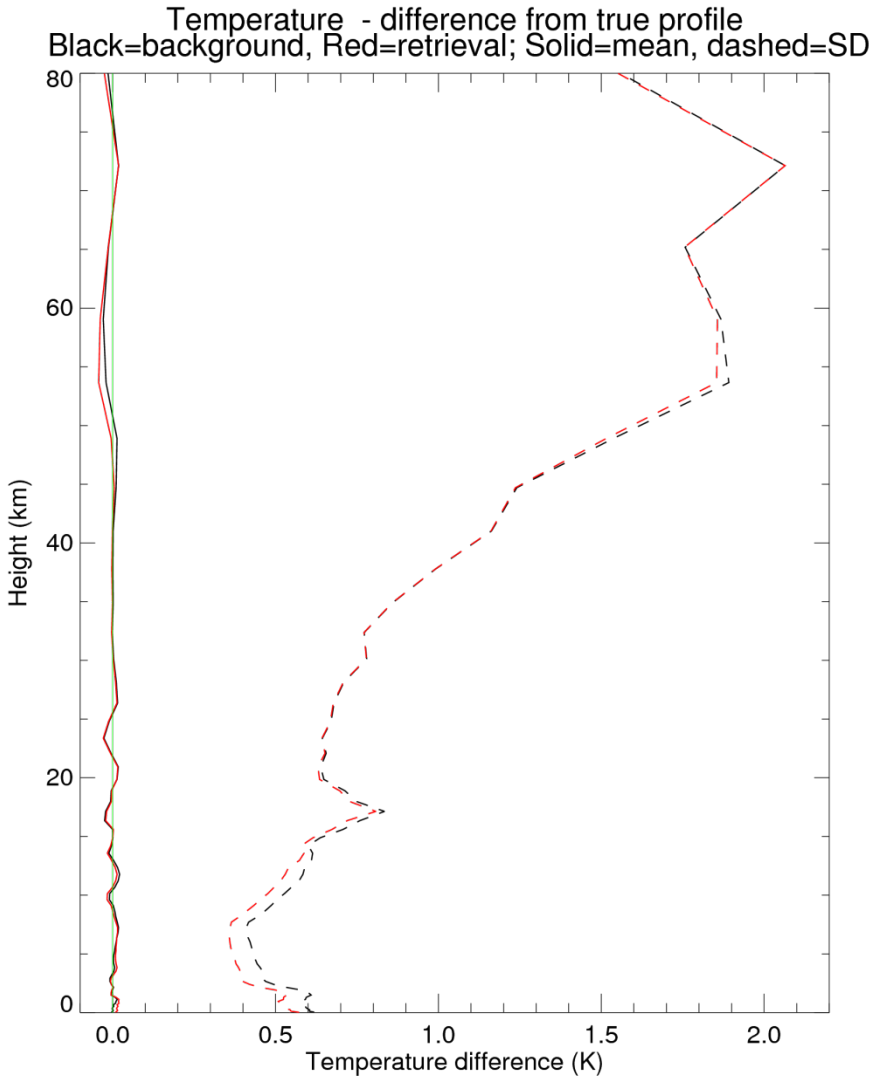
- 4400 Profiles from Met Office model used as truth
- Simulated background errors are added from the MoistCov 1D B-matrix consistent with Met Office 4D-Var
- Observations are simulated from these profiles using RTTOV-10, and observation errors are added consistent with the (diagonal) instrument noise used above
- The observations are then transformed into PC-RTTOV PC scores
- Experiments repeated adding forward model error

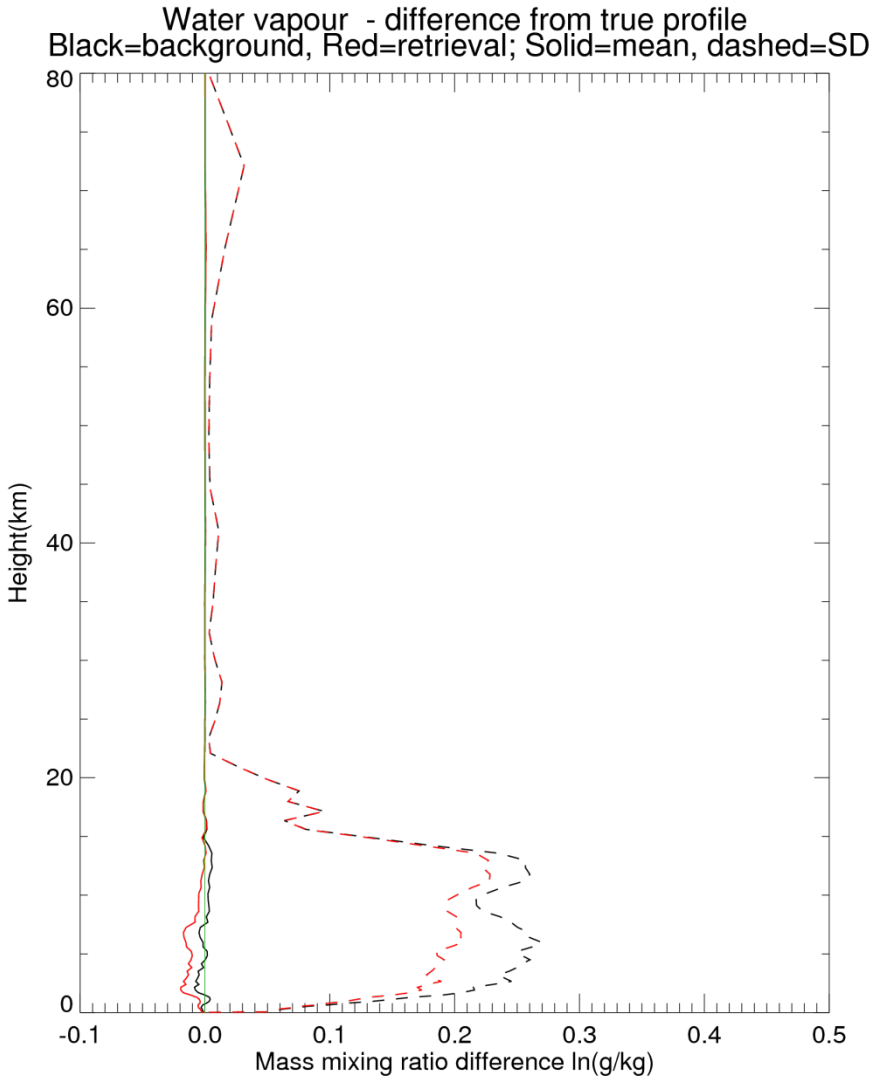
4D-Var Channels – Inst Noise



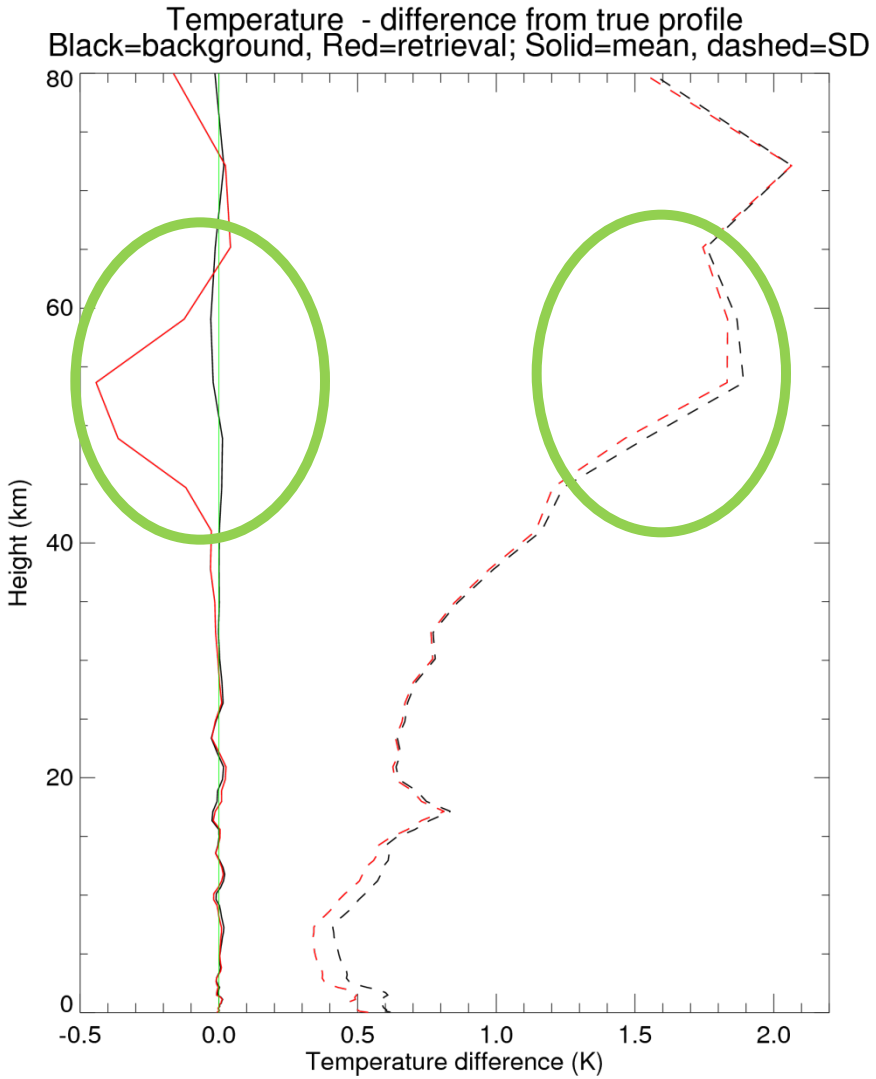


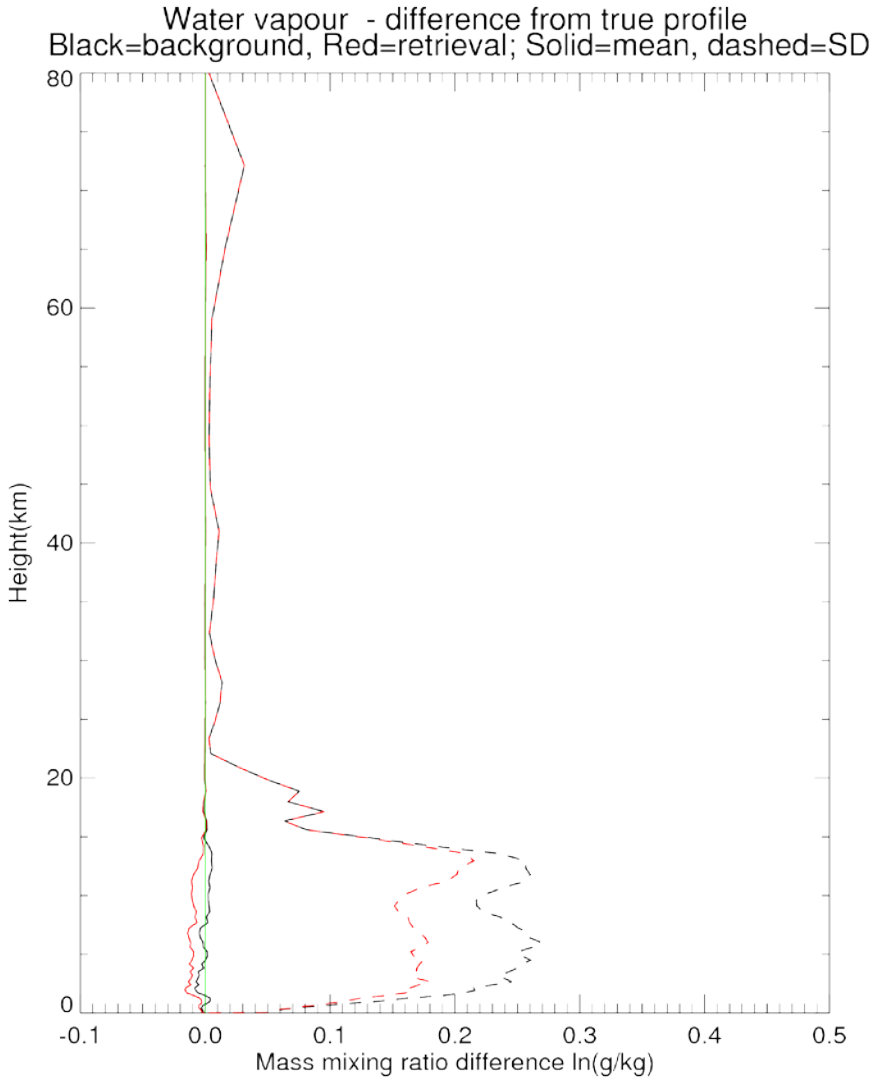
4D-Var Channels – Inst Noise + RT Error



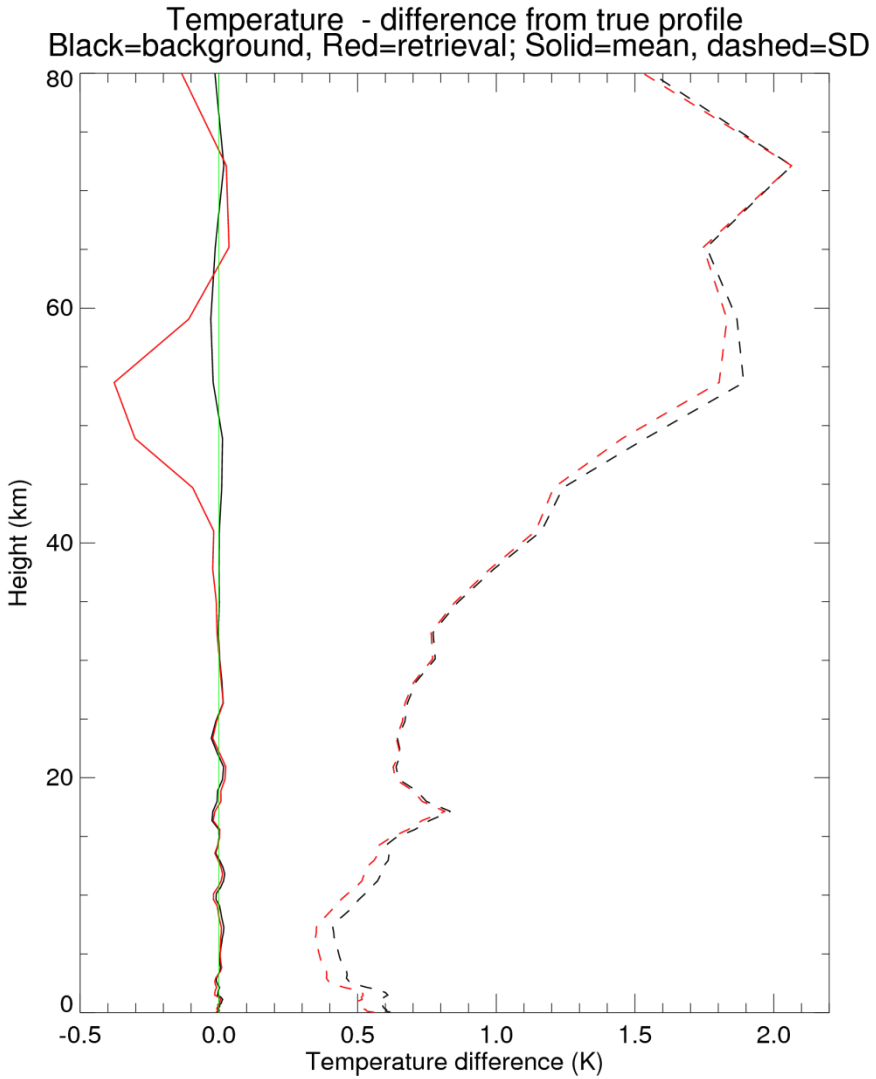


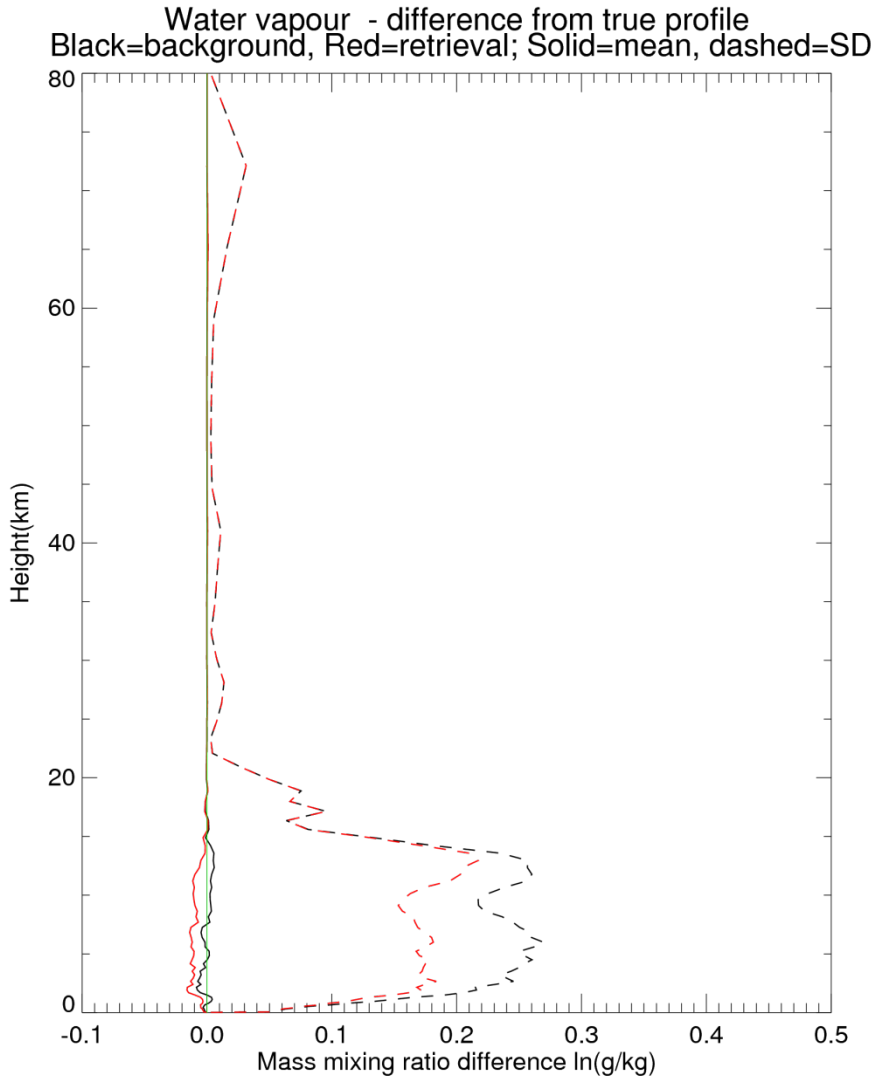
PCRTTOV Predictor Channels – Inst Noise



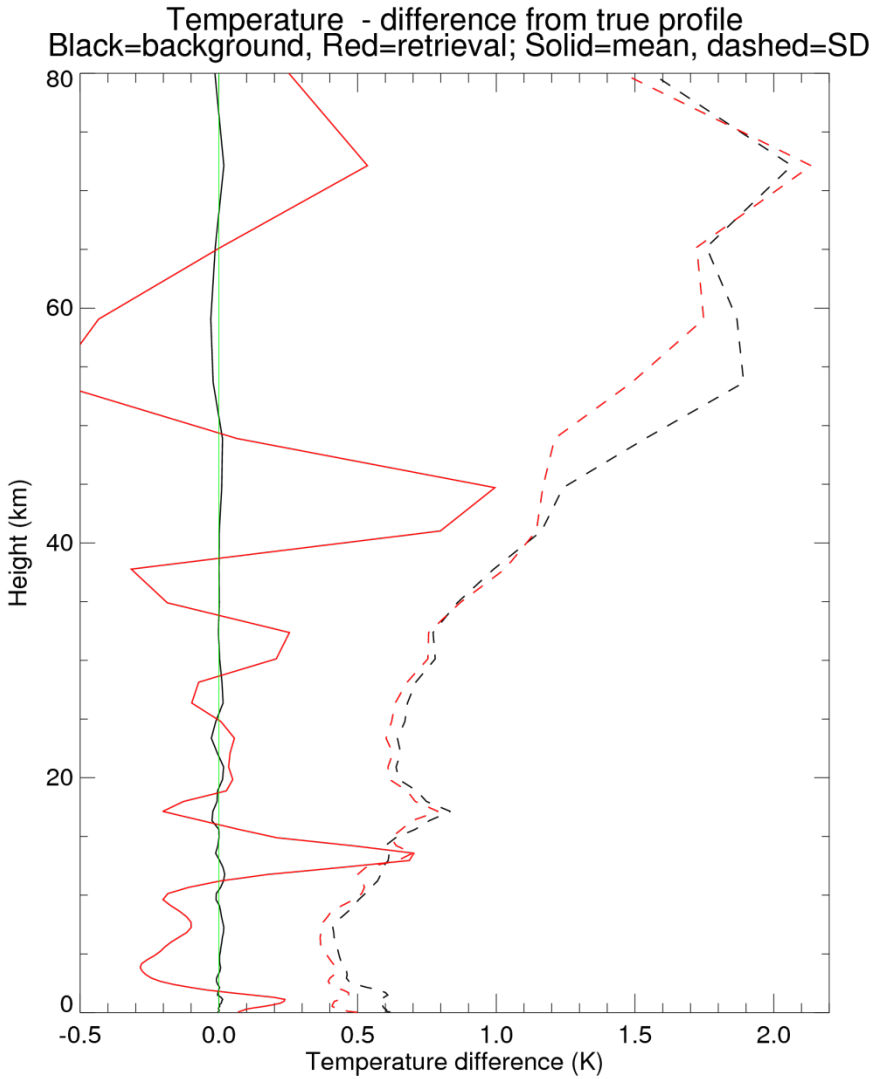


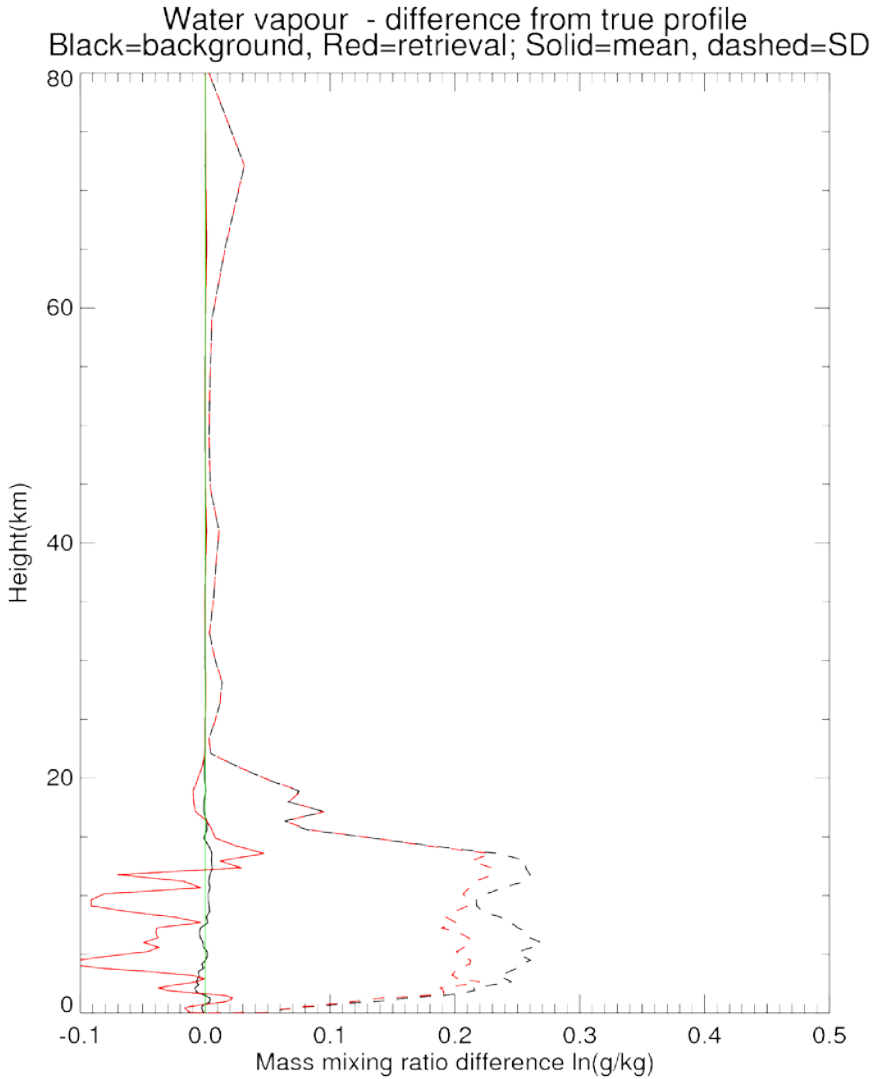
PCRTTOV Predictor Channels – Inst Noise + RT Error



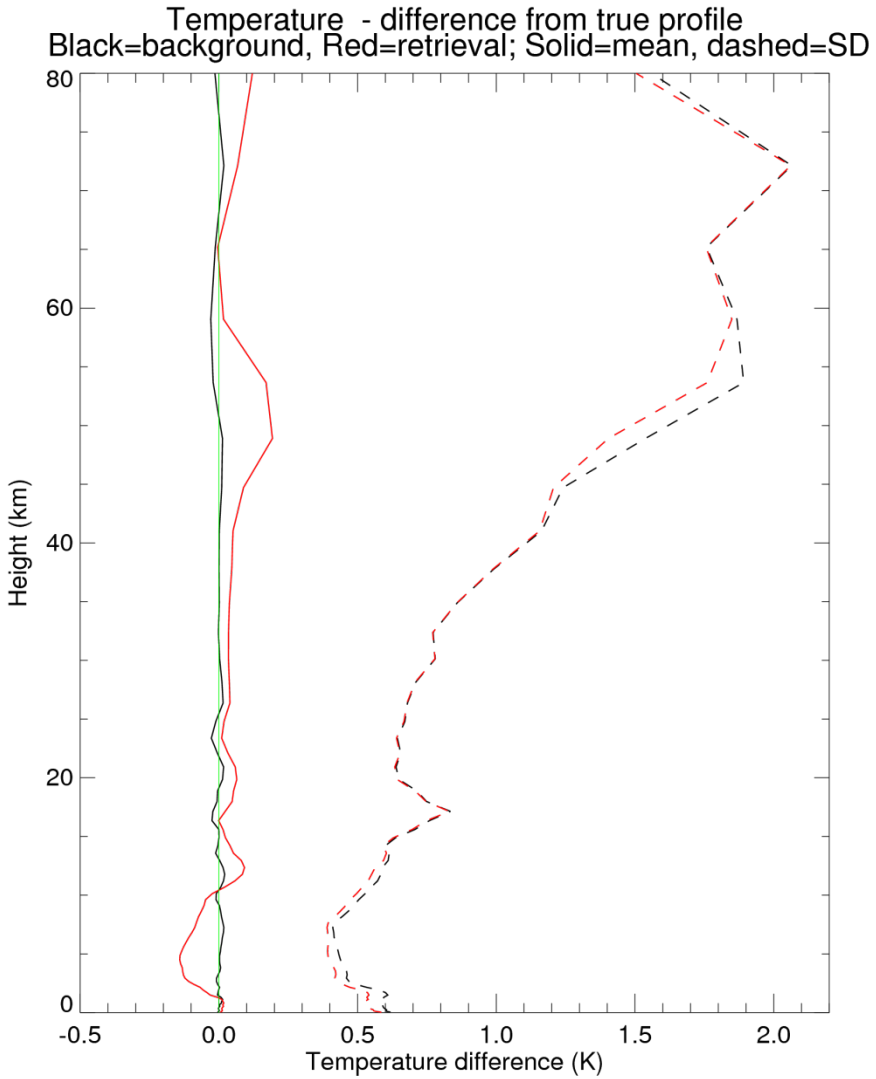


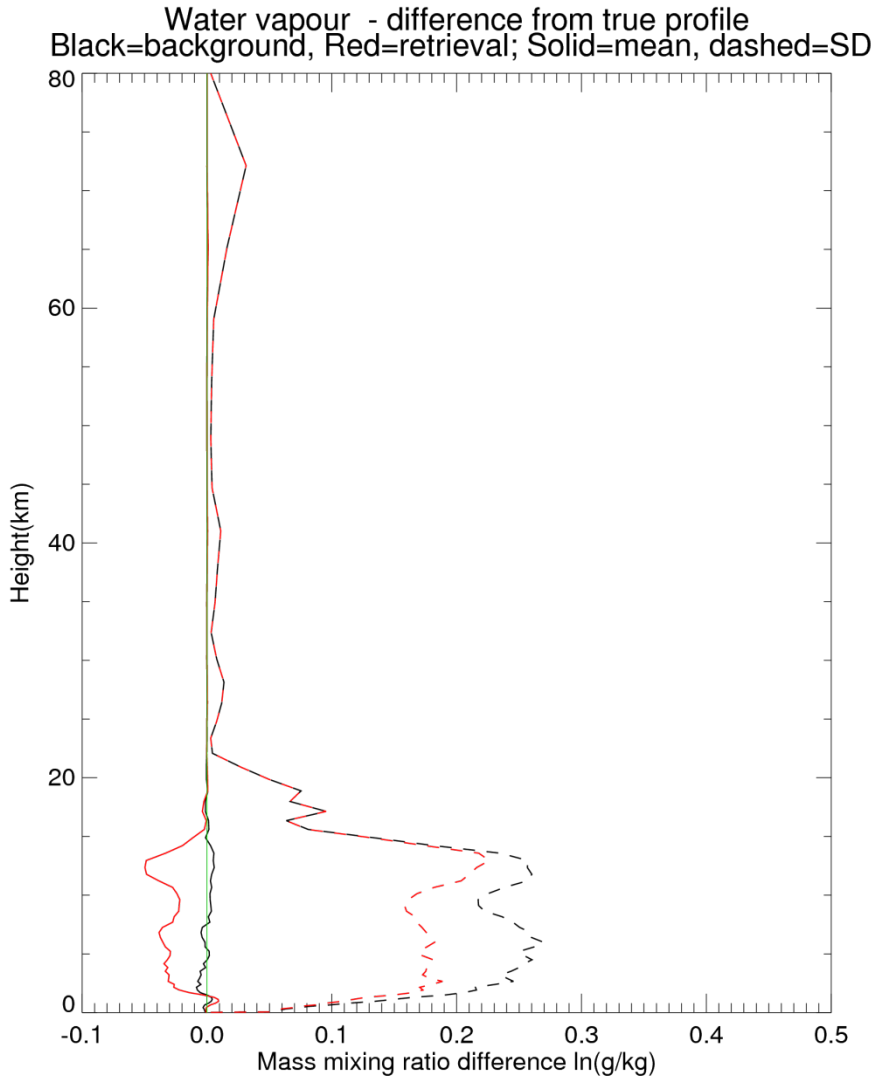
PCRTTOV PC Scores – Inst Noise





PCRTTOV PC Scores – Inst Noise + RT Error Diagonalised





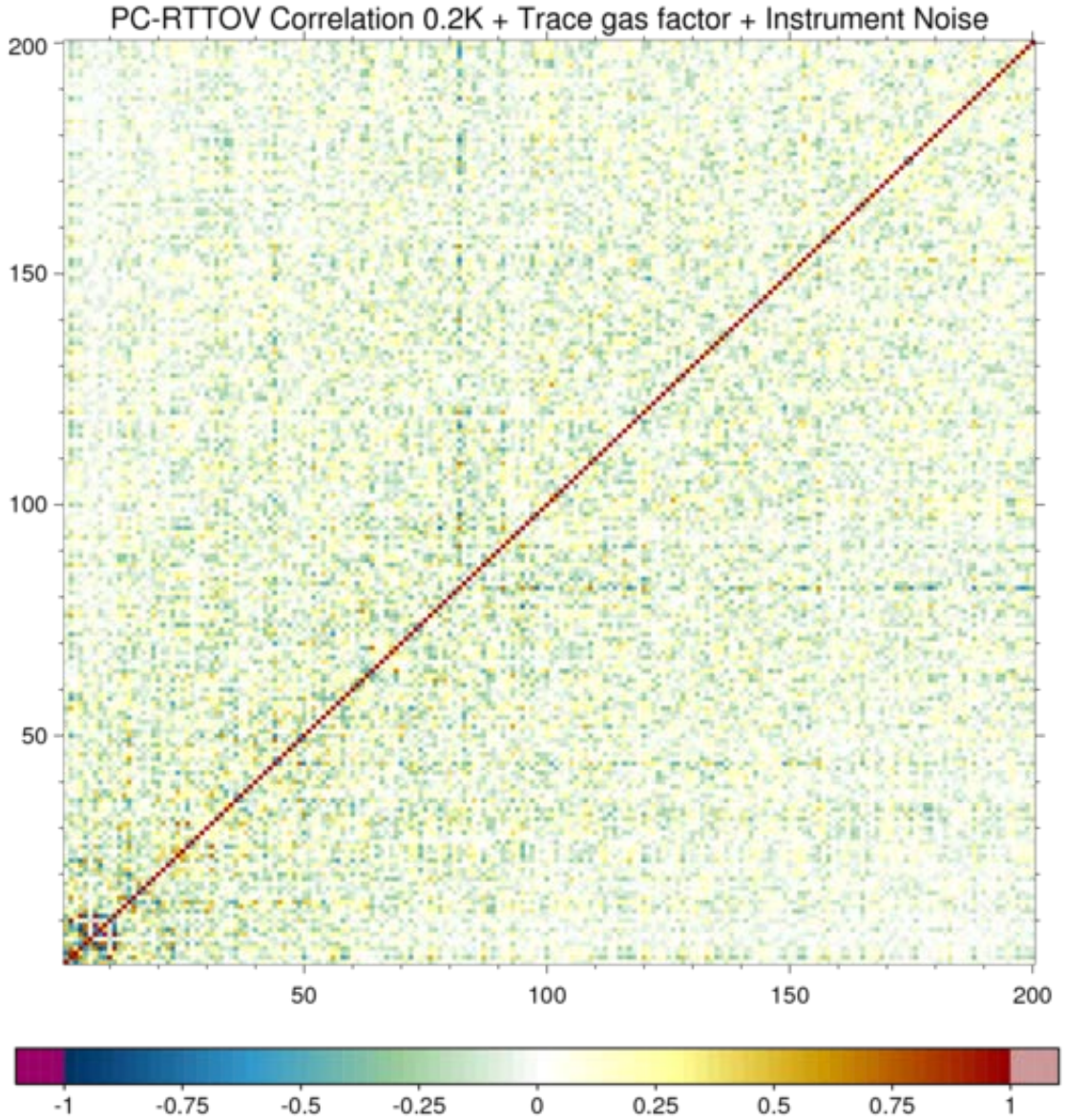
Zigzagging bias for PCs

- Could this be due to the reduction in dimensionality of the observation?
- R matrix in the case of instrument noise:

$$L_{npc}\, L^{T}_{nchan}\, R^{-1}\, R\, L_{nchan}\, L^{T}_{npc} = I_{npc}$$

- You can see that I_{npc} will have reduced dimensionality when npc < nchan
- When you add instrument noise the mathematical properties are different
  - Especially when you “delete” the correlations!

Instrument noise + RT error in PC-RTTOV PC Space



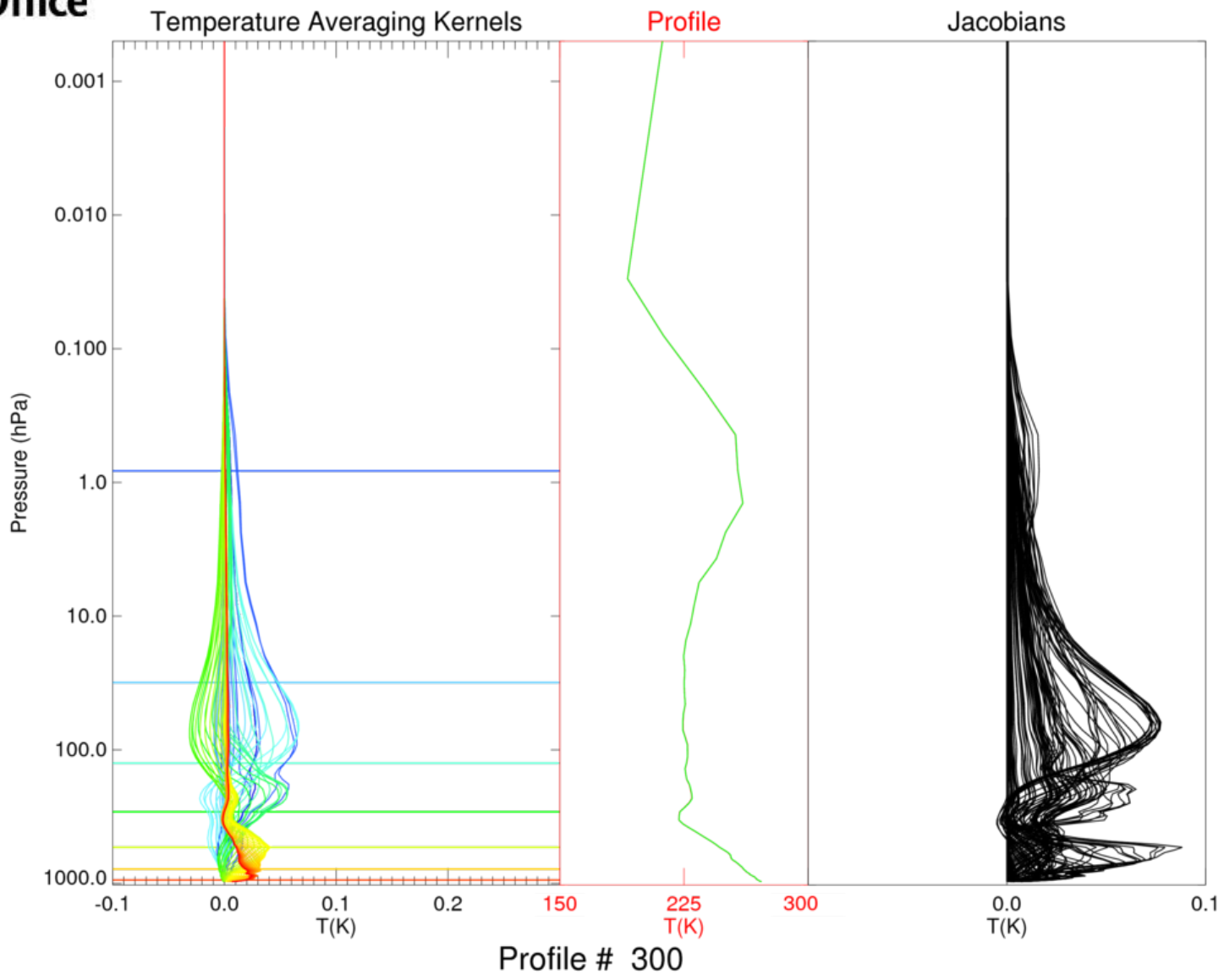
Summary of 1D-Var results

- Adding more channels improves Analysis SD fit to truth, especially for water vapour…
- …but introduces bias in stratospheric temperature
- Analysis in PC space produces unstable biases
- Adding RT error to the observations has little effect for channel retrieval results…
- … but regularises the error covariance matrix in PC space allowing a more stable retrieval performance

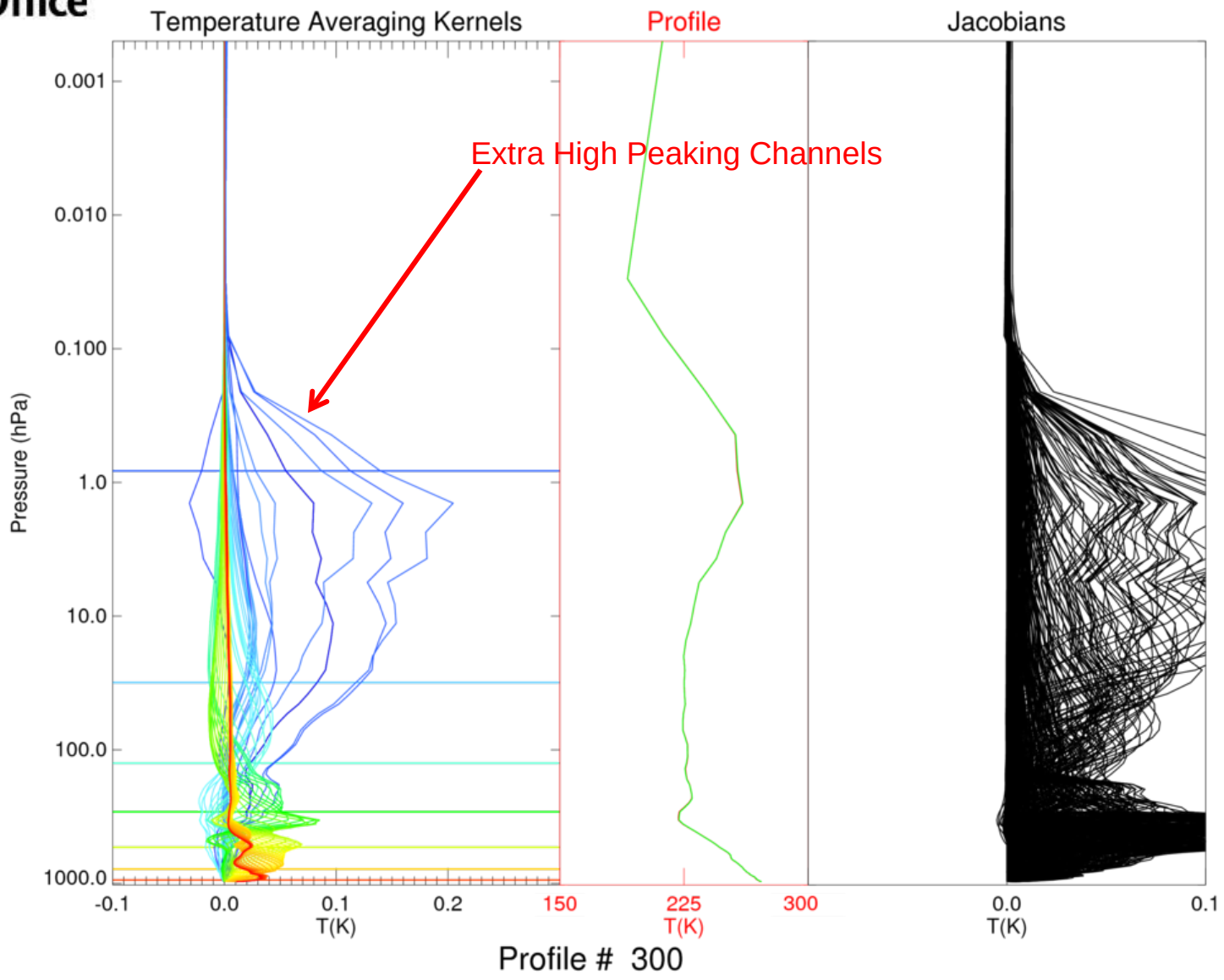
- This is not the full story

Minimisation:
How much structure can the analysis see?

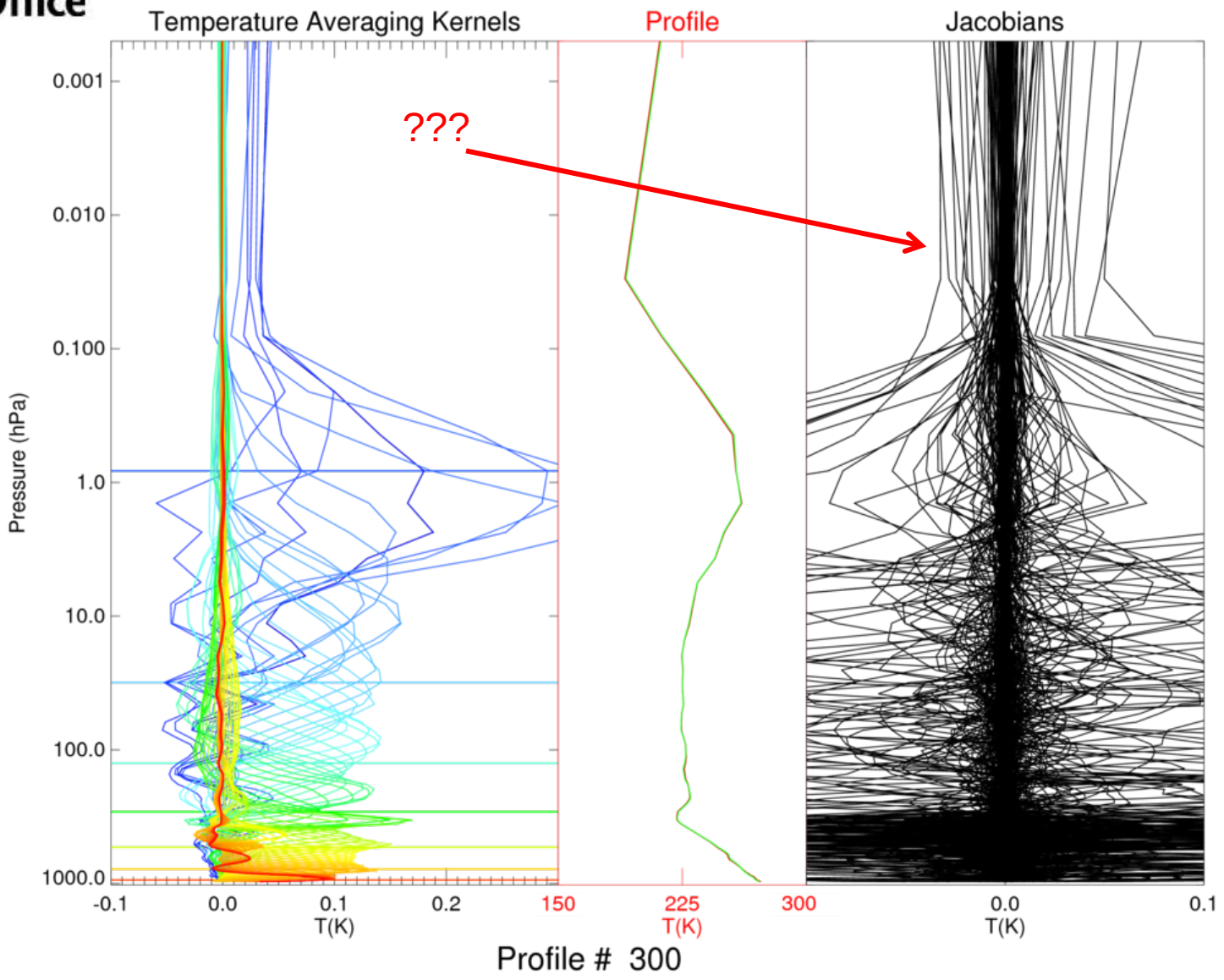
Jacobians and Averaging Kernels 4D-Var Channels - Temperature



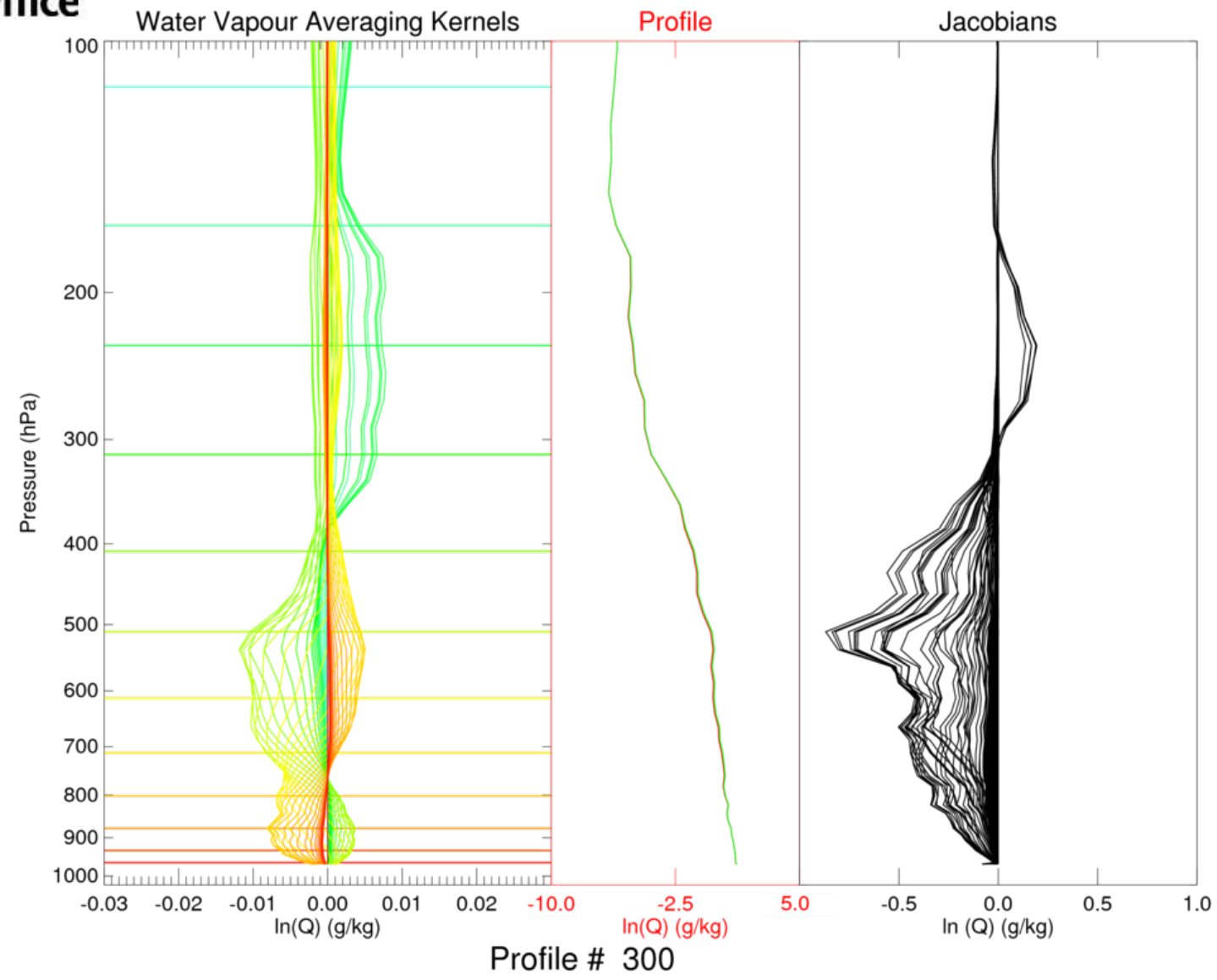
Jacobians and Averaging Kernels PC-RTTOV Channels - Temperature



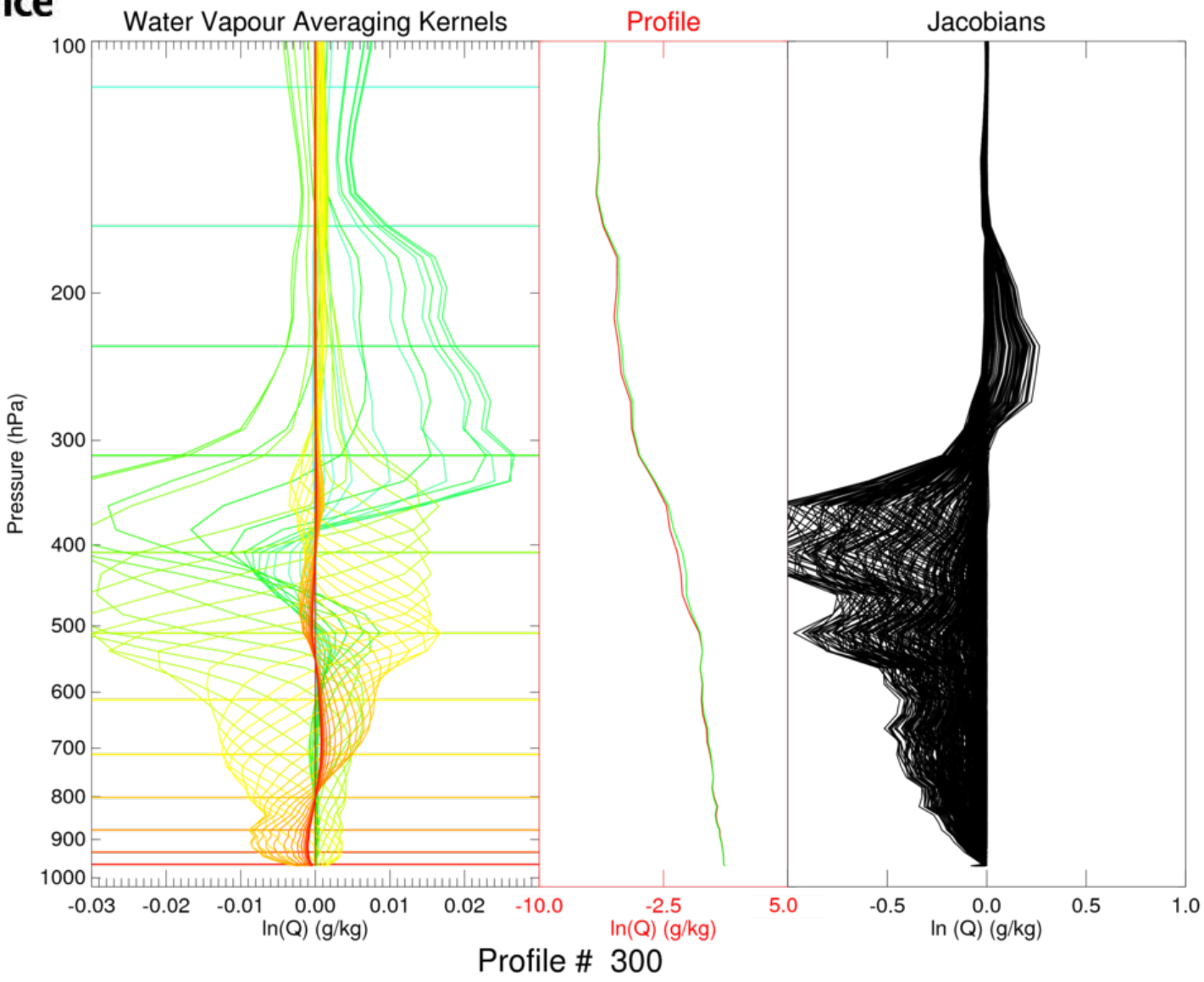
Jacobians and Averaging Kernels
PC-RTTOV PC Scores - Temperature



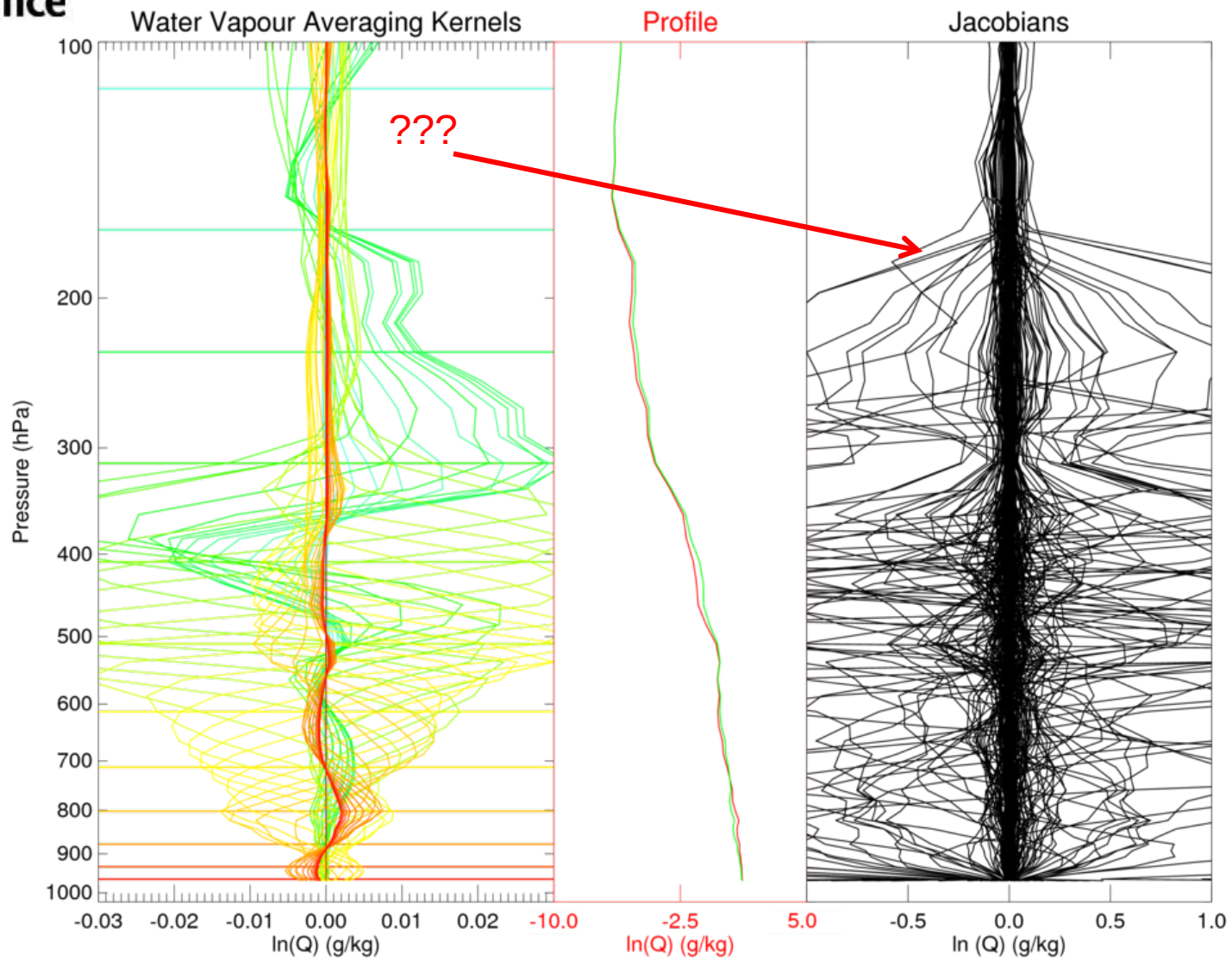
Jacobians and Averaging Kernels 4D-Var Channels – Water Vapour



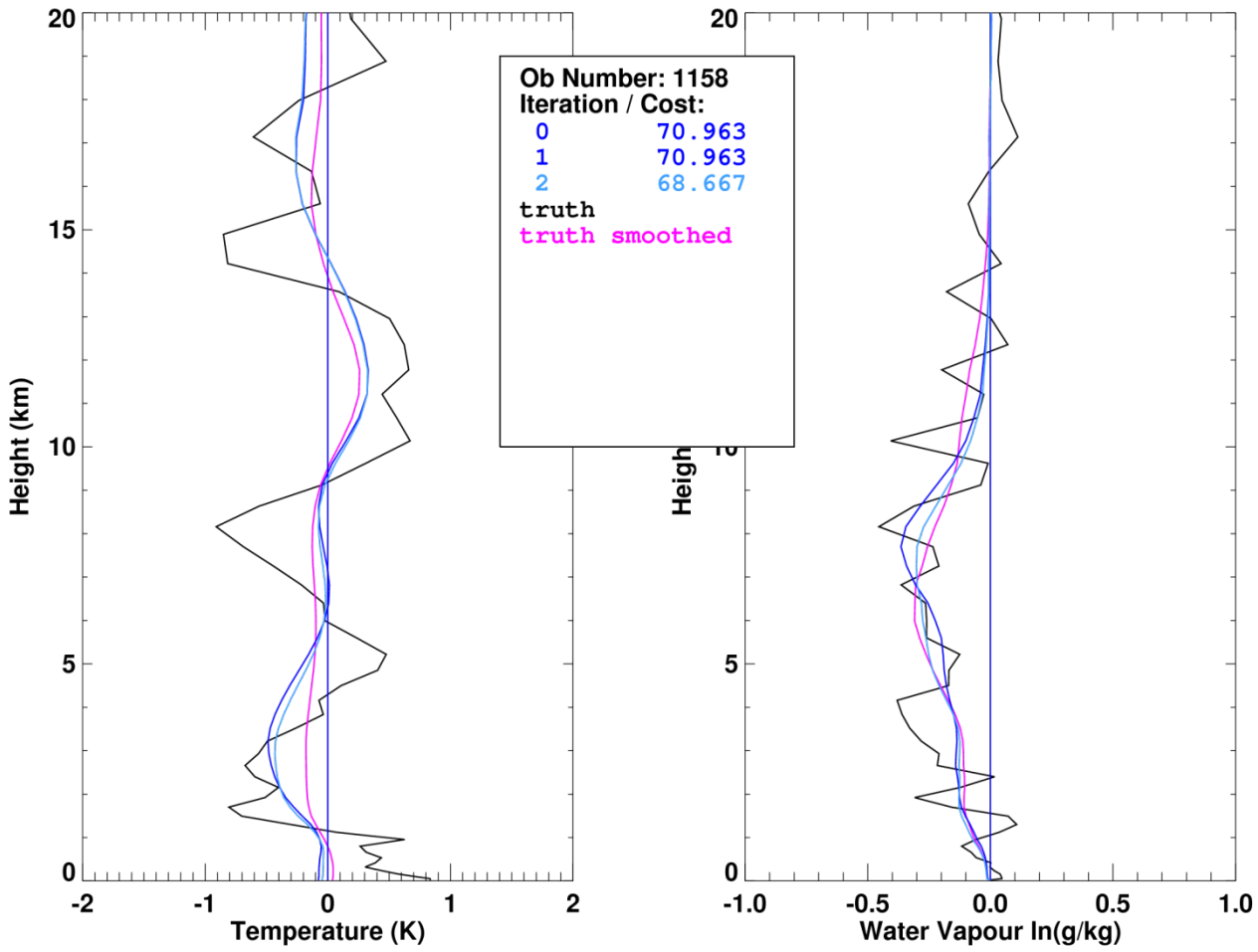
Jacobians and Averaging Kernels PC-RTTOV Channels – Water Vapour



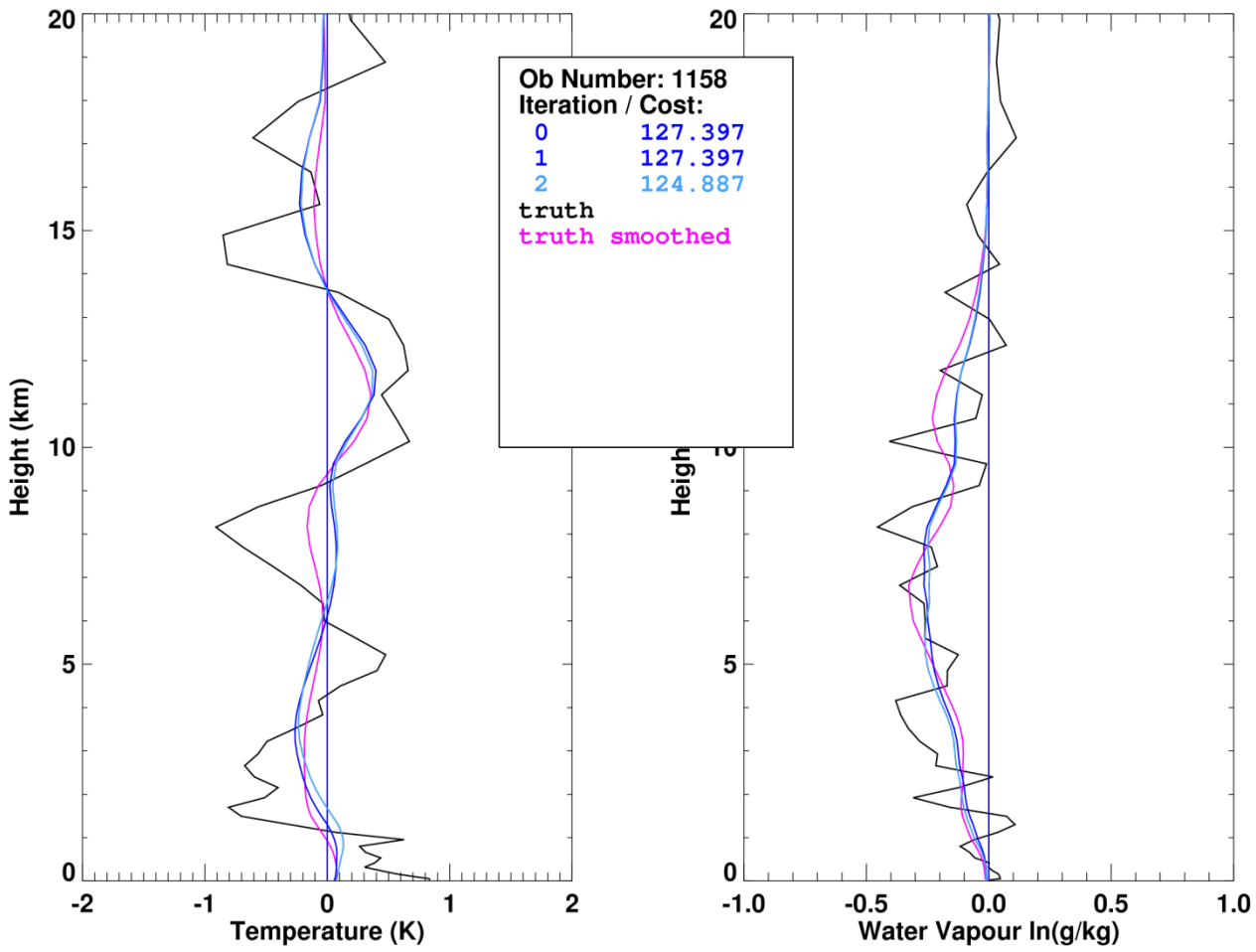
Jacobians and Averaging Kernels
PC-RTTOV PC Scores – Water Vapour



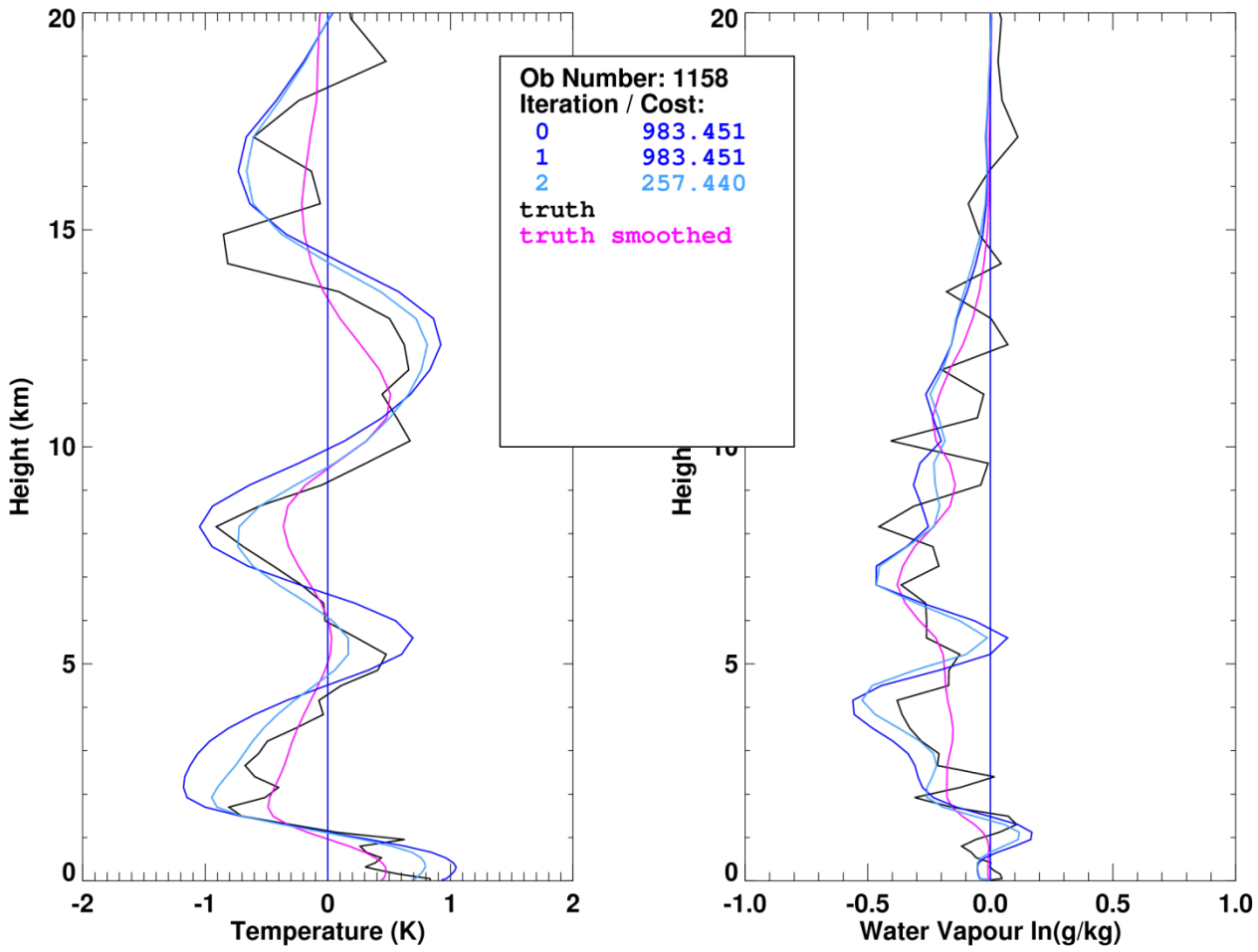
Minimisation – 4D-Var channels Instrument noise



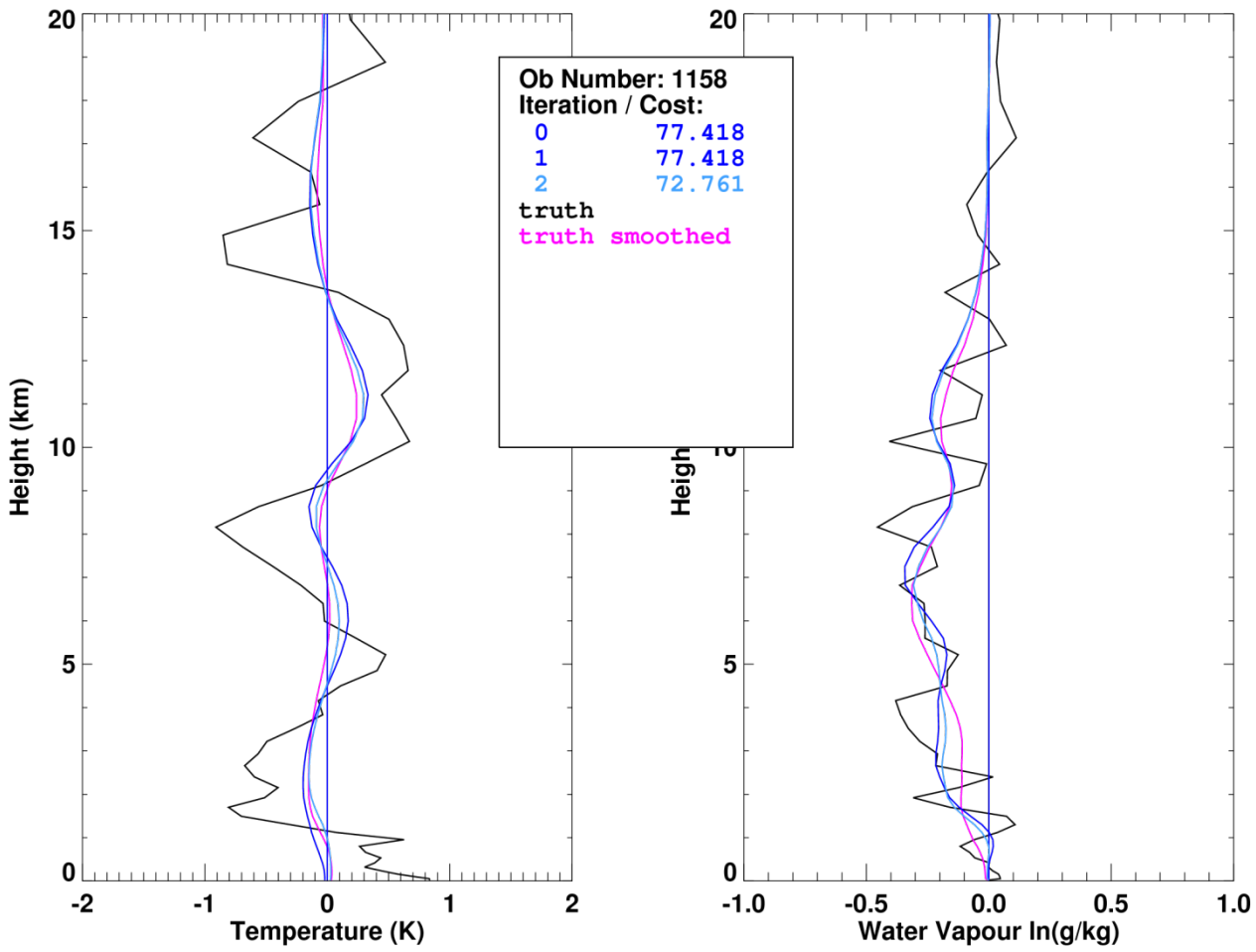
Minimisation – 500 PCRTTOV chans Instrument noise



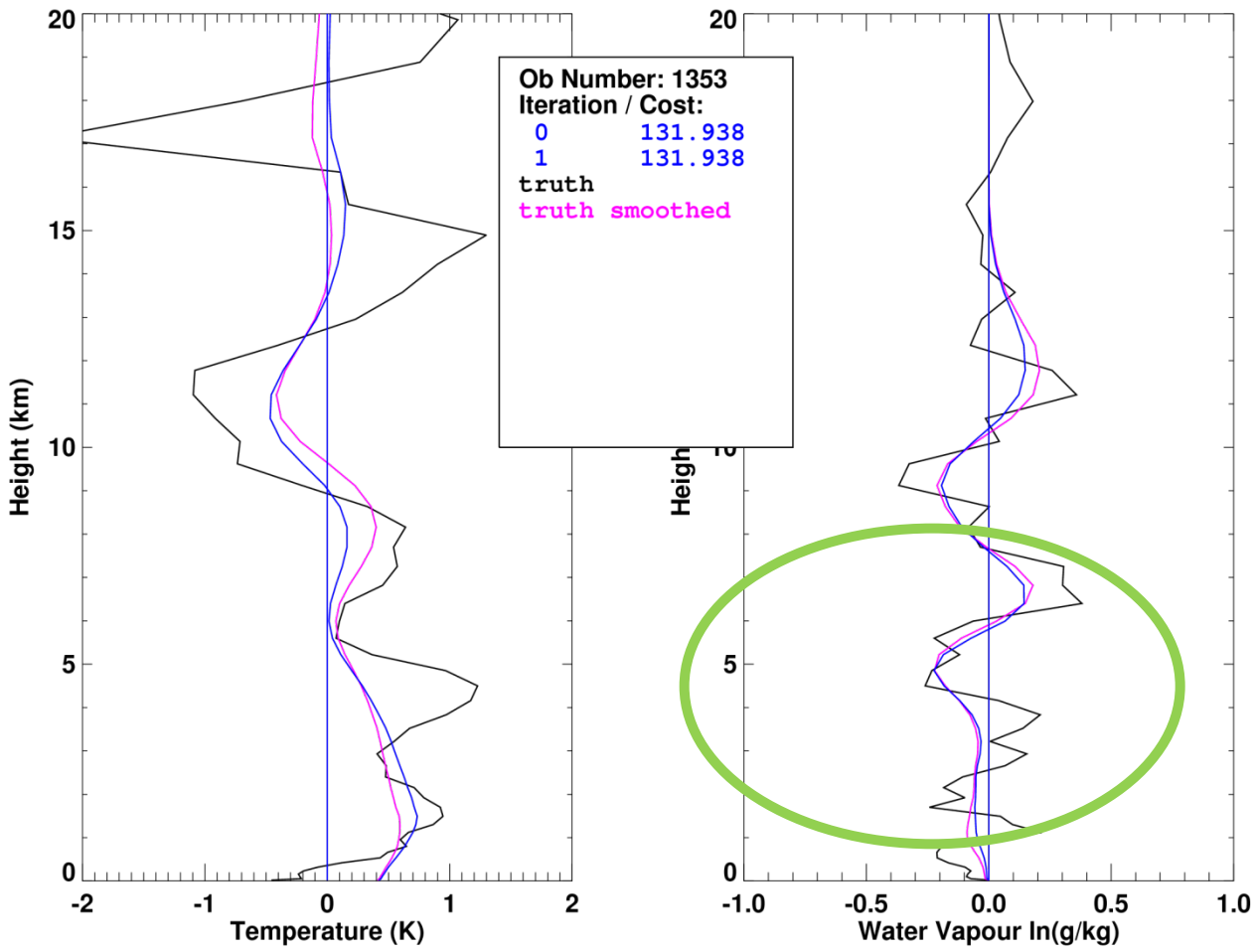
Minimisation – PC Scores Instrument noise



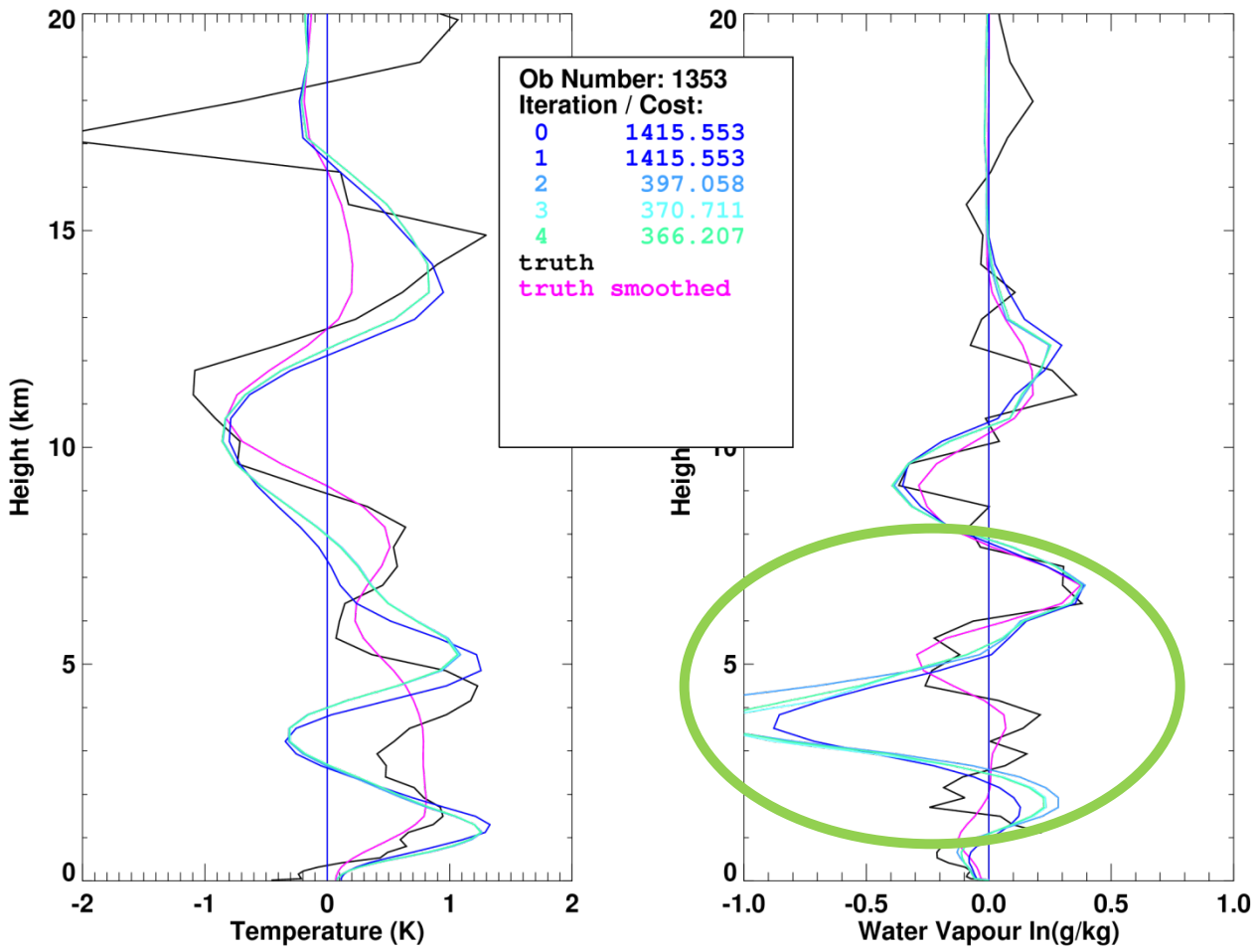
Minimisation – PC scores
Instrument noise + RT error



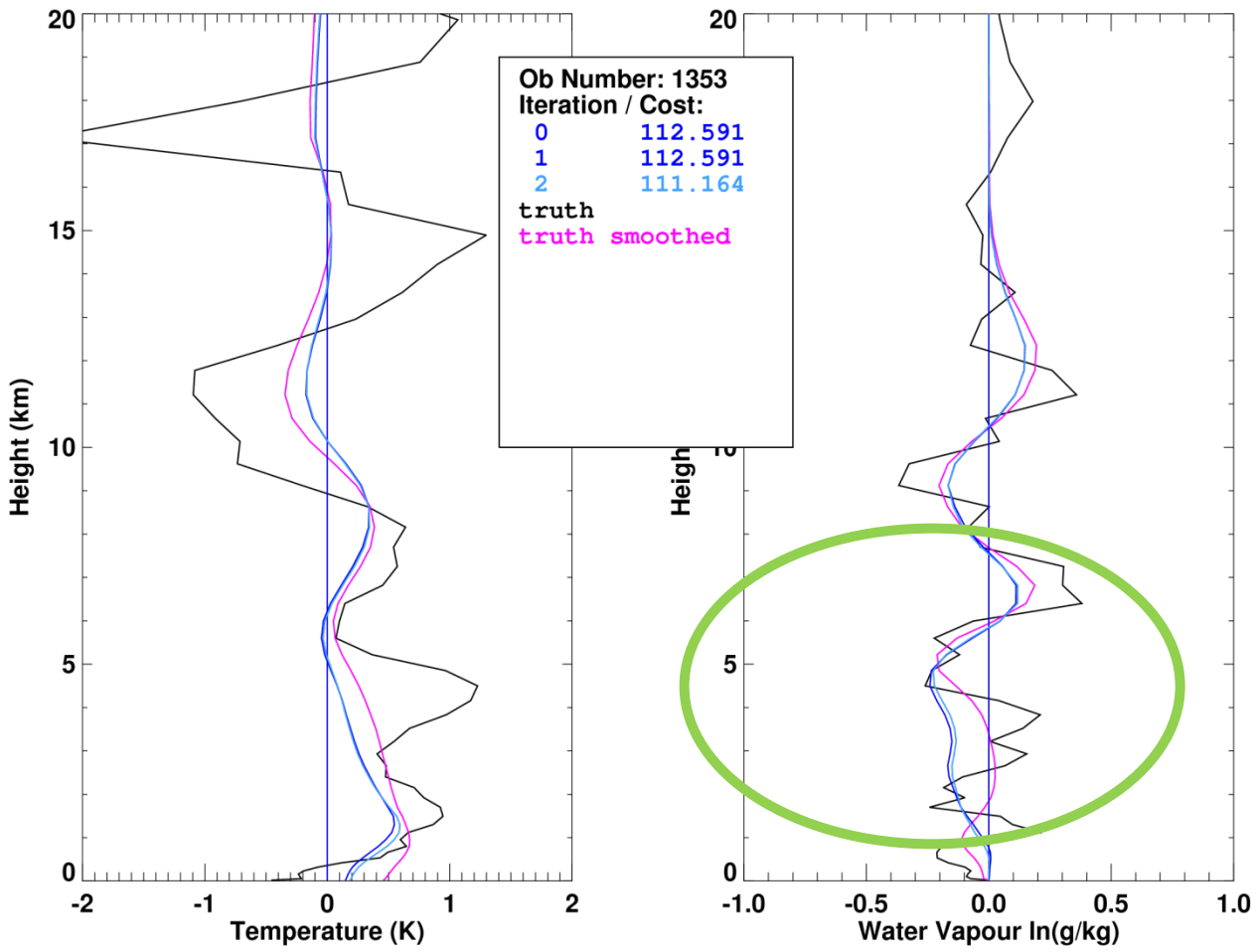
Minimisation – 500 PCRTTOV chans Instrument noise



Minimisation – PC Scores Instrument noise



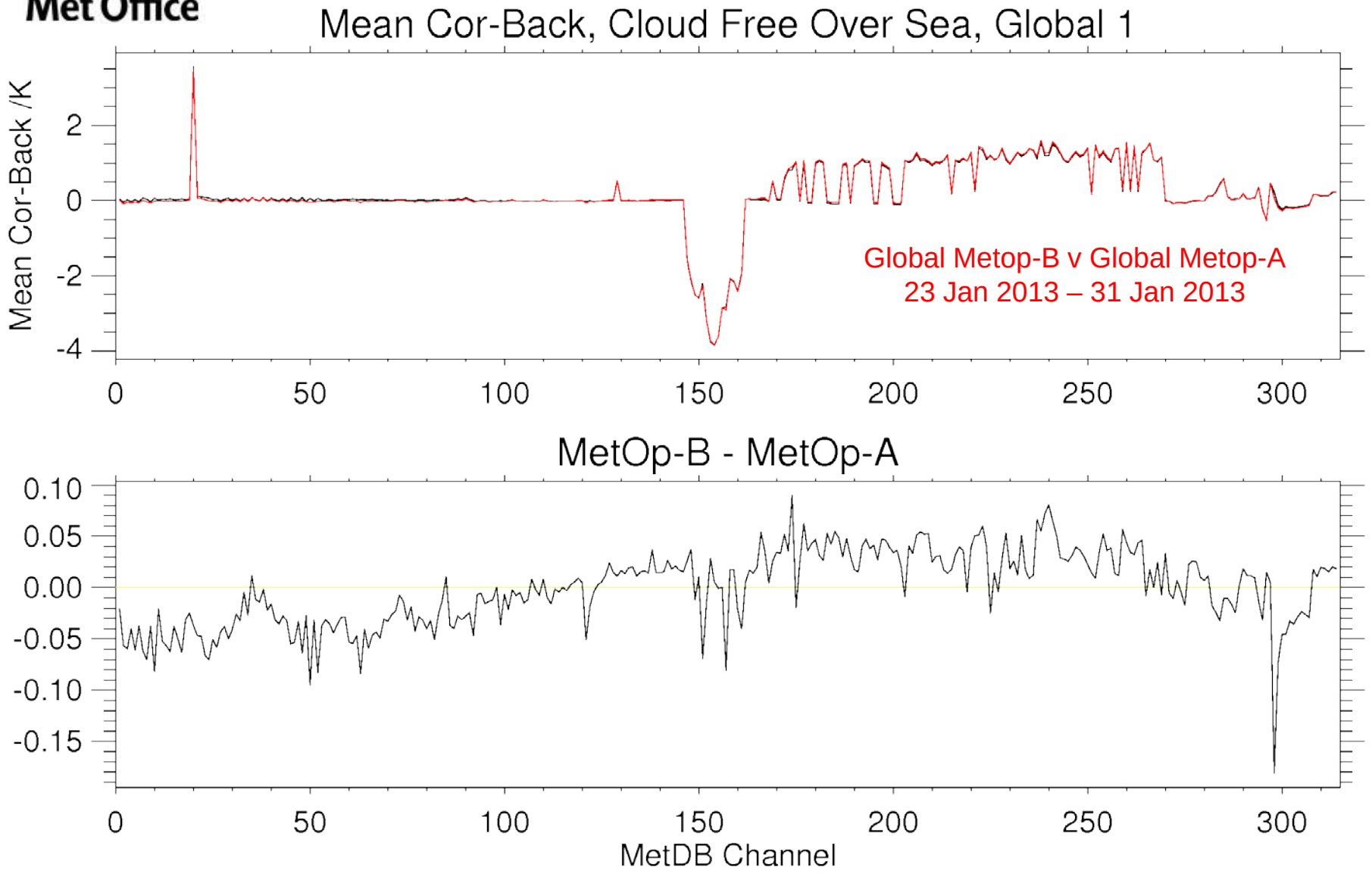
Minimisation – PC scores
Instrument noise + RT error



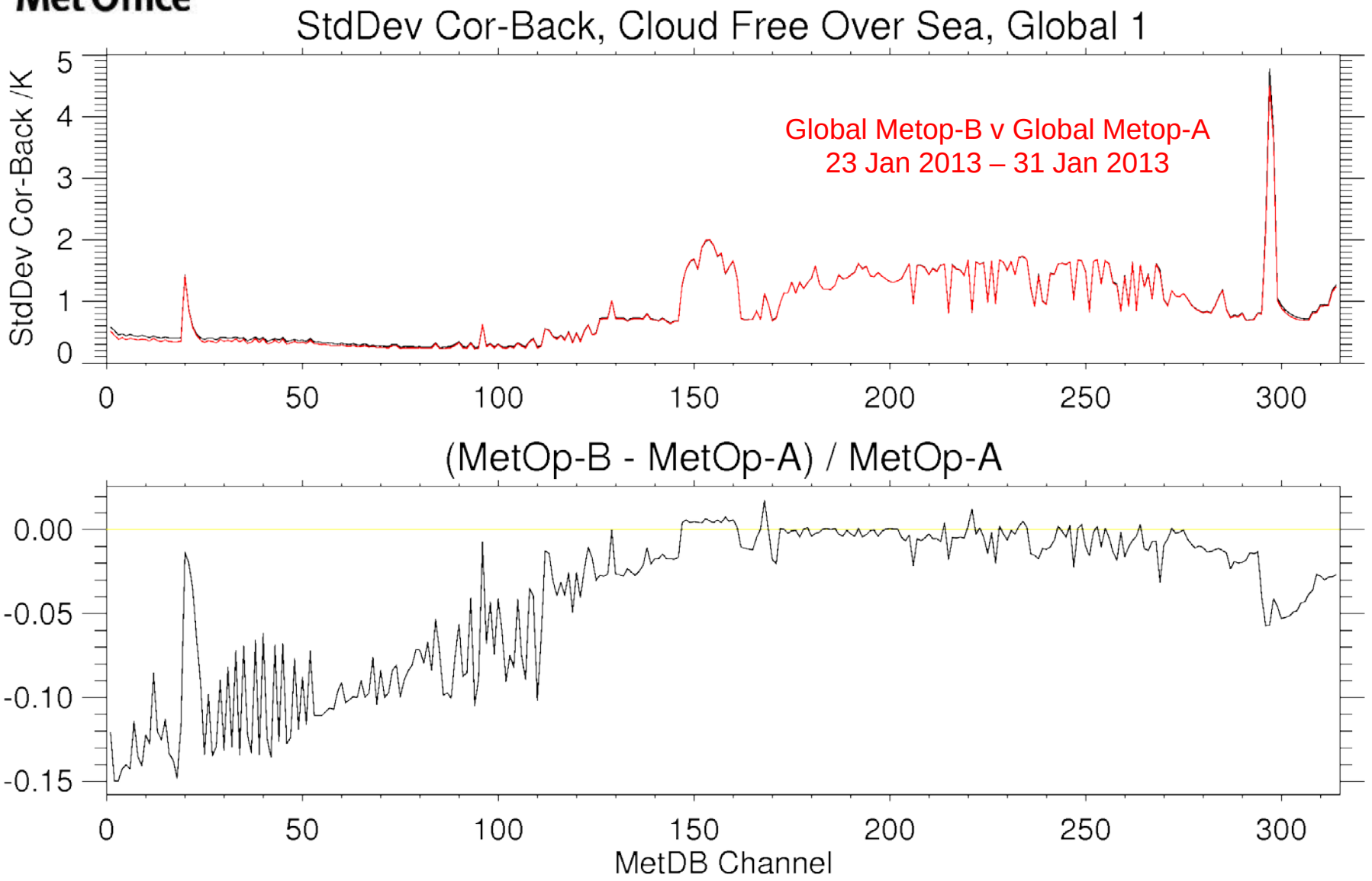
Evaluation of MetOp-B IASI

James Cameron

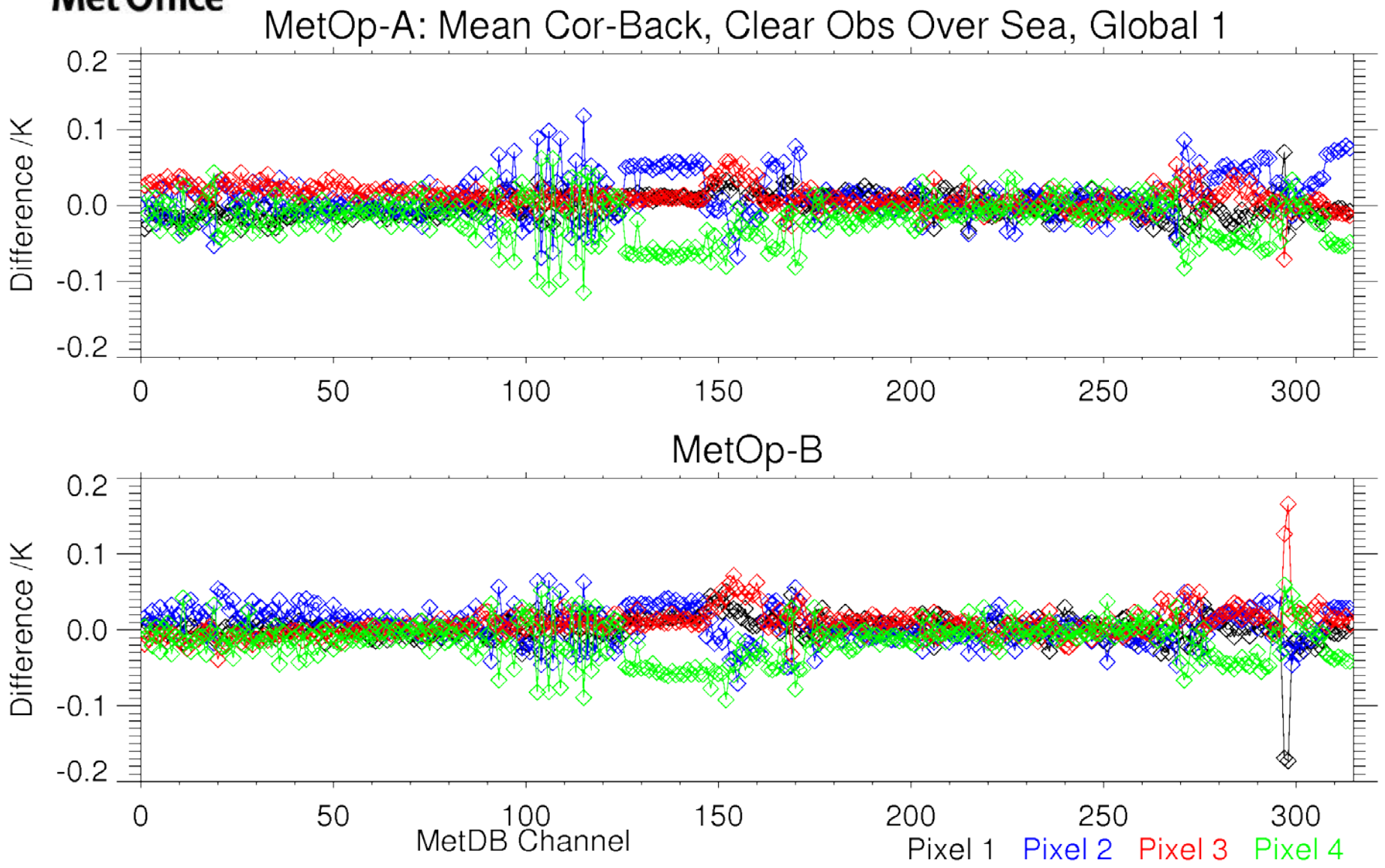
Global Data: Obs-Calc Clear Scenes Mean



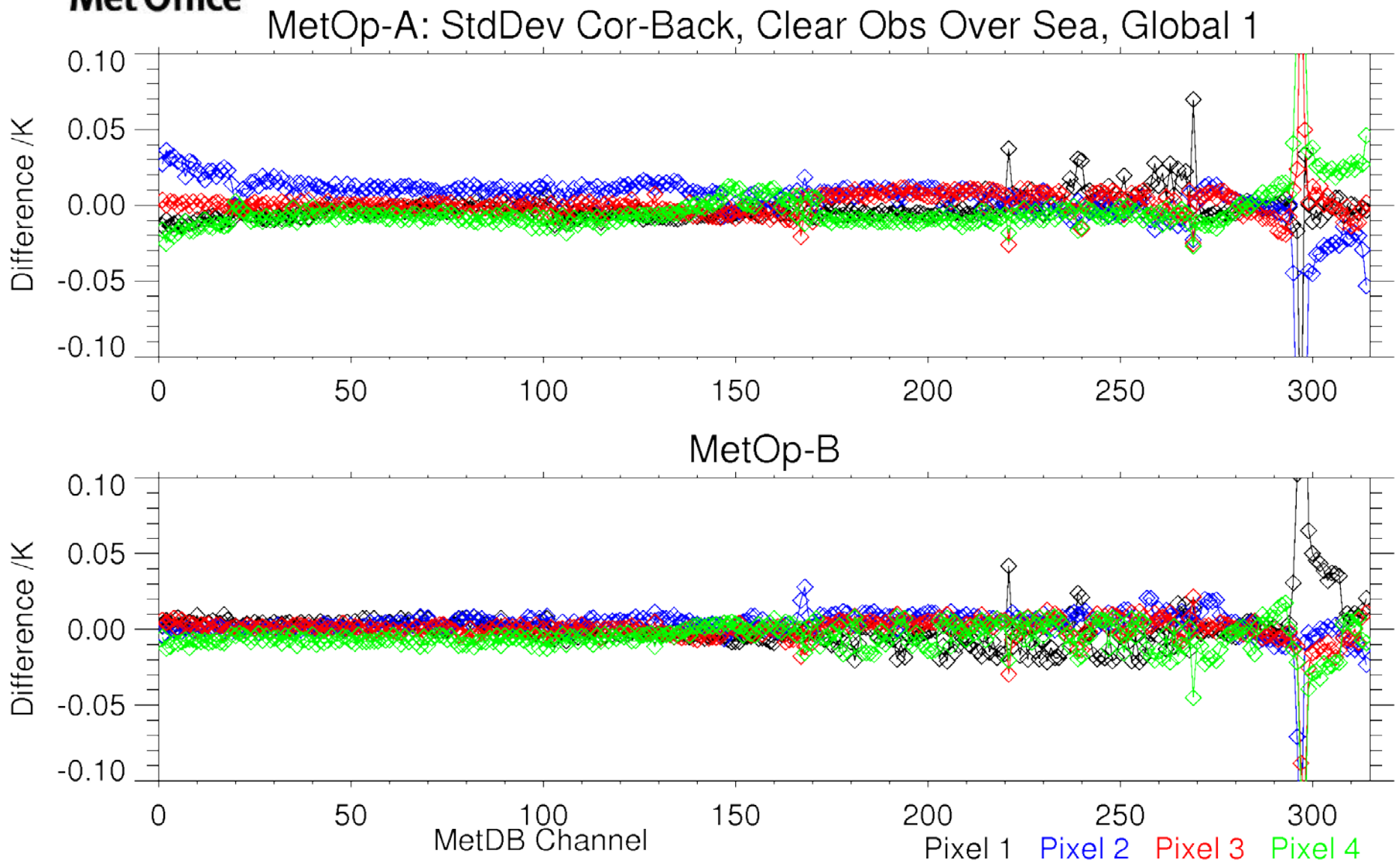
Global data: Obs-Calc Clear Scenes Standard Deviation



Global Data
Inter-pixel differences - Mean



Global Data
Inter-pixel differences – St Dev

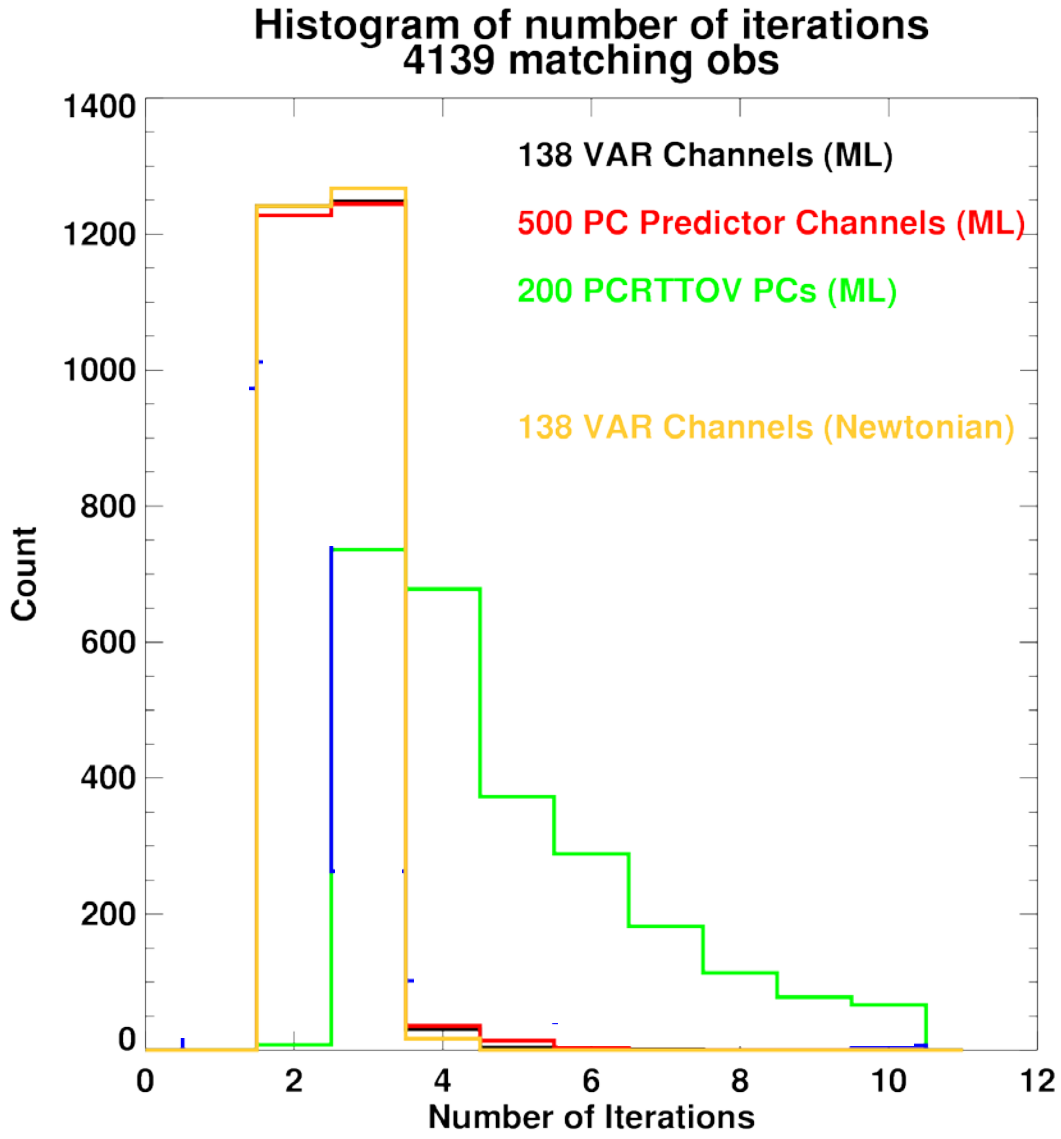


SO_2

- The Met Office is a volcanic ash advisory centre
- We have a new member of staff, Georgina Sawyer, who will be looking at deriving SO_2 amounts from IASI to support this work.

Thanks for listening! Any questions?

Minimisation



Timing and Failures

- For 4348 1D-Var runs
  - RTTOV 4D-Var Channels: 2 hours 8 minutes
  - RTTOV 500 PC-RTTOV predictor channels: 1 hour 51 minutes
  - PC-RTTOV 200 PC scores: 49 minutes

Inst Noise	Not processed	Not converged
RTTOV 4D-VAR	-	-
RTTOV 500	41	-
PCRTTOV	-	168

Inst Noise + RT	Not processed	Not converged
RTTOV 4D-VAR	-	-
RTTOV 500	687	-
PCRTTOV	-	-