A Principal Component Based Version Of The RTTOV Fast Radiative Transfer Model

Marco Matricardi

ECMWF

Work funded by EUMETSAT in support to the product developments for EPS/IASI

IASI 2010 Conference, Sevriers 25-29 January 2007

- A principal component based version of RTTOV has been developed with the aim of :

    1) Investigate the direct assimilation into the ECMWF NWP system of the IASI Principal Component (PC) scores in the short-wave spectral band (band 3).

    2) Improve the computational efficiency of RTTOV.

- The principal component based version of RTTOV allows the fast computation of PC scores and/or IASI Level-1C and AIRS radiances reconstructed from the PC scores.

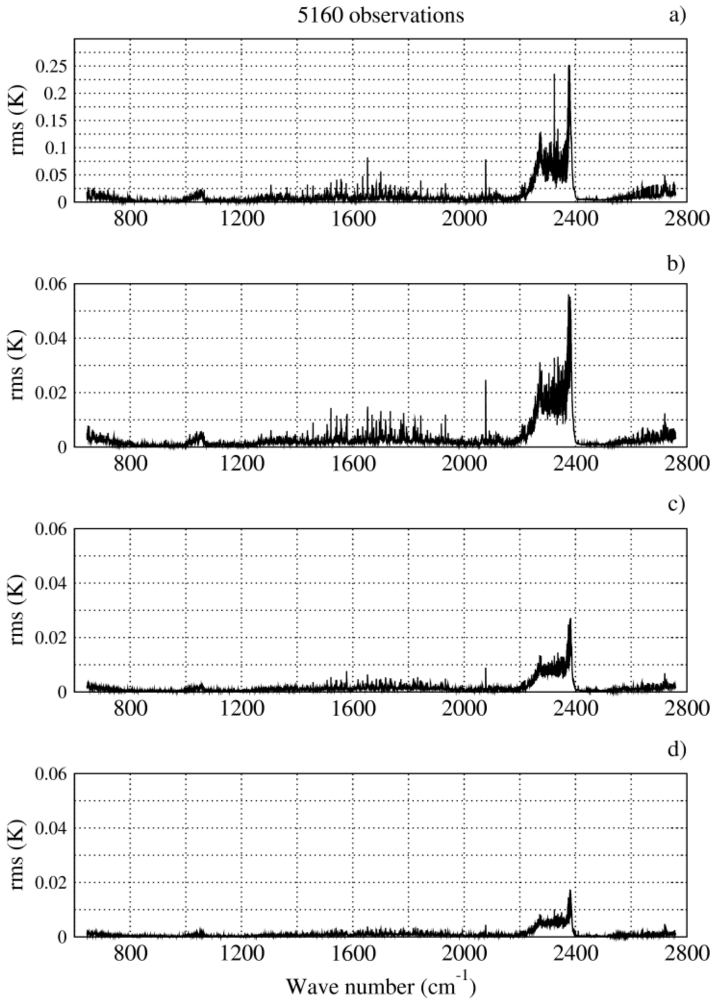
- By using PC analysis we can replace a large number of correlated variables (e.g. radiances computed or observed for a number of different atmospheric situations), with a smaller number of orthogonal variables (called principal components) that still contain most of the information in the original data set.

- If we consider a data set of n observations on m radiances x_1, $x_2,\ldots.x_m$, the radiances for the j^{th} observation can be expressed as a linear combination of the principal component scores z (the values of the principal component variables for the j^{th}observation are known as principal component scores) :

$$x_{i,j} = \sum\nolimits_{l=1,m} a_{l,i} z_{j,l} \qquad i = 1, m$$

Here the a's denote the eigenvectors of the covariance matrix of $x_1, x_2,\ldots.x_m$.

- *Often fewer than m principal components are needed to represent most of the variation in the data.*



100 PC’s

200 PC’s

300 PC’s

400 PC’s

- The scores for the j^{th} principal component can be written as:

$$z_{j,i} = \sum_{l=1,m} a_{j,l}\, x_{i,l} \qquad i=1,n$$

- *If we can establish a direct or indirect statistical relationship between the principal component scores and the atmospheric situations included in a training set, we can then generate a radiance spectrum for any given input atmospheric profile.*

- The development of the PC based RTTOV has involved the following stages:

    1) The selection of a training set of atmospheric profiles.

    2) The computation of accurate line-by-line radiances for the atmospheric situations included in the training set.

    3) The computation of the covariance matrix of the simulated spectra and subsequent computation of the principal component scores.

    4) ***<u>The formulation of a prediction scheme for the computation of the principal component scores</u>.***

- The profiles used to compute the database of line-by-line radiances used in the training of the fast model are chosen to represent the whole range of variation in temperature, absorber amount and surface parameters found in the real atmosphere.

- The training dataset of diverse atmospheric profiles used in this study is based on the database generated by Chevallier et al. (2006) using the operational suite of the ECMWF forecasting system during the period July 2006- June 2007.

- We selected a total of 1039 profiles of temperature, water vapour and ozone over the sea. In addition to the profiles of temperature and absorber amount, the 1039 profile data set contains ancillary information on surface temperature, skin temperature, surface pressure and wind speed at 10m.

- The line-by-line database used in this study was calculated using version 11.1 of the LBLRTM model in conjunction with a model of the surface emissivity that takes into account the effect of reflected emission from the surface.

- The line parameters were obtained by merging data from HITRAN_2000, HITRAN_2004/2006 and GEISA 2003.

- In our line-by-line computations we included 17 atmospheric constituents and have allowed to vary only H_2O and O_3.

- Radiances were computed for each atmospheric profile and five surface zenith angles, namely, the angles for which the secant has equally spaced values from 1 to 2. This makes a total of 5195 observations for each channel.

- In the PC based RTTOV, the simulated PC scores are computed using a linear regression scheme. The independent variables (predictors) used in the regression are a selected number of polychromatic top of the atmosphere radiances computed using the standard RTTOV fast transmittance model. The advantage of this approach is two-fold:

  1) The use of polychromatic radiances as predictors requires only a limited number of modifications to the RTTOV suite.

  2) No major additional effort is required for the long term maintenance of the model.

- In addition, the use of RTTOV radiances as predictors makes the architecture of the model suitable for utilization in a scheme where the adjoint technique is used.

- The spectra of IASI contain many channels that have very similar spectral signatures and channels that are similar in content can cause problem in a linear regression since they can be very highly correlated.

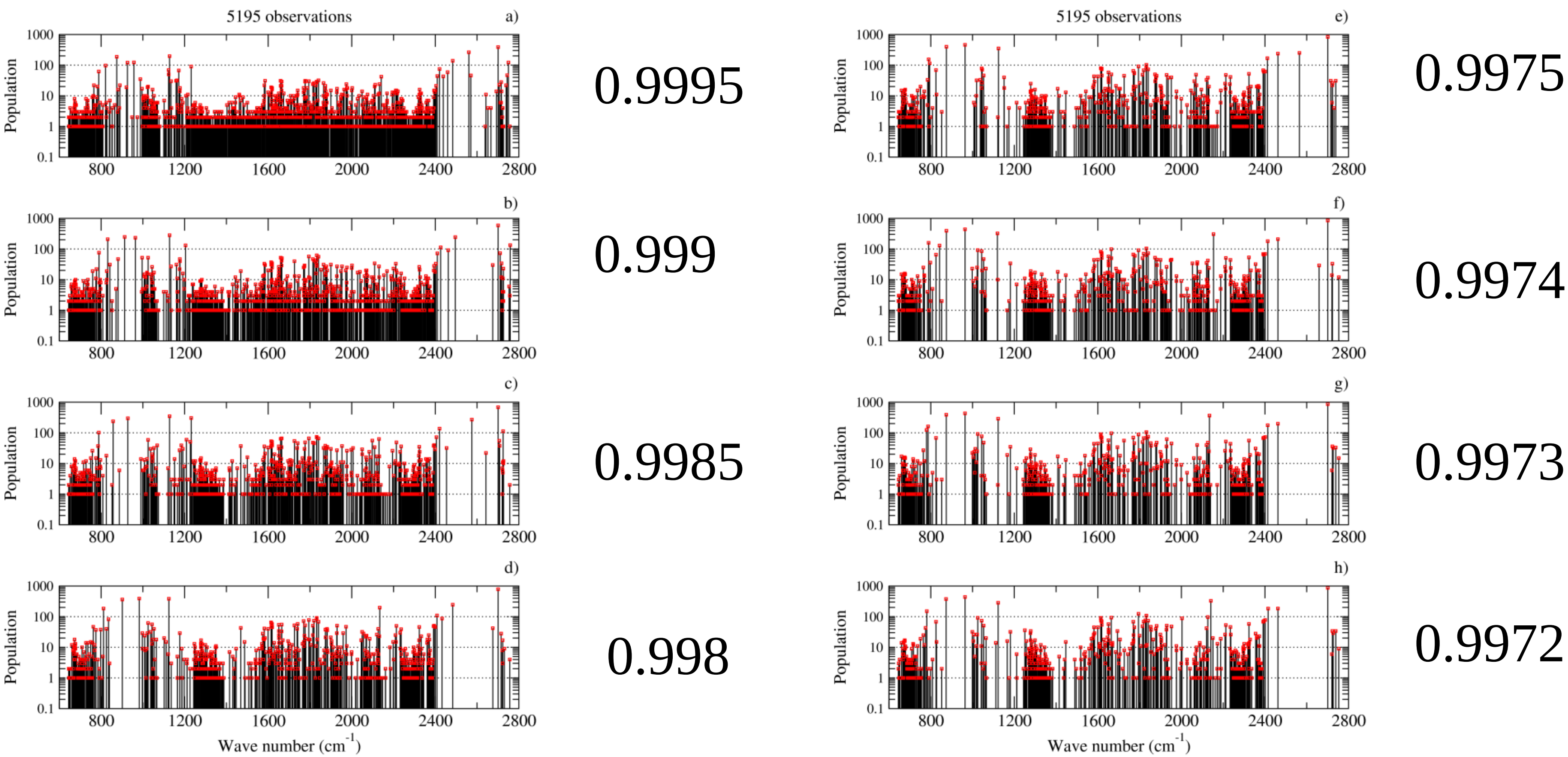
- To overcome this problem we have arranged the channels in clusters that contain highly correlated channels (McMillin and Goldberg, 1997).

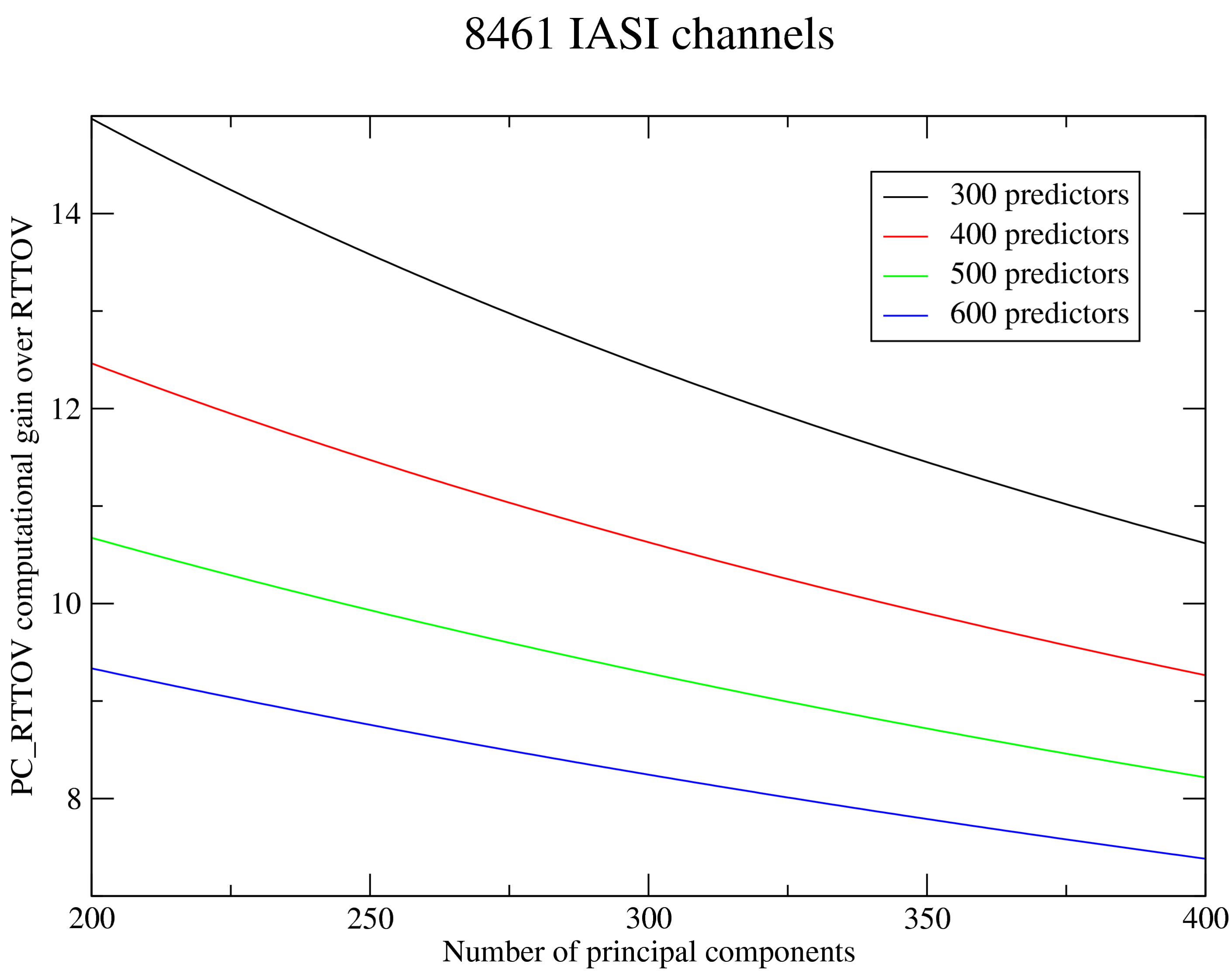
- To compute the clusters we search for the channel with the largest standard deviation (lead channel) and find channels that are correlated to the lead channel with a correlation greater that a given threshold.

- We then remove the channel cluster from the selection pool and repeat the process until no more clusters can be obtained. The lead channels in the clusters are then used as predictors in the regression.

- To choose the optimal value of the threshold we have finely tuned the threshold to the value that allows the best fit to the line-by-line radiances.

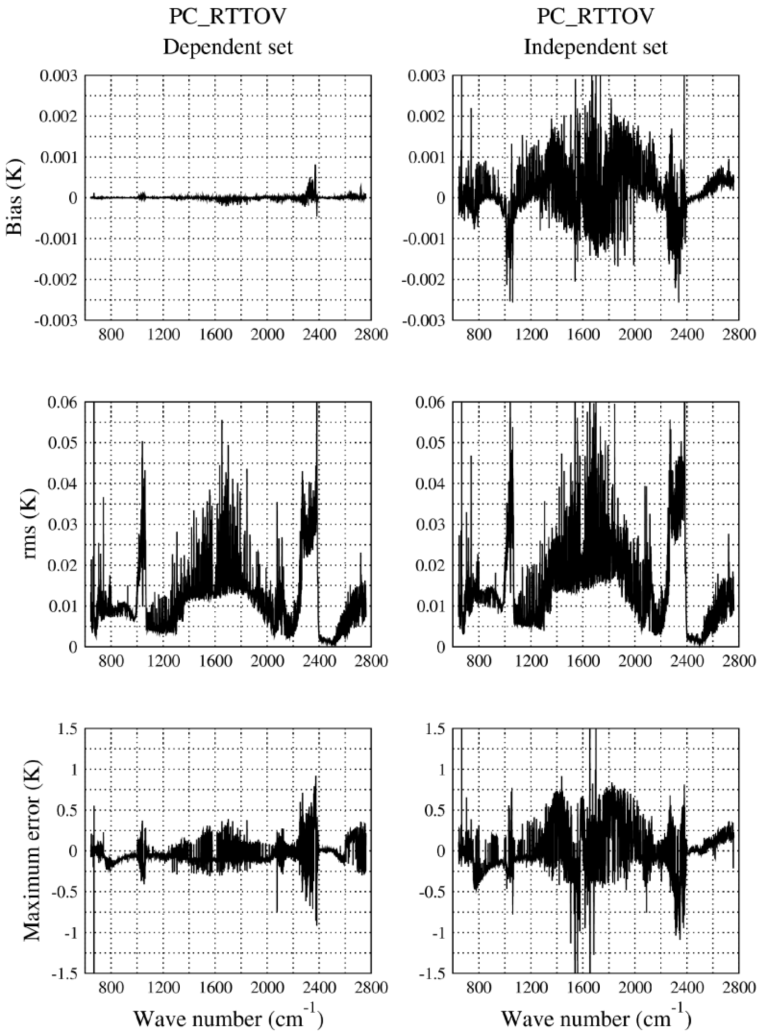
- For IASI we have selected a correlation threshold equal to 0.9976. This value allows a maximum number of 669 lead channels (i.e. predictors) to be used in the regression.



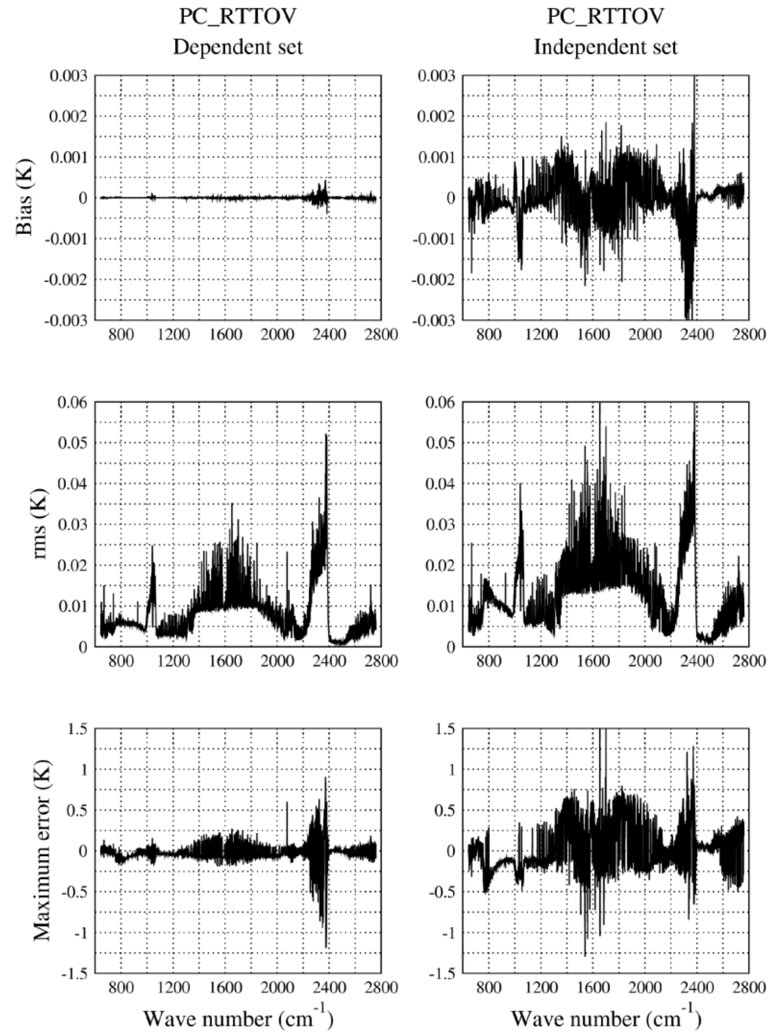

8461 IASI channels
300 predictors
400 predictors
500 predictors
600 predictors
PC_RTTOV computational gain over RTTOV
Number of principal components
8
10
12
14
200
250
300
350
400

Fit of PC_RTTOV to line-by-line radiances

200 PC’s
300 Predictors

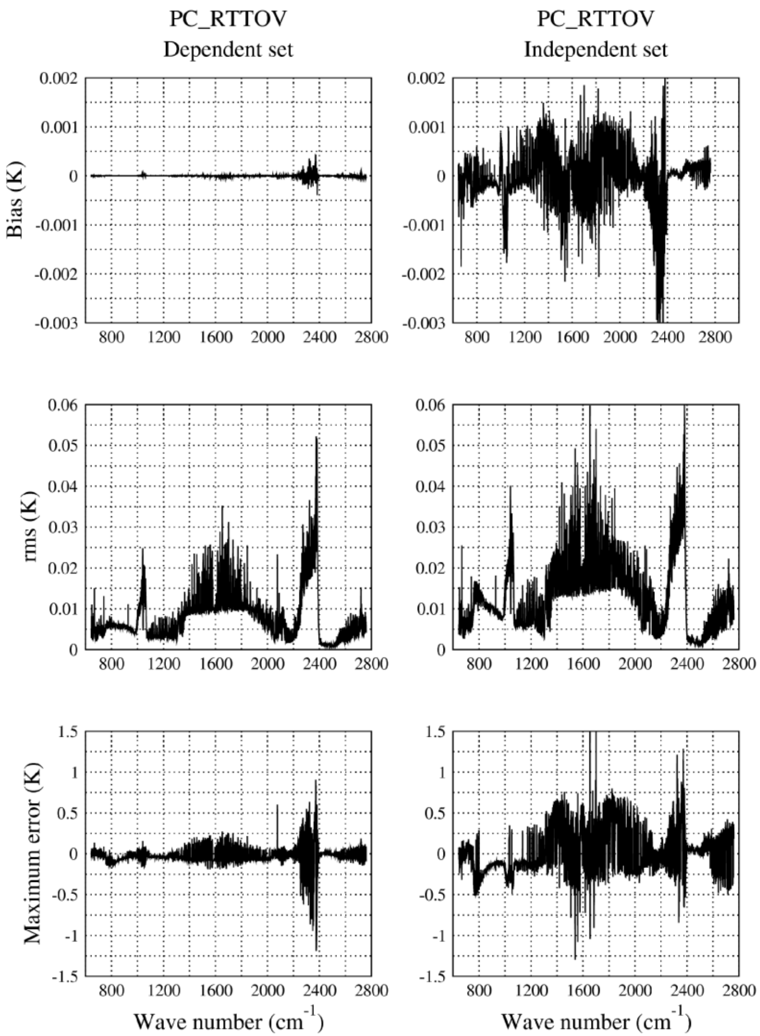


200 PC’s
600 Predictors

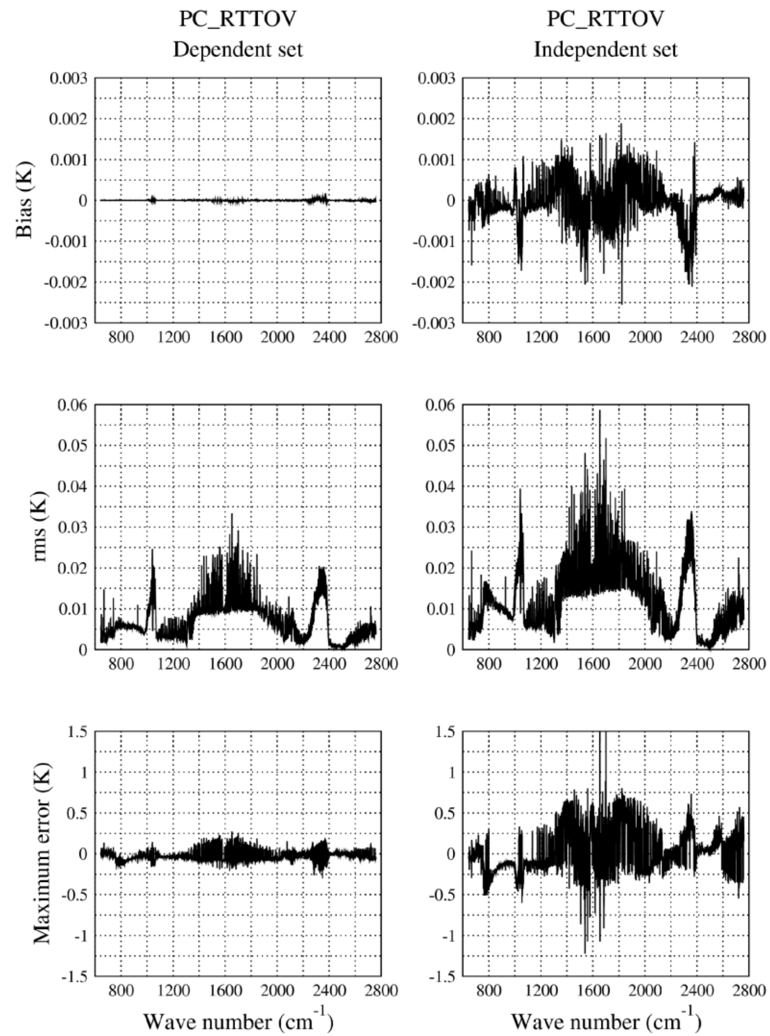


Fit of PC_RTTOV to line_by_line radiances

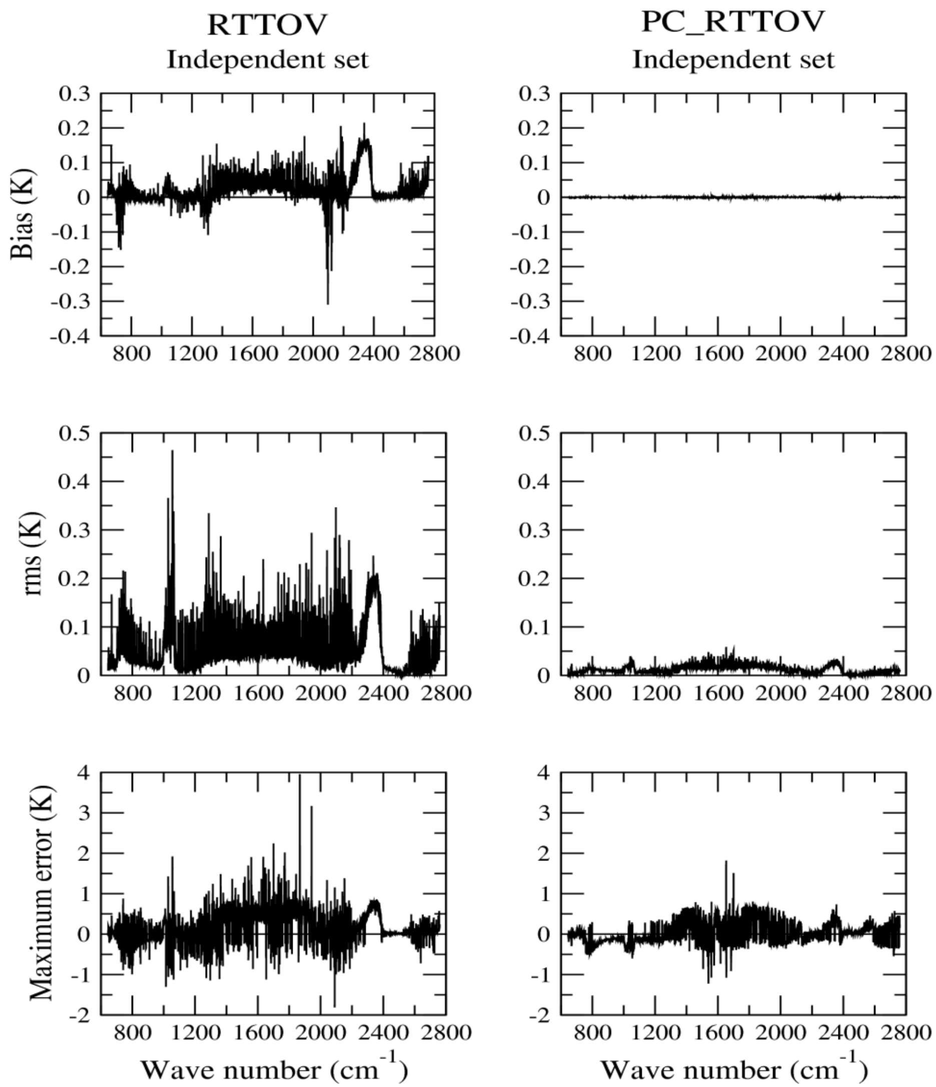
200 PC's
600 Predictors



400 PC's
600 Predictors

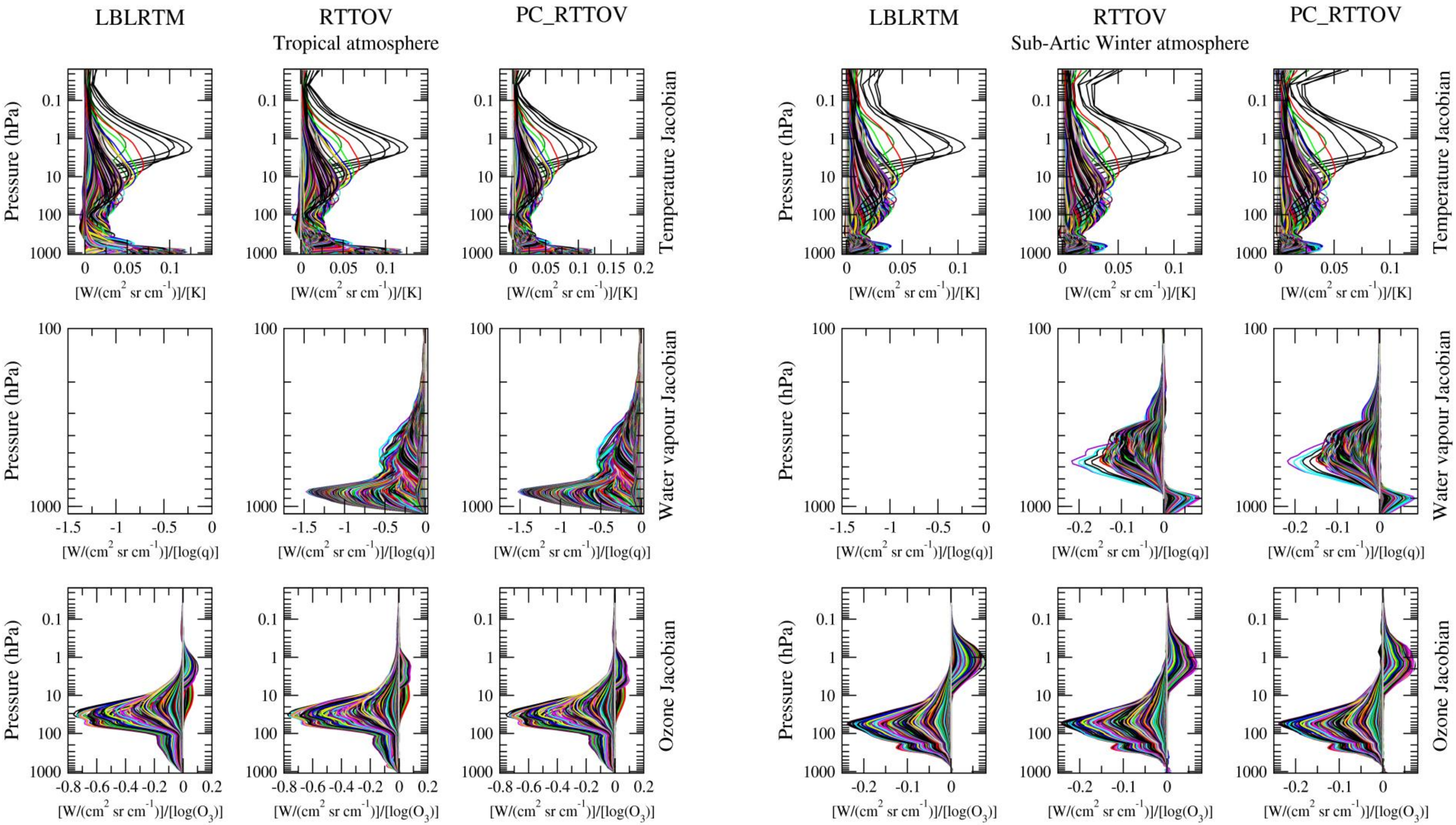


Fit of RTTOV and PC_RTTOV to line-by-line radiances



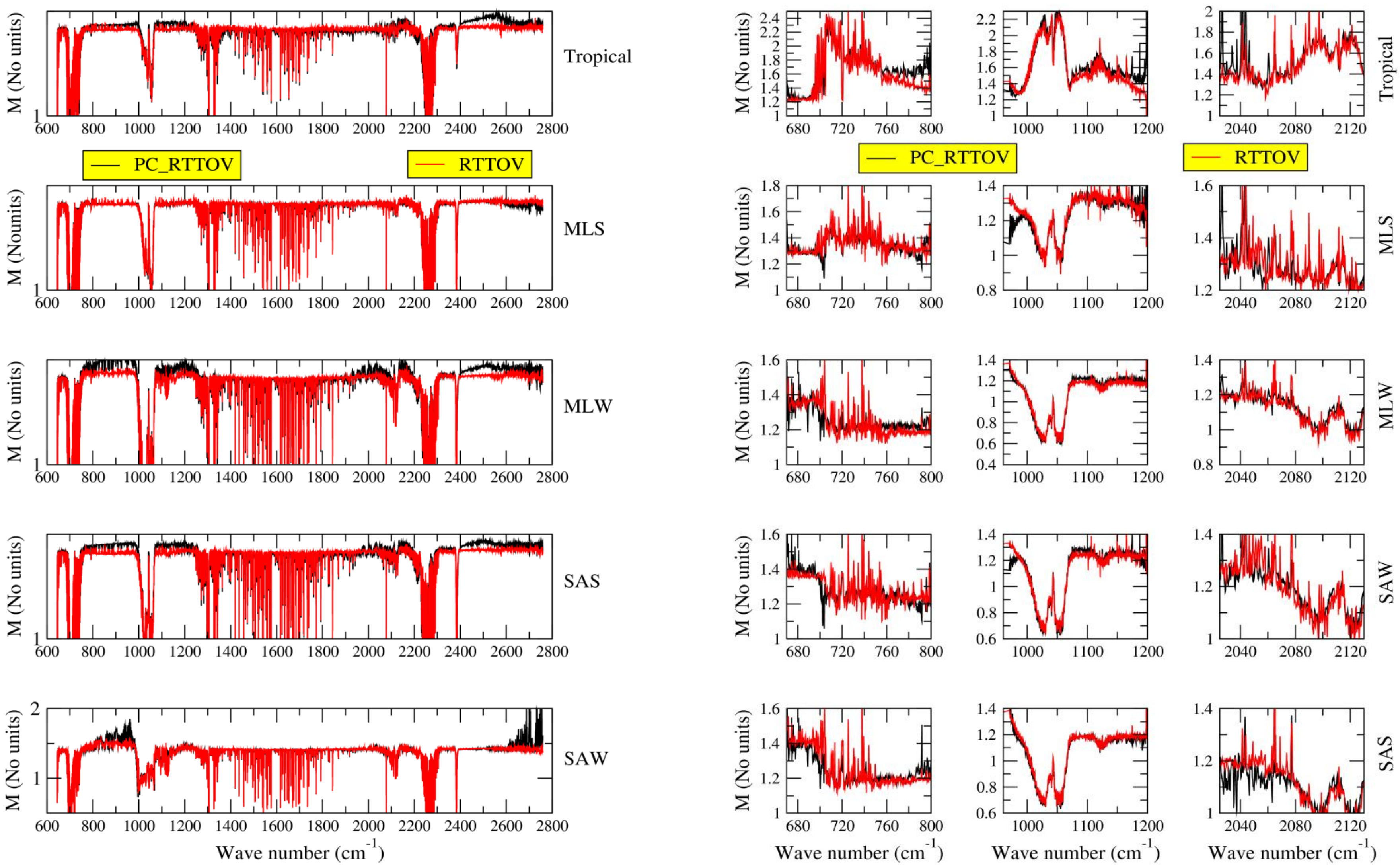
PC_RTTOV:
400 PC's
600 Predictors

Jacobians

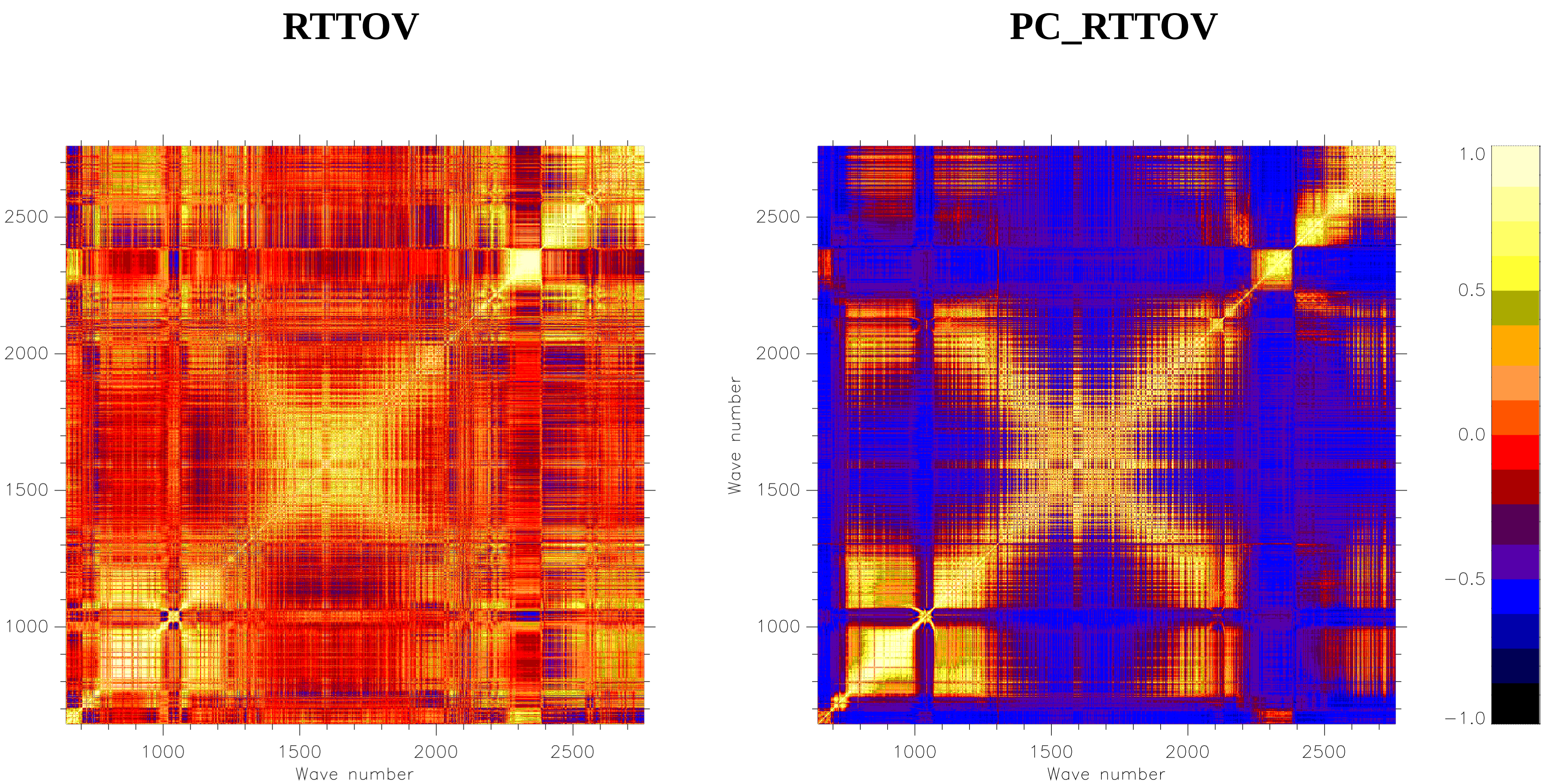


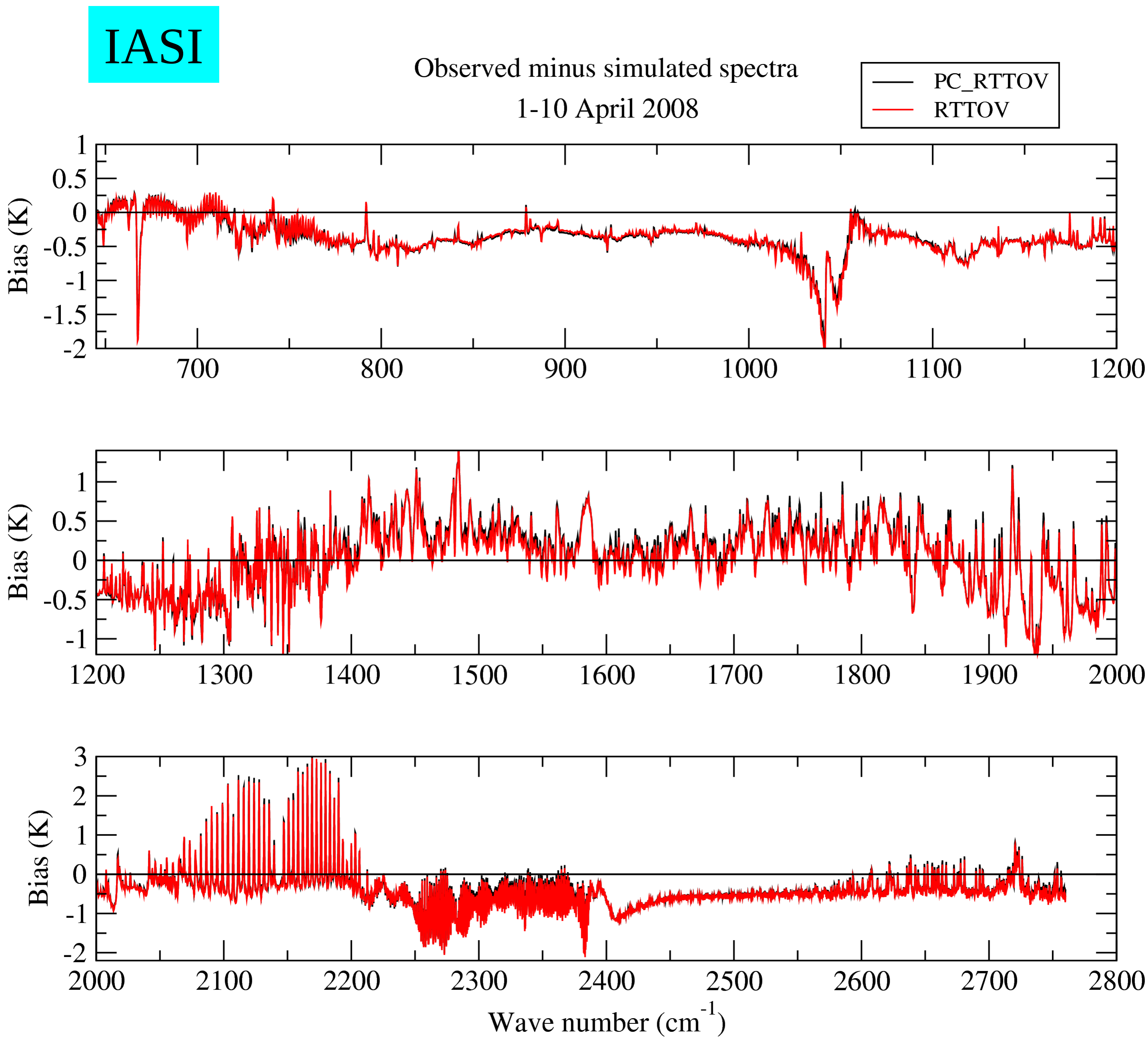
LBLRTM
RTTOV
PC_RTTOV
Tropical atmosphere
LBLRTM
RTTOV
PC_RTTOV
Sub-Artic Winter atmosphere
Temperature Jacobian
Water vapour Jacobian
Ozone Jacobian
Pressure (hPa)
[W/(cm2 sr cm-1)]/[K]
[W/(cm2 sr cm-1)]/[log(q)]
[W/(cm2 sr cm-1)]/[log(O3)]

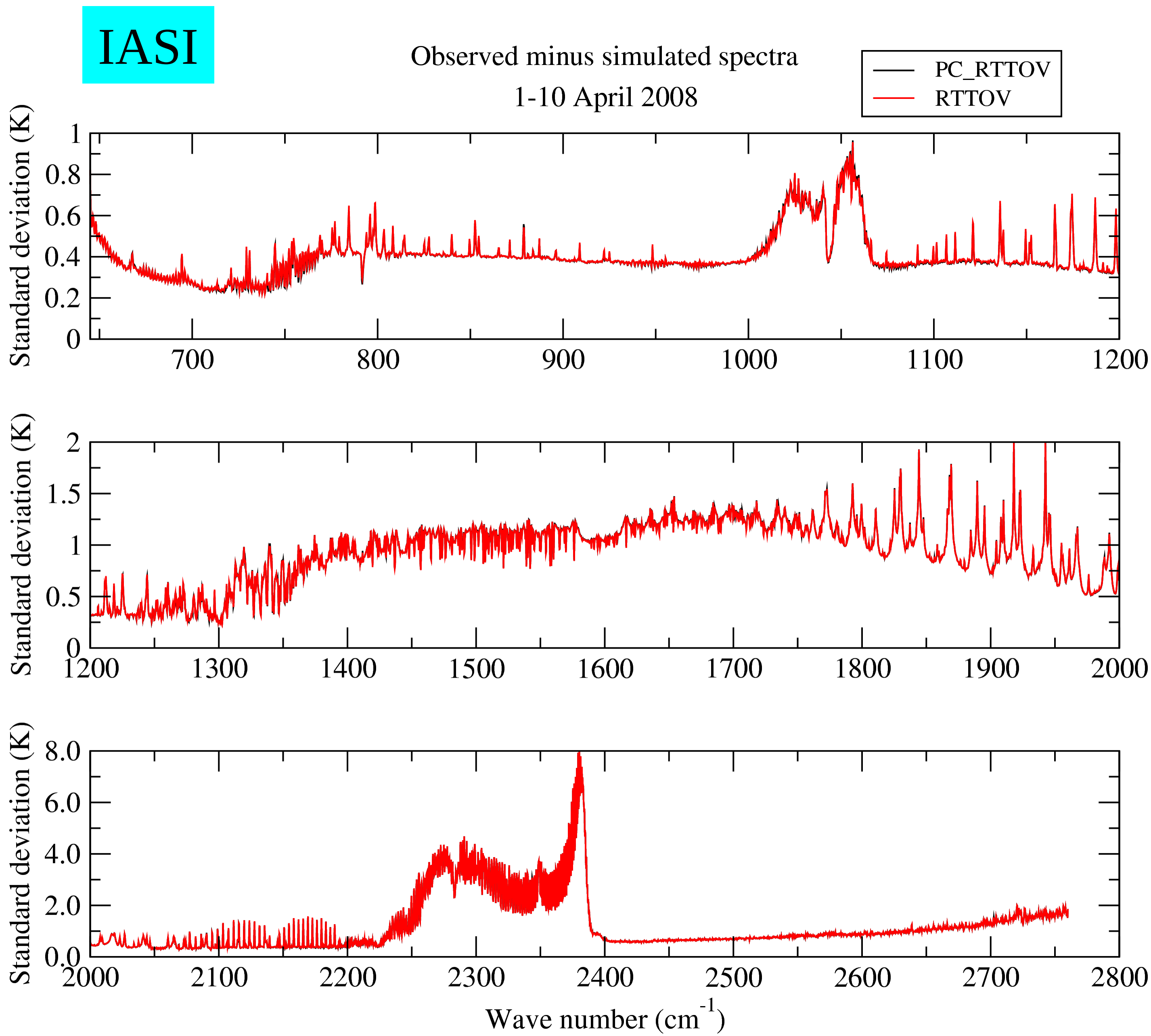
$$M = \sqrt{\frac{\sum_{j=1}^{100}(J_j^{RTTOV} - J_j^{LBLRTM})^2}{\sum_{j=1}^{100}(J_j^{LBLRTM})^2}}$$



Fast model error correlation matrix: 5195 profile independent set




IASI
Observed minus simulated spectra
1-10 April 2008
PC_RTTOV
RTTOV
Bias (K)
Wave number (cm-1)


IASI
Observed minus simulated spectra
1-10 April 2008
PC_RTTOV
RTTOV
Standard deviation (K)
Wave number (cm-1)