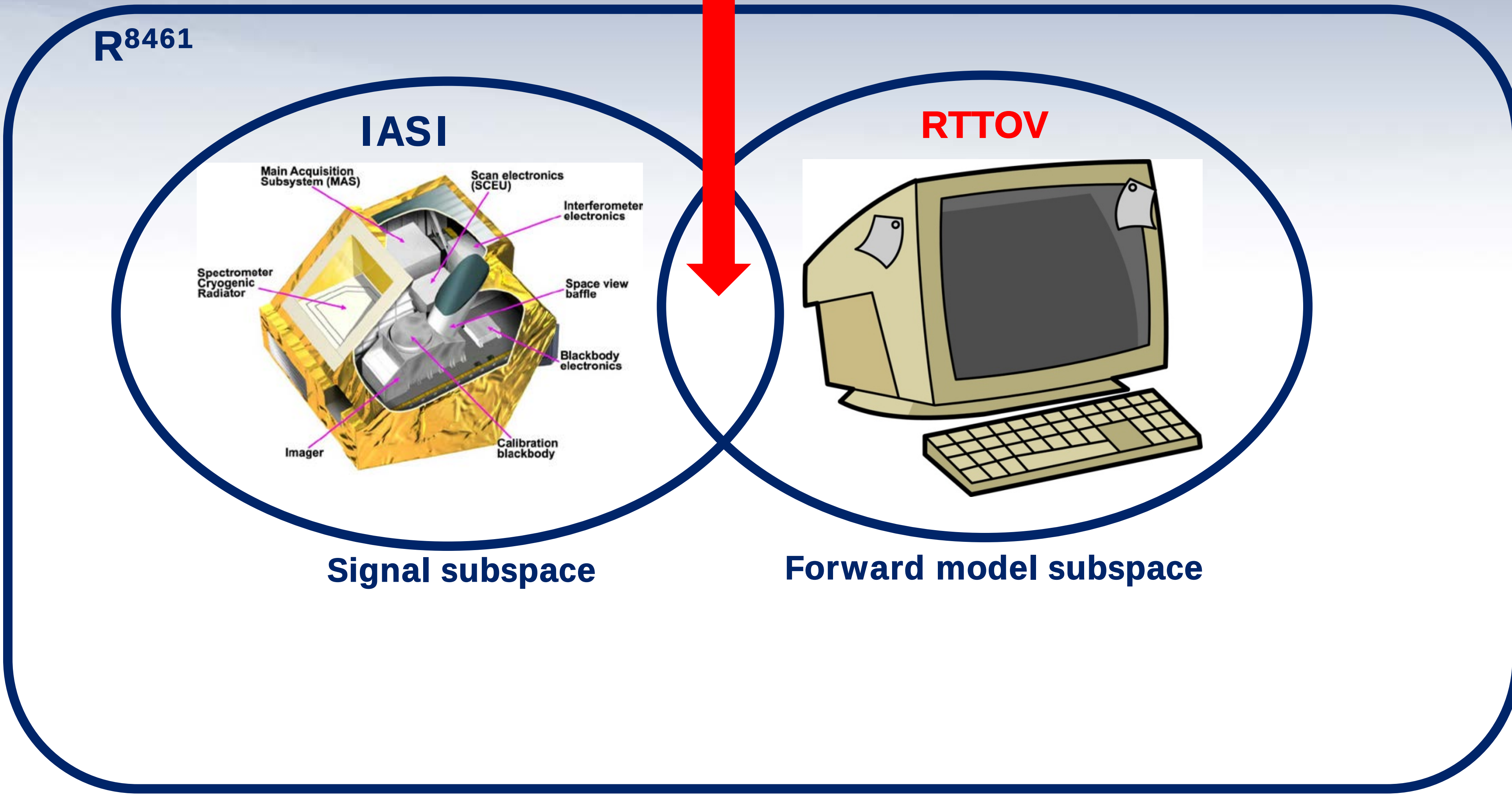
IASI retrievals in the intersection of the signal and forward model subspaces

Tim Hultberg , Thomas August, Marc Crapeau

Outline

1. Signal and forward model subspaces
2. Reconstructed radiances and channel selection
3. Linear regression with IASI, AMSU and MHS
4. Bias correction
5. Retrieval results and wrap up

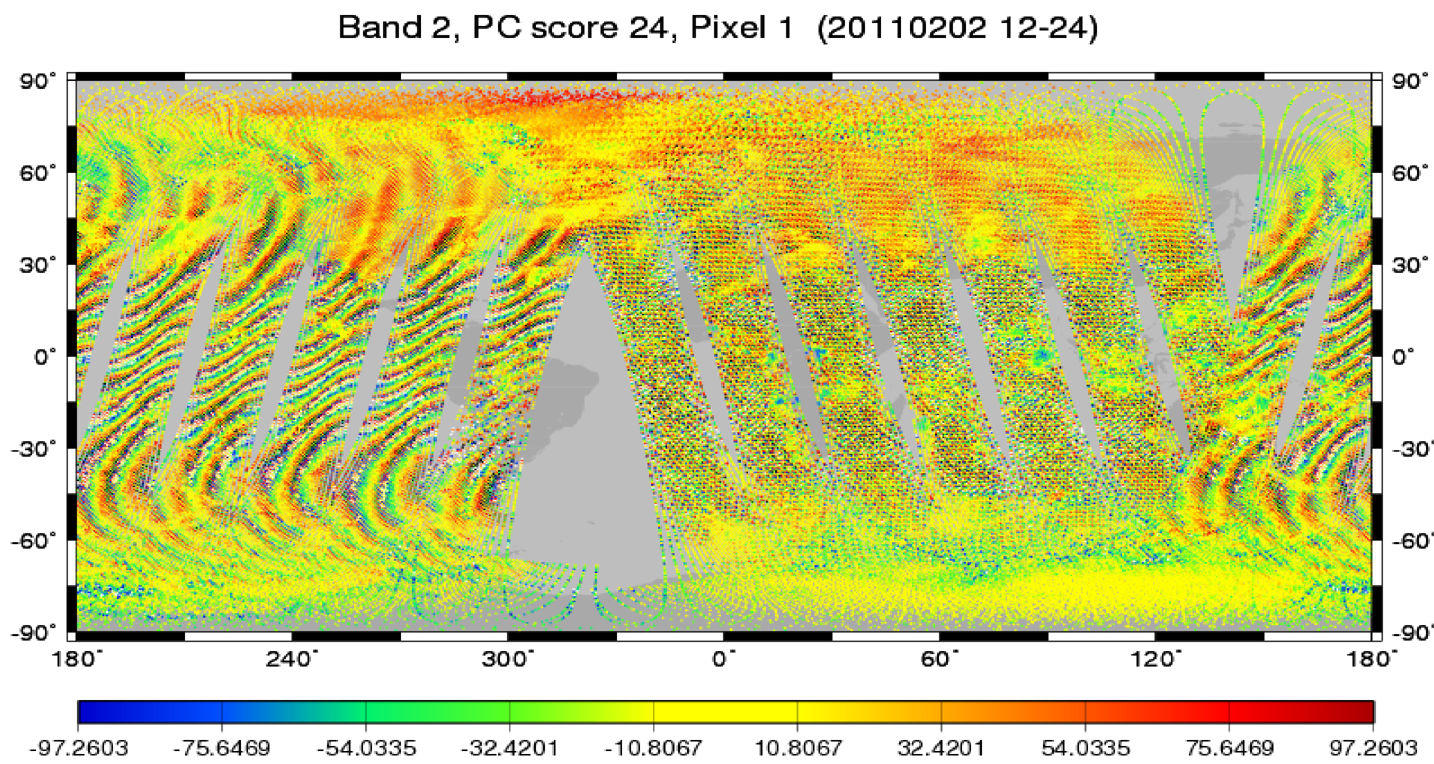
The intersection subspace



Measurements and forward model simulated spectra

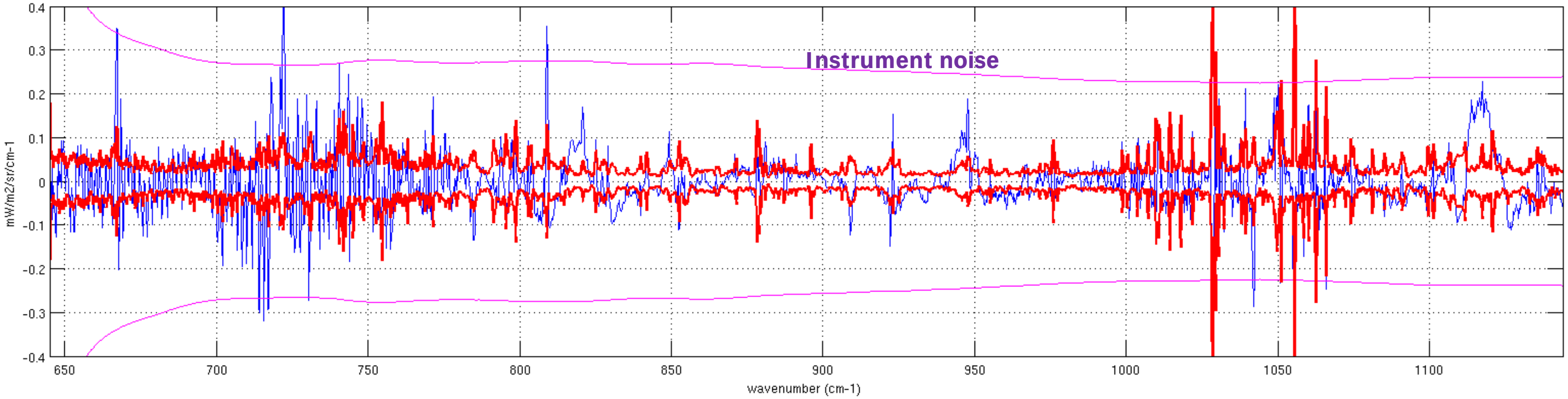
Live in two different linear subspaces

Features found in measurements but not in simulated spectra due to: undetected cloud, un-modelled trace gases, instrument and processing artefacts.



EUMETSAT

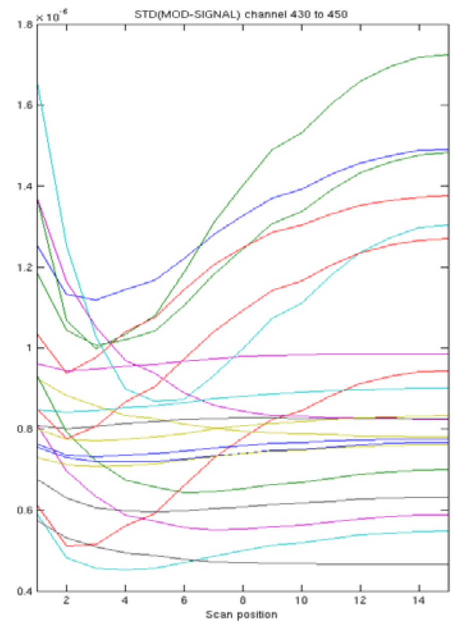
Features found in simulated spectra but not in measurements

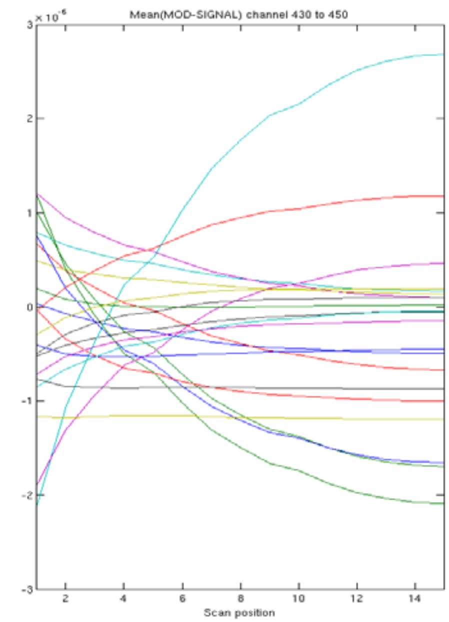


Bias and standard deviation of residual after projecting RTTOV 10 spectra onto signal space.

Lower limit on forward model error.

Strong scan angle dependence of forward model error for some channel.





The signal and forward model subspaces

The subspaces are determined by truncated set of eigenvectors of the covariance matrices of the measured and simulated radiances respectively.

$$E_S \in R^{m \times p} \qquad E_F \in R^{m \times p}$$

The dimension of the intersection space equals the dimension of the null space of the m times 2p matrix [E_S E_F]

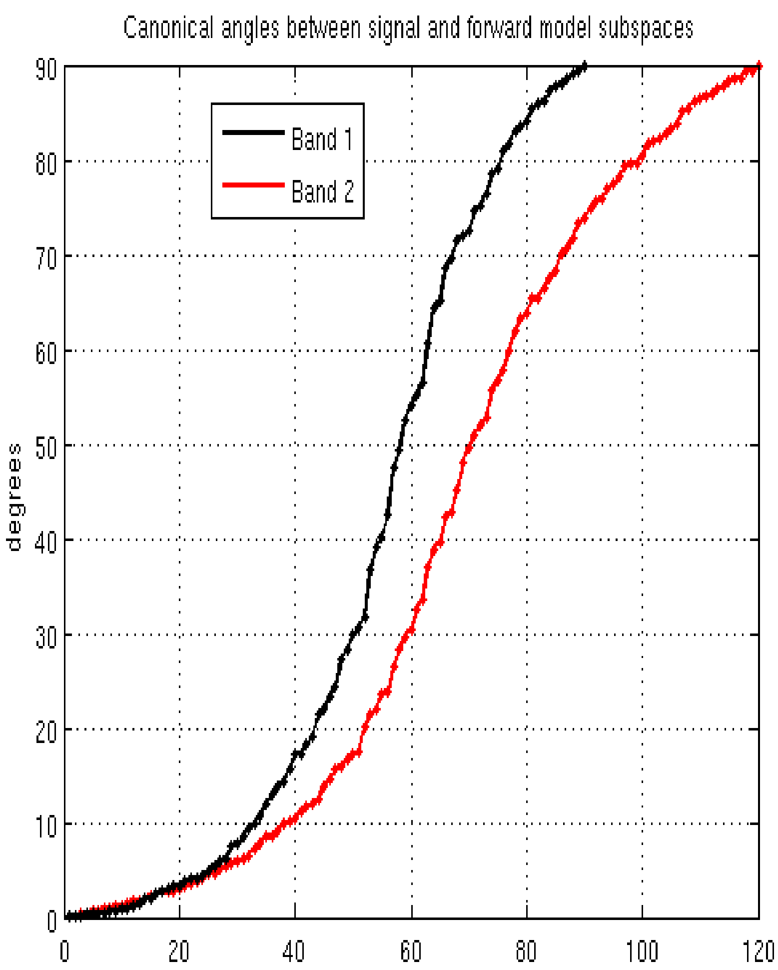
It turns out that the intersection of the two subspaces is empty. But clearly directions very close to each other can be found in the two subspaces.

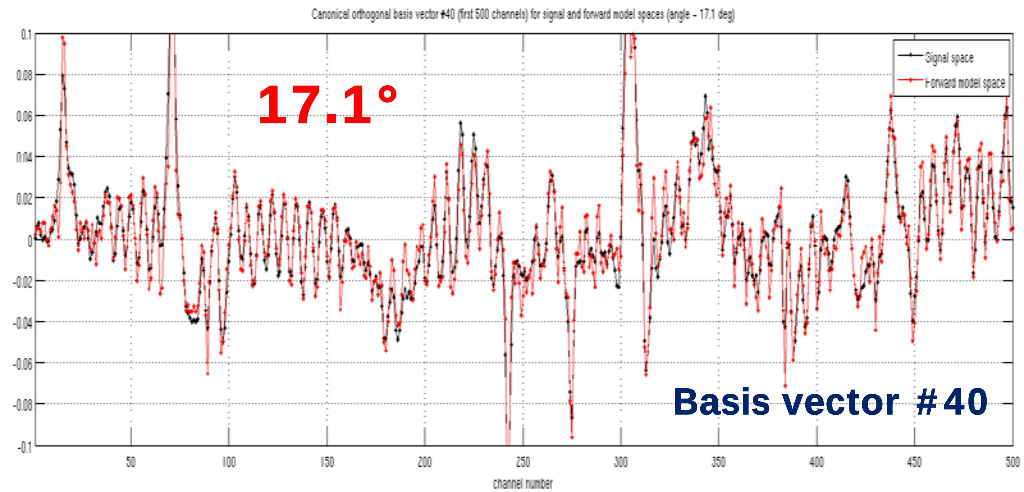
$$E_S^T E_F = USV^T$$

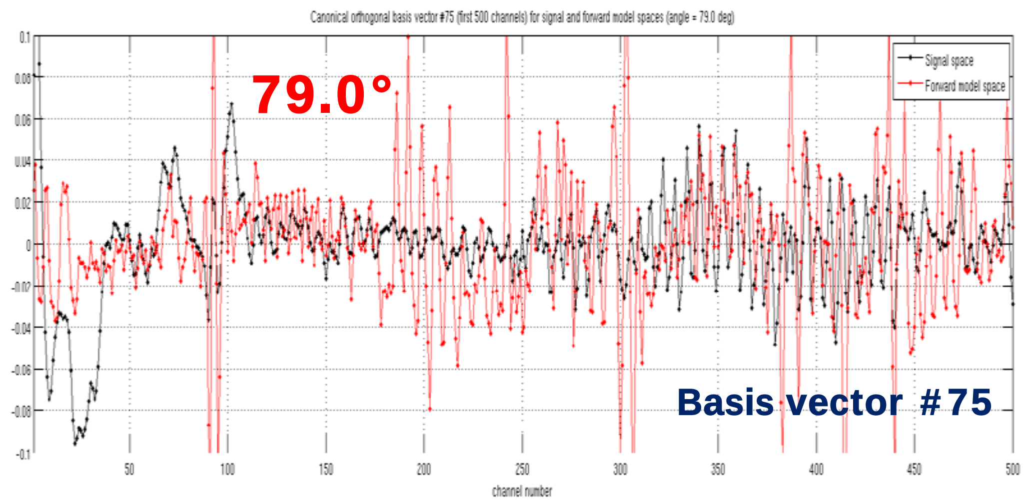
$$\widehat{E_S} = E_S U \qquad \widehat{E_F} = E_F V$$

$\hat{E}_F$ and $\hat{E}_S$ are bi-orthogonal and the canonical angles between the two subspaces are given by arccos(S_{ii}) in ascending order

Canonical angles between subspaces

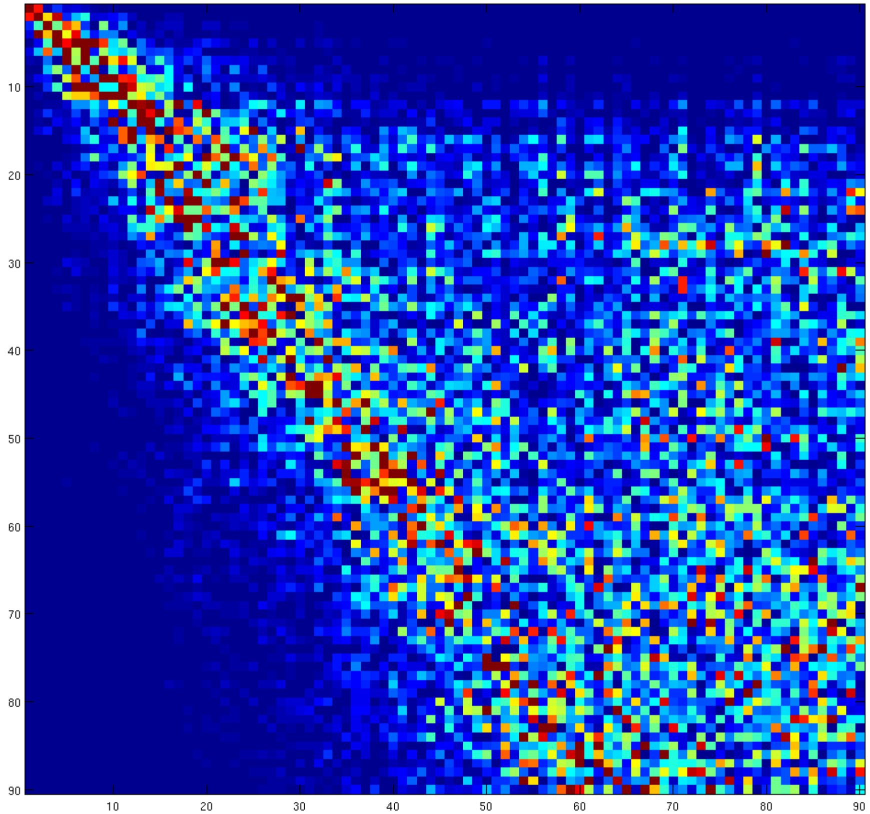




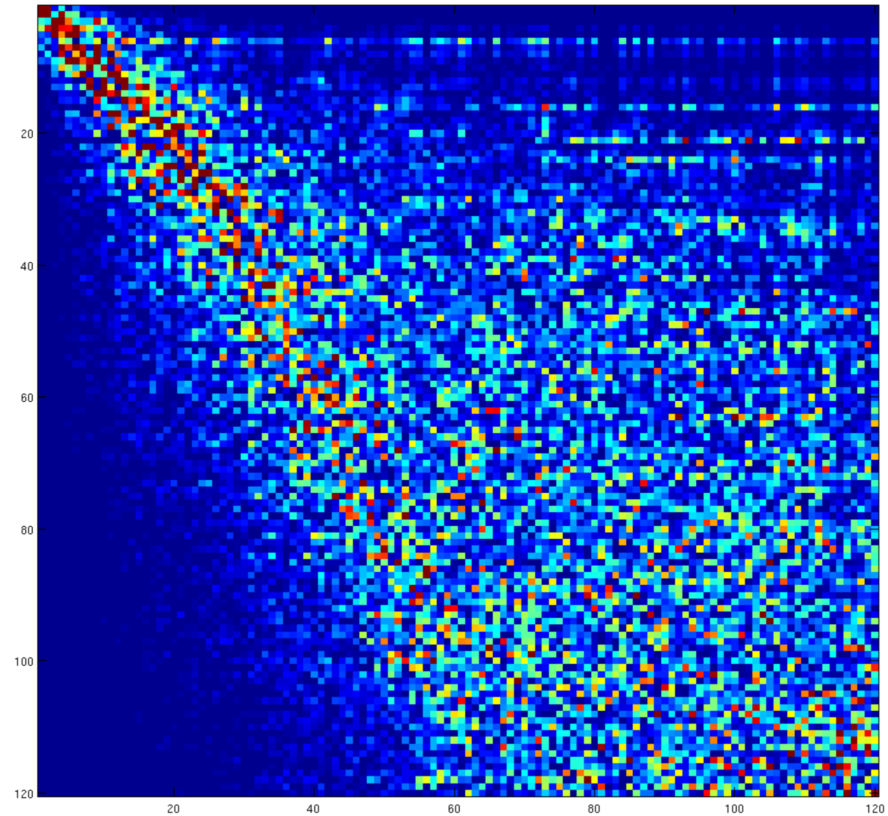


The U matrix (which rotates the signal space basis)

Band 1



Band 2

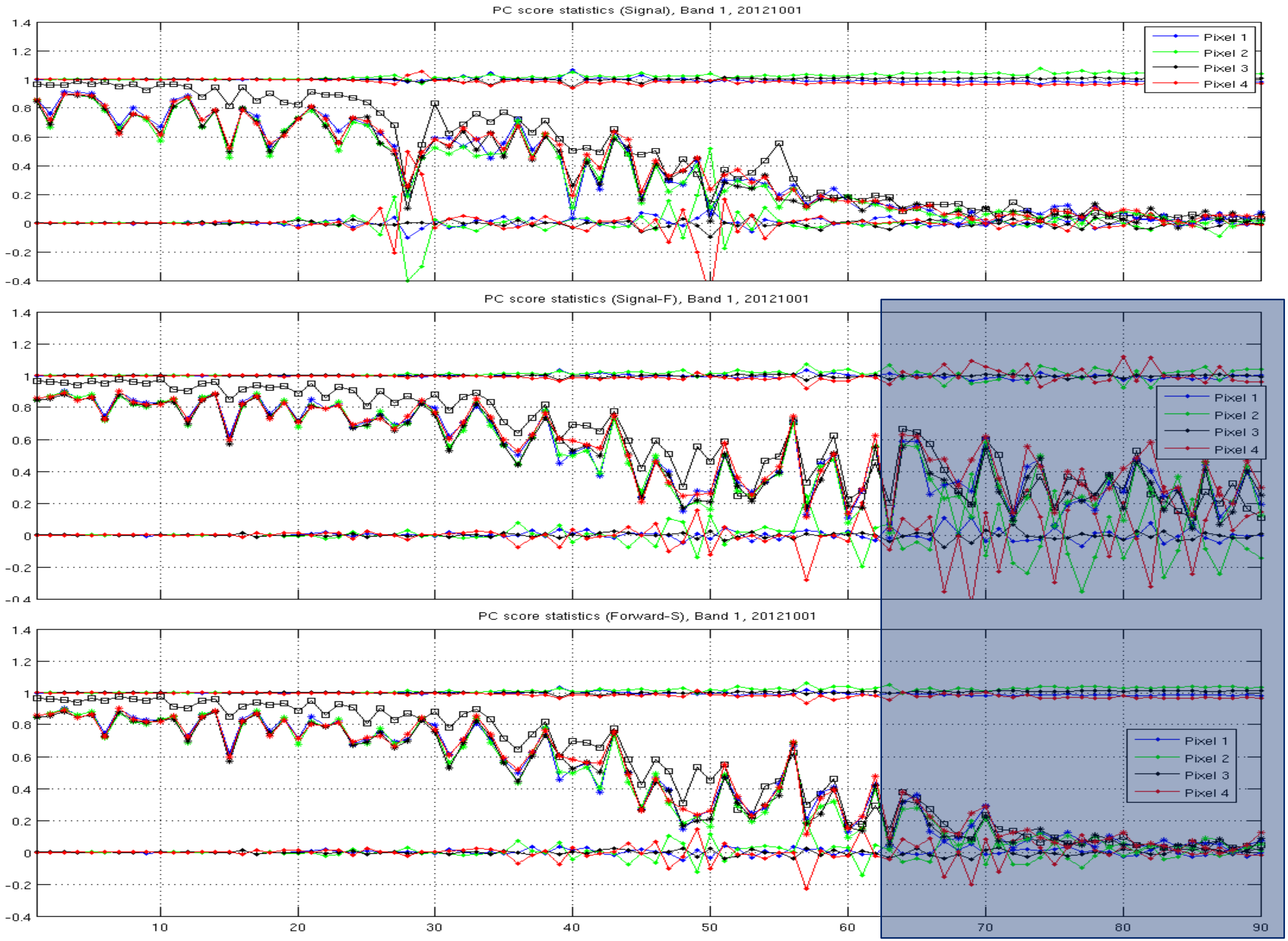


$$\widehat{E_S} = E_S U$$

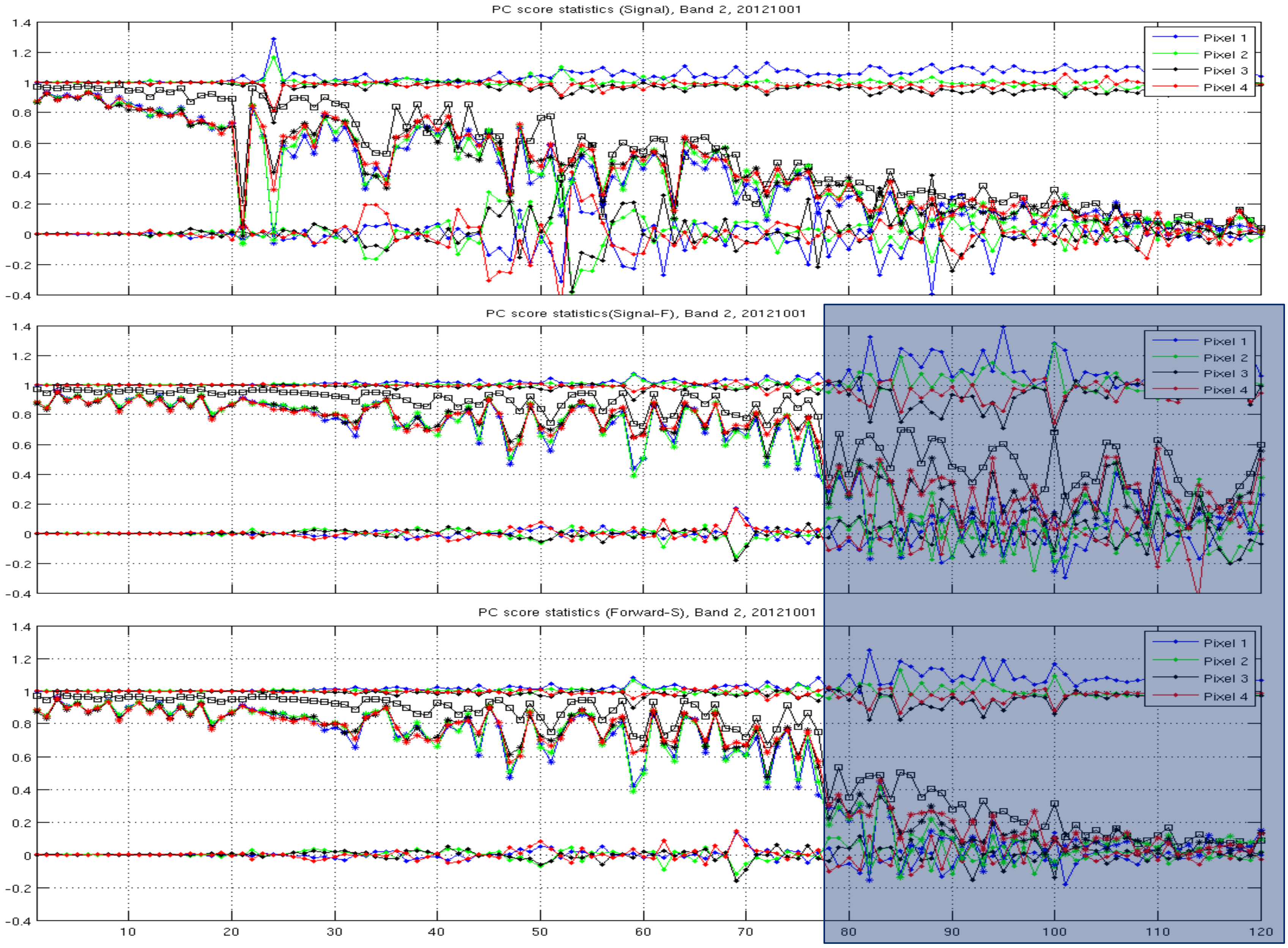
Columns of U indicate which basis vectors of E_S contribute to basis vectors in $\hat{E}_S$

EUMETSAT

PC score statistics, Band 1



PC scores statistics, Band 2



Canonical angles

New bases for the signal and forward model subspaces, in which similar directions in the two spaces are “moved forward”

Truncation can be used to filter out components of the measurements which are not found in the forward model space and vice versa

Instead of the intersection space (which is empty) we obtain two “approximations”: One is a subspace of the signal space (SF), the other is a subset of the forward model space (FS)

For the retrievals we keep 62 basis vectors in Band 1 and 77 basis vectors in Band 2 (corresponding to canonical angles lower than 60 degrees) and project the measurements onto the FS subspace

On the equivalence of representing radiances as PC scores or reconstructed radiances

Well accepted that IASI spectra can be represented by a small number of PC scores with only minor loss of information

The same can be said about the representation by a small number of reconstructed radiances.

In fact the cost function in both cases are identical if:

- **The two forward models are consistent and**
- **The channel subset is chosen such that the observation error covariance matrix in reconstructed radiance space is non-singular**

$$\text{(rank m)} \quad S_y \rightarrow EE^T S_y EE^T \quad \text{(rank s)}$$

EUMETSAT

Channel selection (reconstructed radiances)

Need to select *s* channels such that the corresponding sub-matrix of $EE^TS_yEE^T$ is non singular.

This is equivalent to selecting *s* linearly independent rows of the matrix square root $EE^TS_y^{1/2}$

- **Gram-Schmidt process**
- **Gaussian elimination**

Condition number of observation error covariance for reconstructed radiances can be minimized (heuristically) by choosing numerically large pivot elements.

MW + IR Linear regression as a priori

All sky, linear regression retrievals of temperature, humidity and ozone profiles from co-located IASI, AMSU and MHS measurements.

111 predictors:

- Surface height (km)
- Secant of satellite zenith angle
- Radiance in 14 AMSU channels (channel 7 excluded)
- Radiance in 5 MHS channels
- 30 leading IASI band 1 PC scores
- 30 leading IASI band 2 PC scores
- 30 leading IASI band 3 PC scores

277 + 1 dependent variables:

- Ta (K)
- Wa (K)
- Ts (K)
- Sp (hPa)
- T profile (K) at 91 model levels
- W profile (K) dew point temperature at 91 model levels
- O profile (K) “dew point temperature” (W formula) at 91 model levels
- Quality indicator [average absolute error of Ta, Ts, Wa and T[79],T[86],T[90], W[86],W[90]]

EUMETSAT

Regression based on real data

Training data based on co-located ECMWF analysis 953242 cases from 12 days, Nov-11 to Oct-12

36 regression classes for which individual regression coefficients are retrieved. The class is determined by the surface height, the radiance in AMSU channel 2 and 4, the radiance in MHS channel 2 and the first PC score in IASI band 2.

Regression coefficients to predict the absolute value of the retrieval error (surface). Estimation of the quality of each individual retrieval.

Class A: z==0 and AMSU_4 > 370 + 1.5*AMSU_2 (open sea)
Class B: z < 1 and AMSU_4 <= 370 + 1.5*AMSU_2 (land, sea ice, some clouds over sea)
Class C: z > 1 (high elevation)

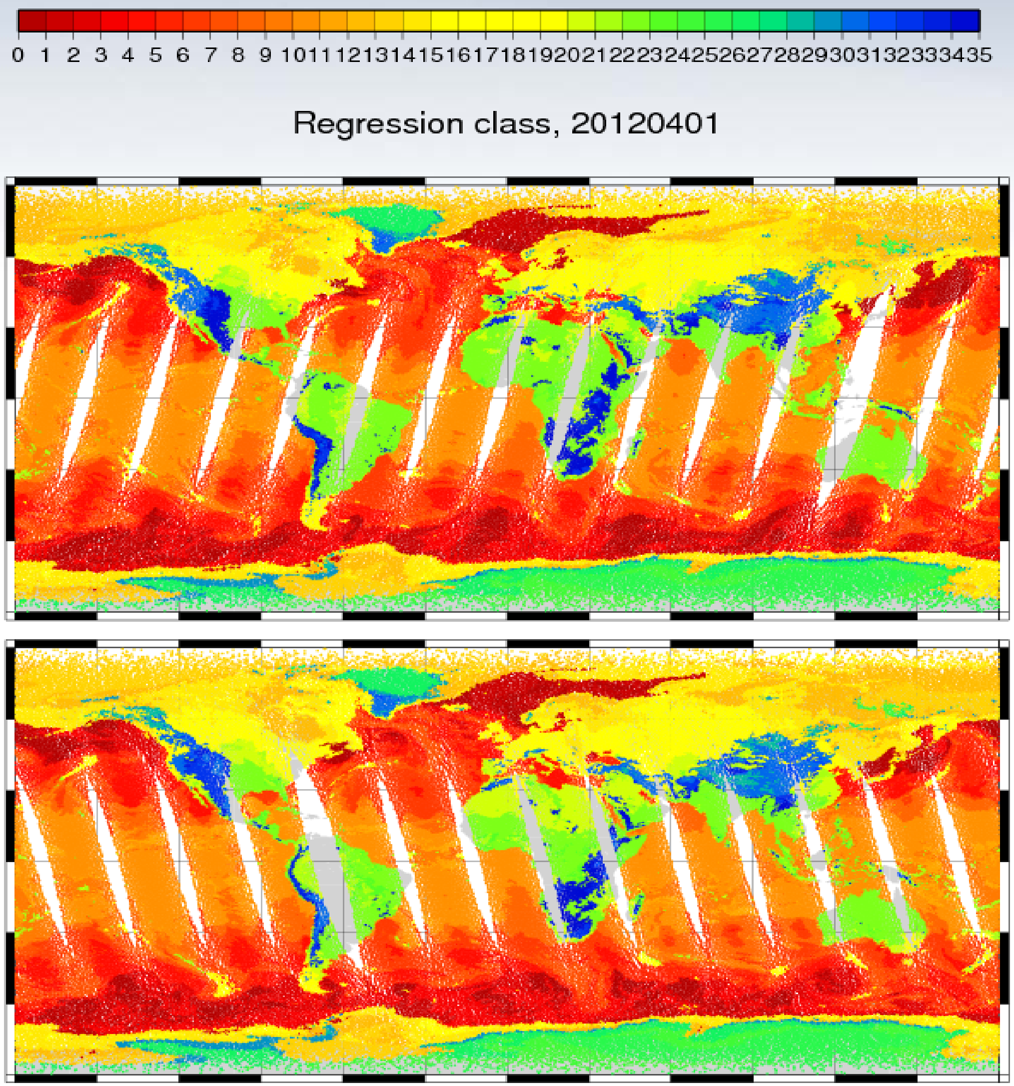
Each class further subdivided in 3 according to AMSU_4

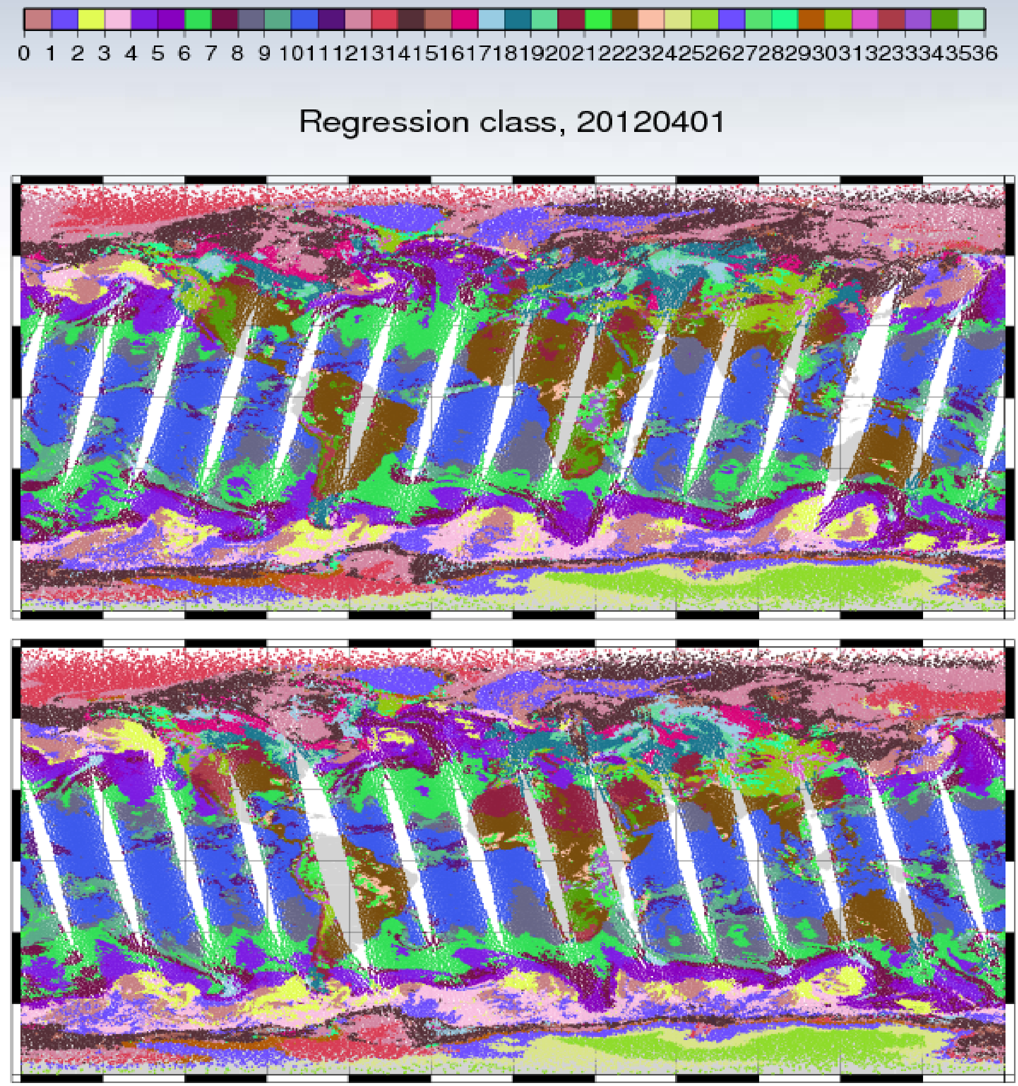
Each class further subdivided in 2 according to MHS_2

Each class further subdivided in 2 according to B2PC1

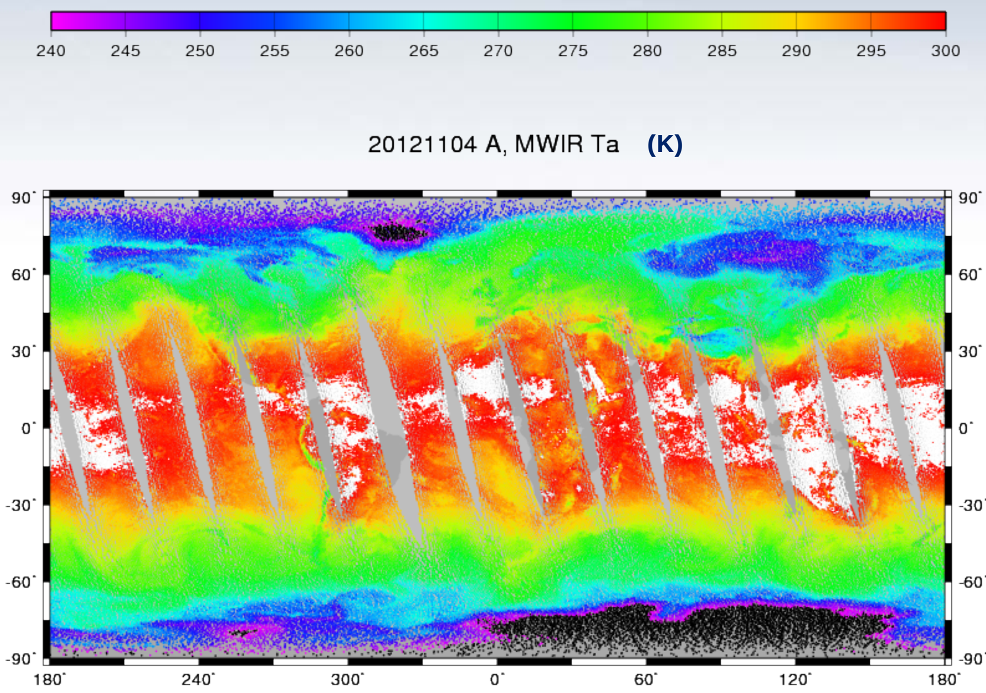
EUMETSAT

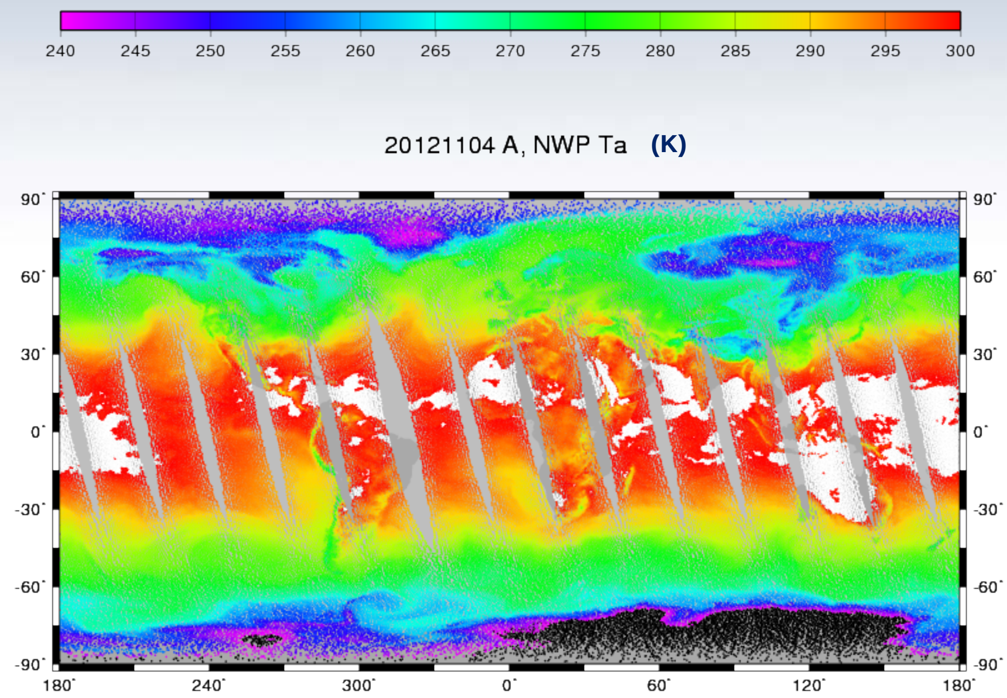
Piece-wise linear regression (36 regression classes)





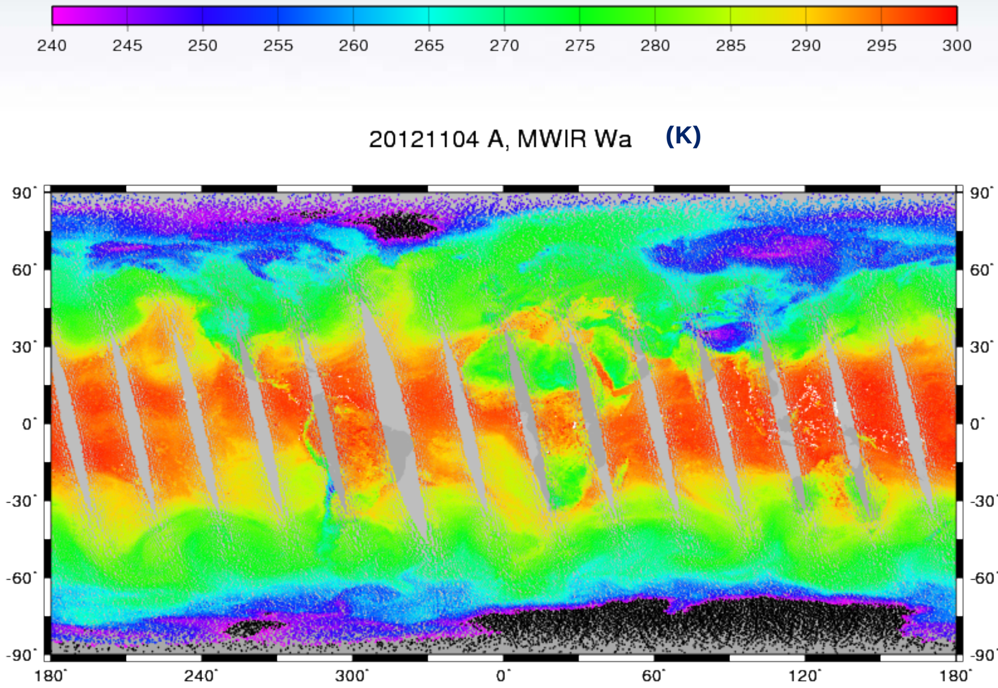
Surface air temperature

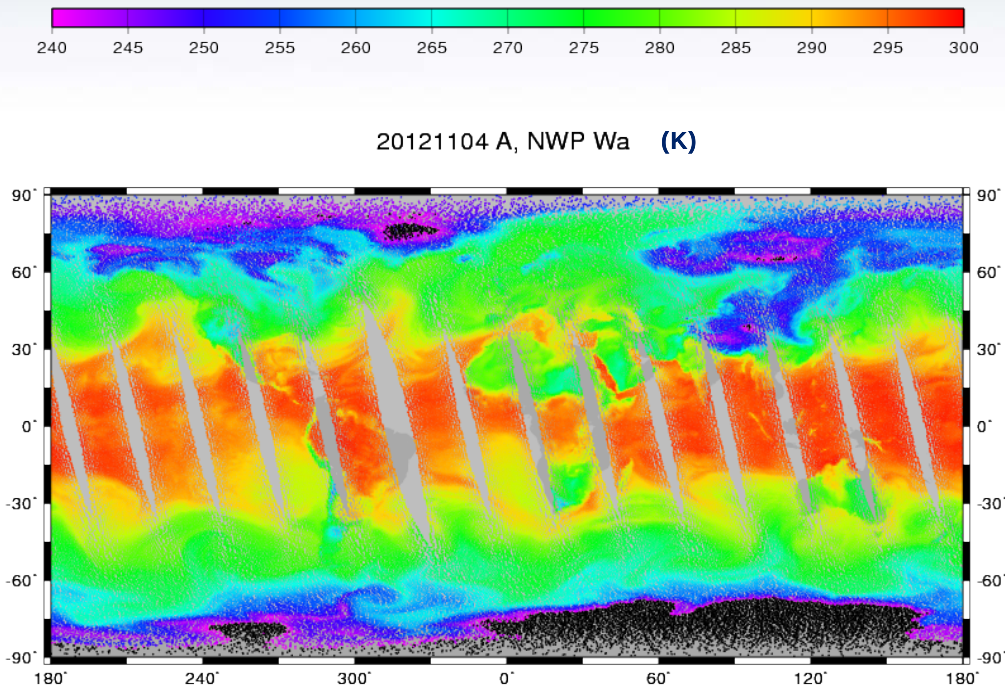




EUMETSAT

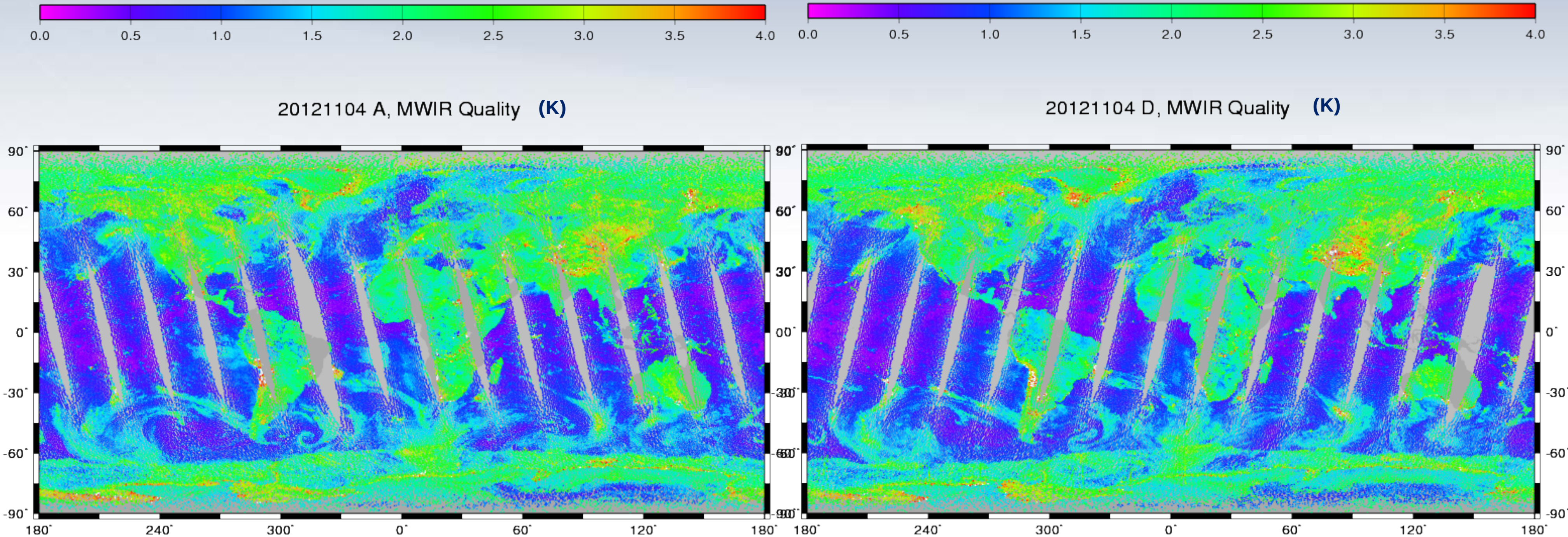
Surface air dew point temperature





EUMETSAT

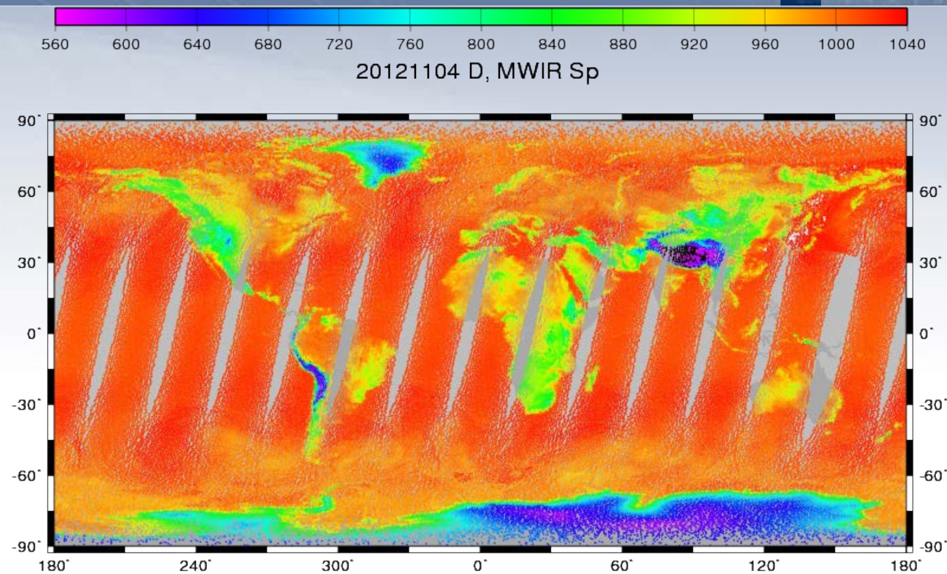
Quality indicator

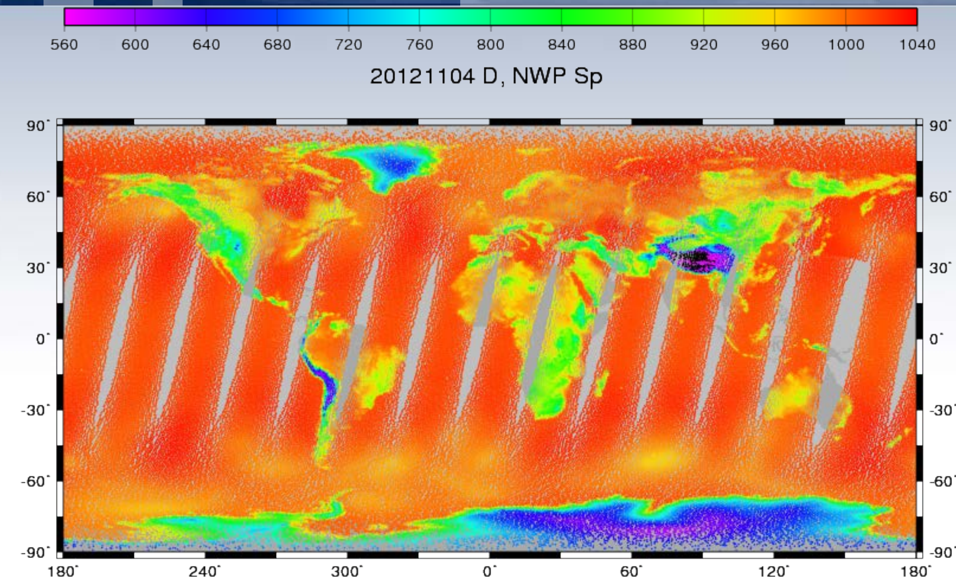


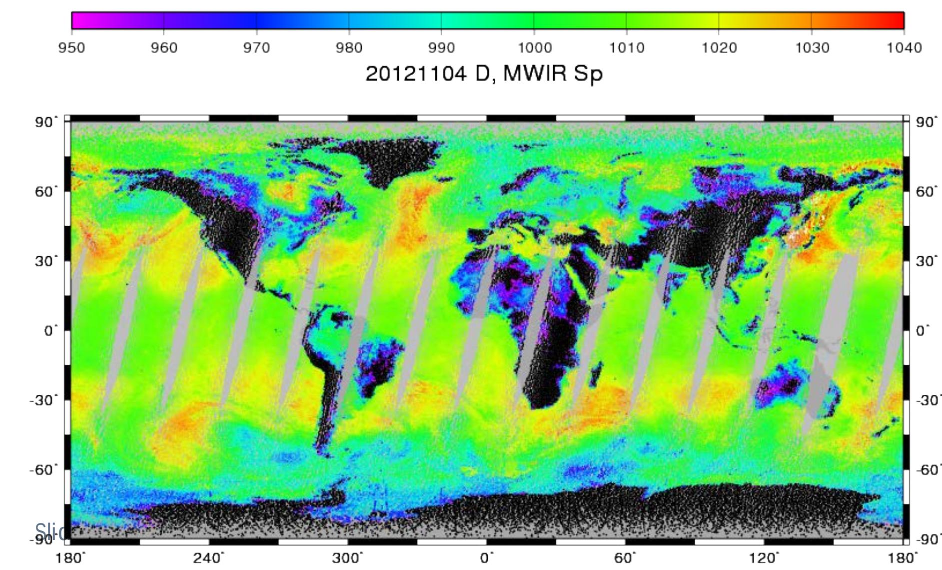
Estimate of error obtained by regression against absolute value of error

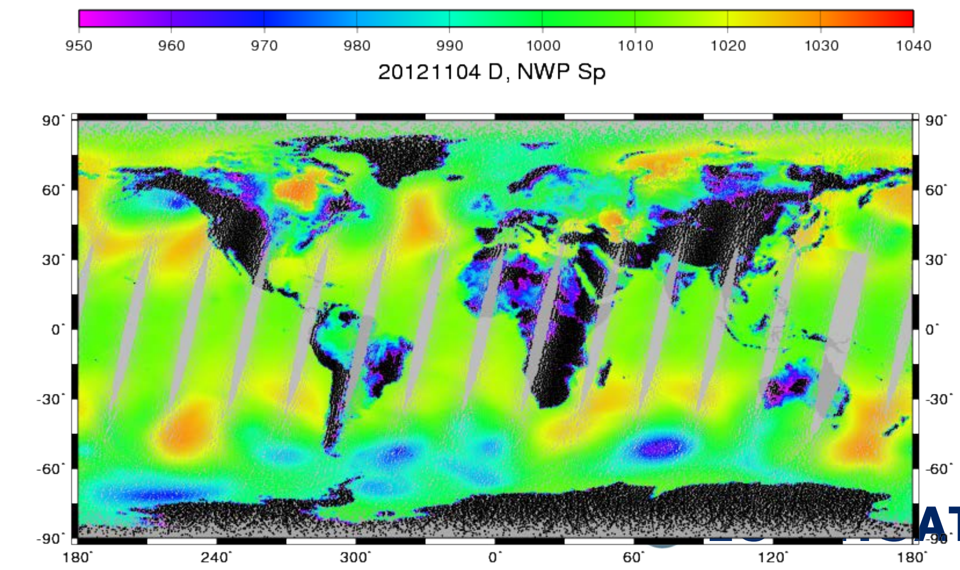
EUMETSAT

Surface pressure

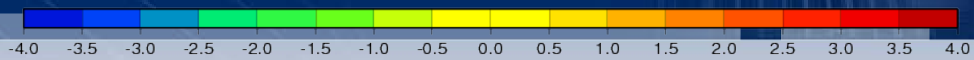


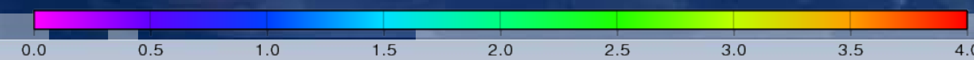




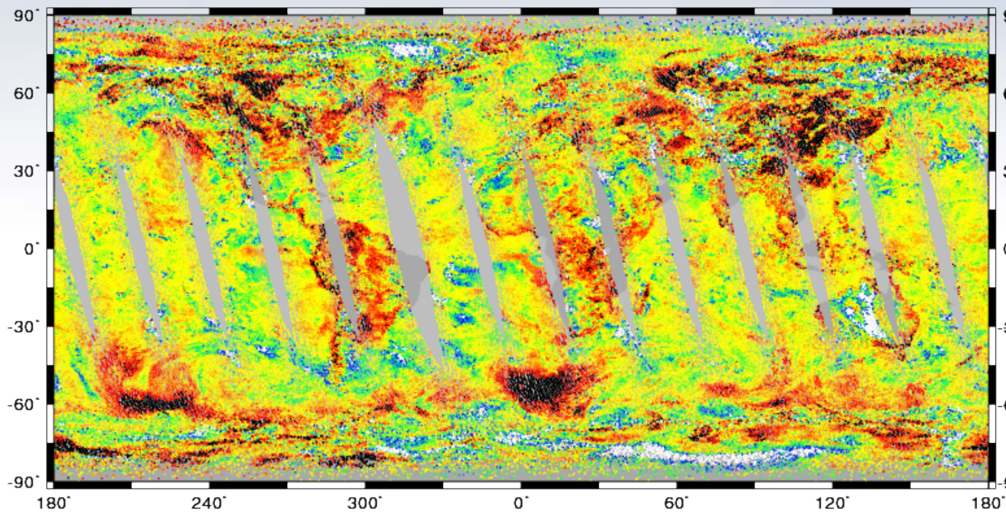


Difference plots

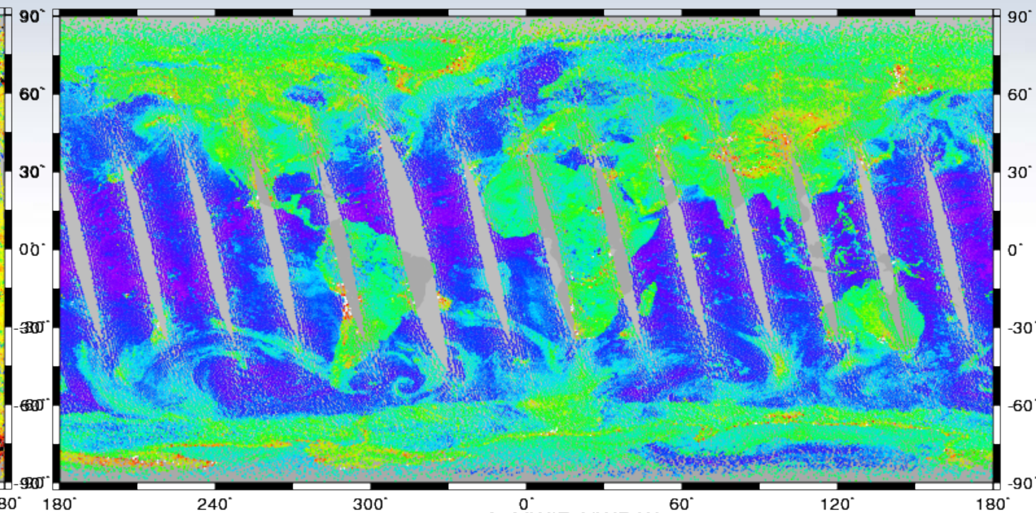




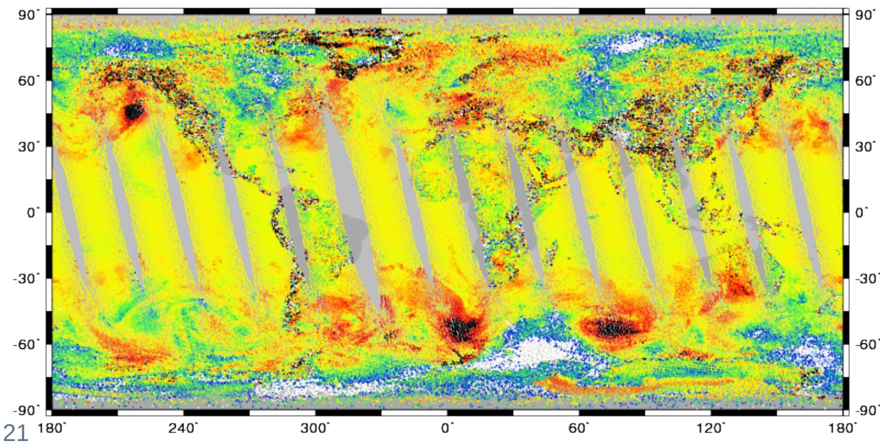
20121104 A, MWIR-NWP Ta



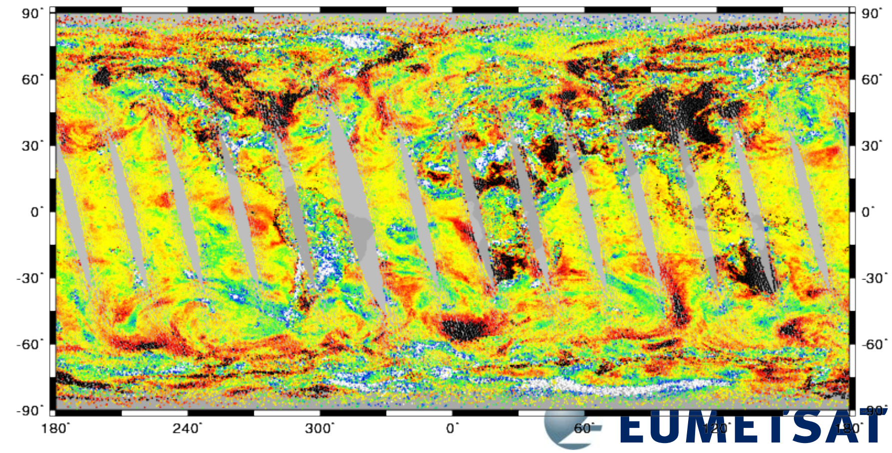
20121104 A, MWIR Quality



20121104 A, MWIR-NWP Sp



20121104 A, MWIR-NWP Wa



Bias correction by linear regression

Tuned against 12 days co-located ECMWF analysis, clear scenes only

41 Predictors:

- **Surface height (km)**
- **Secant of satellite zenith angle**
- **Radiance in 14 AMSU channels (channel 7 excluded)**
- **Radiance in 5 MHS channels**
- **10 leading IASI band 1 PC scores**
- **10 leading IASI band 2 PC scores**

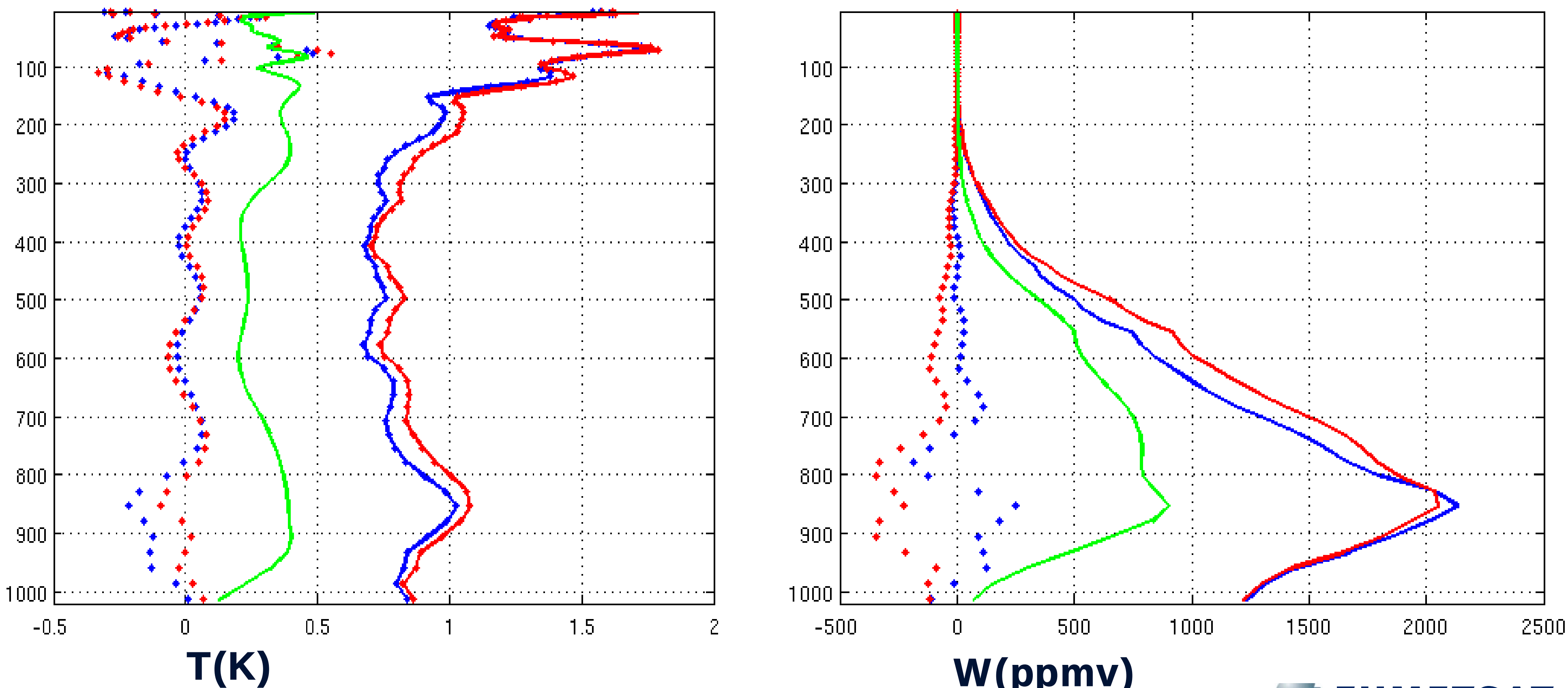
Dependent variables: $IASI_{OBS} - RTTOV_{CALC}(ECMWF)$ in the 139 selected channels

Bias correction makes $COV(IASI_{OBS} - RTTOV_{CALC}(ECMWF))$ decrease!

EUMETSAT

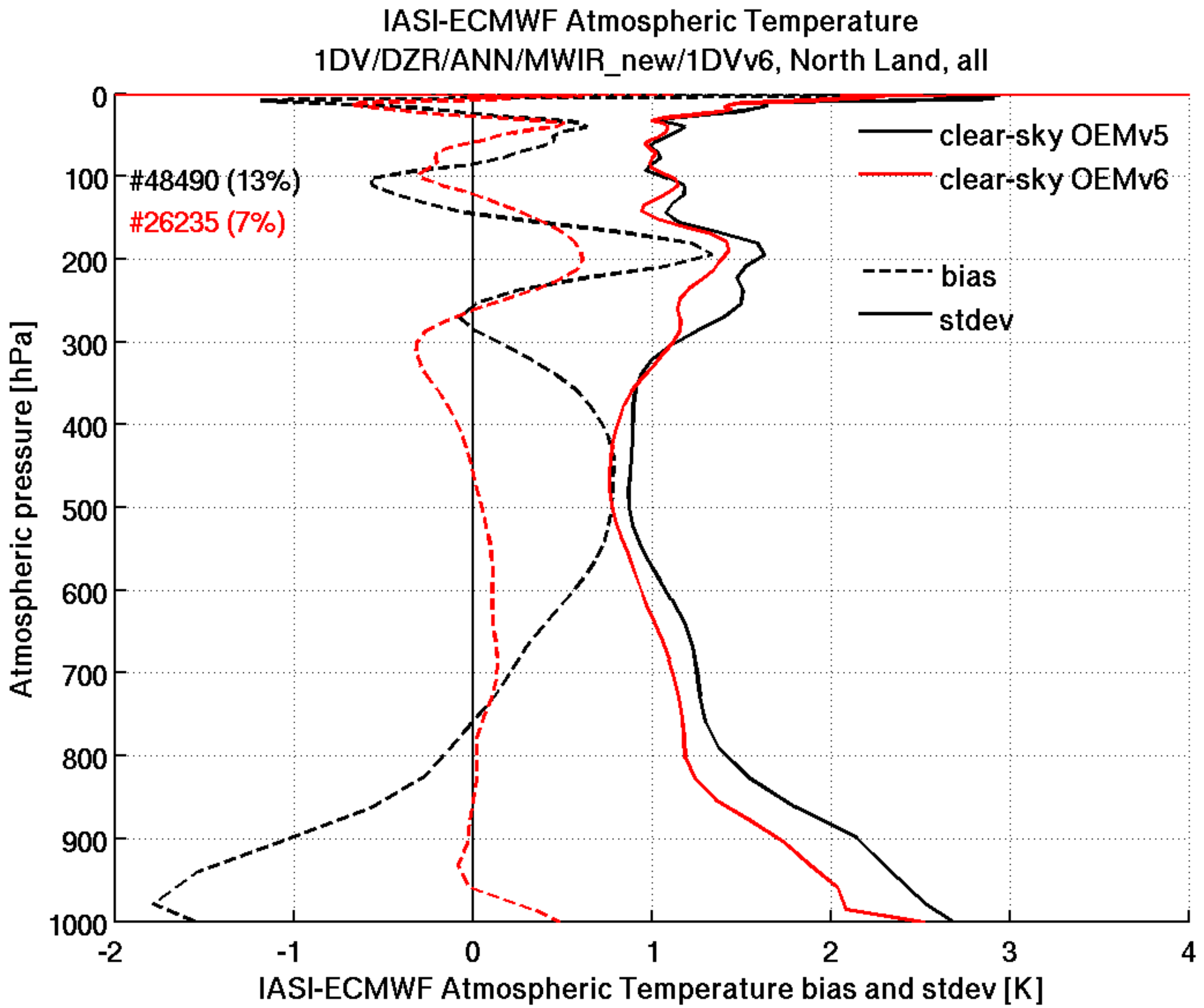
Retrieval results

20120401 sea only. Linear regression, 1DVar, 1DVar - linear regression



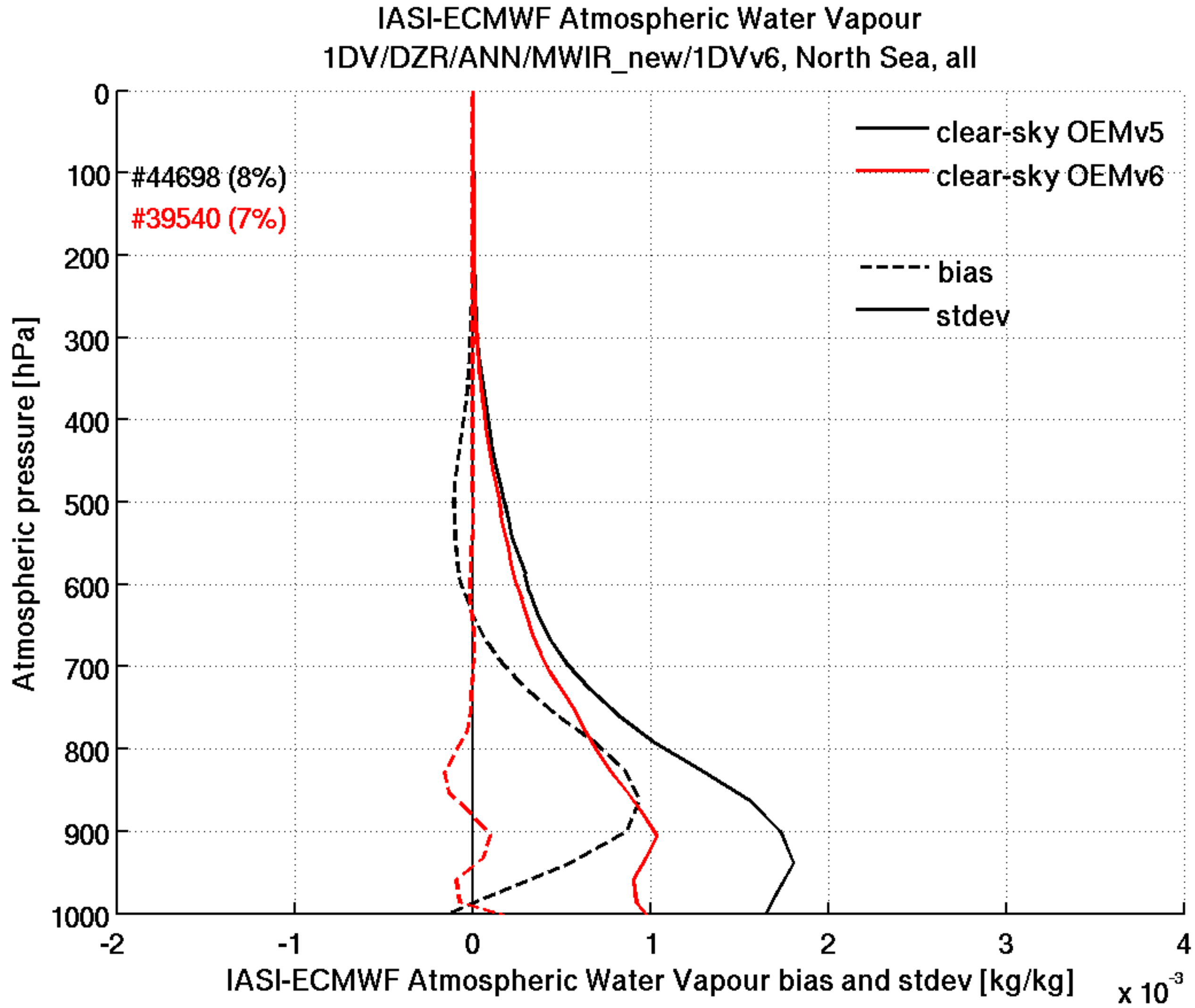
EUMETSAT

Temperature vs v5



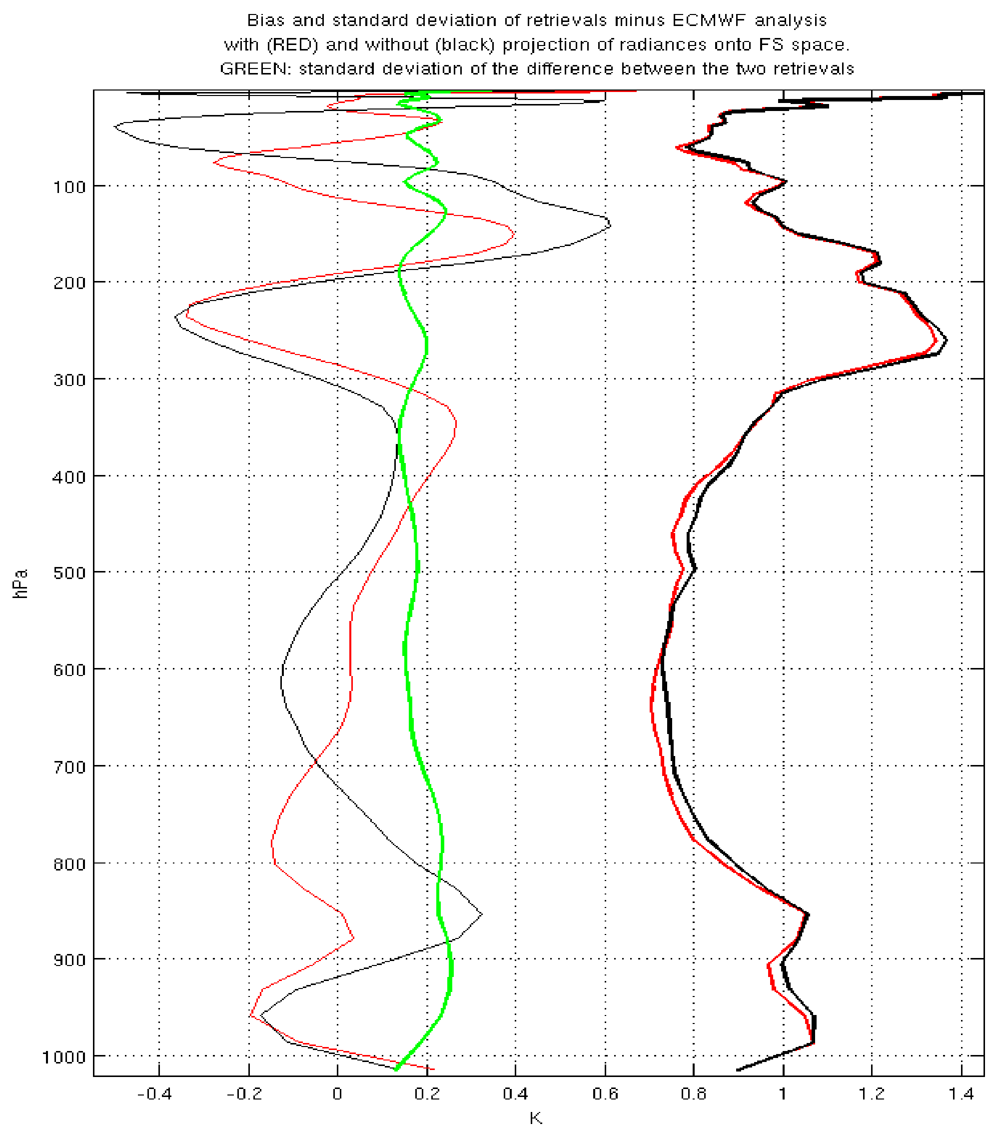
EUMETSAT

Water vapour vs v5



EUMETSAT

Impact of filtering onto the FS space



Radiances projected onto signal space

Radiances projected onto the FS space

Standard deviation of the difference of the retrievals with and without projection onto FS space

Statistics based on 11822 cases over sea, + / - 60° latitude on the 2012.10.01

EUMETSAT

Wrap-up / Final slide / The end

Regression profiles very good even in cloudy scenes ➔ can be used for cloud detection and characterisation

Fast 1DVAR:

- Usually converges in 1 or 2 pure Newton iterations
- Forward model only invoked for 139 channels

1DVAR improves the retrieved profiles even further

Configuration tuned against ECMWF, but no forecasts used in retrieval step

Transformation of signal space basis used for operational dissemination of PC scores would give users an easy way to suppress instrument artefacts from the measurements

Improvements mainly due to incorporation of microwave information in the a-priori, but also additional improvement from projecting radiances onto the FS space

EUMETSAT