Super Channel Selection for IASI Retrievals

or

How to Deal with **8461** Spectral Samples

Peter Schlüssel
EUMETSAT

IASI Level 2 Product Generation
High Level Break-Down Structure

IASI Level 1c	AVHRR Cloud Mask and S/CTT	Configurable Databases

ATOVS Level 2	AMSU-A Level 1	MHS Level 1	NWP Forecast

Pre-Processing

Cloud Processing

Geophysical Parameters Retrieval

Monitoring Information	Level 2 Product	Quality Information

EUMETSAT

Geophysical Parameters Retrieval

- The geophysical parameters retrieval consists of two steps:

  – First retrieval using EOF regression for temperature and water vapour profiles, ozone columns in deep layers, surface temperature and emissivity, ANN retrieval for trace gases (alternatively also for temperature and water vapour profiles)

  – Second, iterative retrieval using sub-set of IASI spectral samples combined to "super channels"

- Constraint of near real time processing restricts the number of super channels in the iterative retrieval to about 500

Final, Iterative Retrieval

- Simultaneous iterative retrieval, seeking maximum probability solution for minimisation of cost function by Marquardt-Levenberg method

- Initialisation with results from first retrieval

- Other choices of initialisation may be selected, depending on configuration setting and availability (e.g. NWP forecast, climatology, ATOVS Level 2 product)

- Background state vector from climatology, ATOVS Level 2 product, adjacent retrieval, or NWP forecast, depending on configuration and availability

- State vector to be iterated depends on cloud conditions and configuration setting (clear, cloudy, variational cloud clearing)

EUMETSAT

Maximum Probability Solution Formulation of Cost Function

- Minimisation of cost function

$$J = (\mathbf{y(x)} - \mathbf{y^m}) \cdot \mathbf{E}^{-1} \cdot (\mathbf{y(x)} - \mathbf{y^m})^{\mathbf{T}} + (\mathbf{x} - \mathbf{x^b}) \cdot \mathbf{C}^{-1} \cdot (\mathbf{x} - \mathbf{x^b})^{\mathbf{T}}$$

- $\mathbf{x}$ = atmospheric state vector as calculated iteratively

- $\mathbf{x^b}$ = atmospheric state vector from background field

- $\mathbf{C}$ = error covariance matrix associated with the background

- $\mathbf{y^m}$ = measurement vector

- $\mathbf{E}$ = combined measurement and forward model error covariance matrix

- $\mathbf{y(x)}$ = forward model operator for given state vector $\mathbf{x}$

- $\mathbf{K_x}$ = partial derivatives of $\mathbf{y}$ with respect to the elements of $\mathbf{x}$ (Jacobians)

Super-Channel Composition

- Search for highly correlated radiances among all IASI spectral samples (noise-normalised radiances) and collect them in clusters
    - No need to have adjacent spectral samples
    - Correlation must be pre-determined at certain level: determines the number of clusters

- Determine a lead channel in each cluster: Radiance and Jacobians will only be calculated for lead channel radiances $y_L(\mathbf{x})$

- Represent the measured radiance $y_L{}^m$ of a cluster by a weighted average of all radiance samples which are members of the same cluster
    - Weights consist of regression coefficients, taking into account the correlation and the noise of the respective sample
    - Errors include measurement and regression errors

$$y_L^m = \frac{1}{N} \sum_{i=1}^{N} a_i + b_i \, y_i^m$$

EUMETSAT

IASI FM-2 Noise (self-apodised)
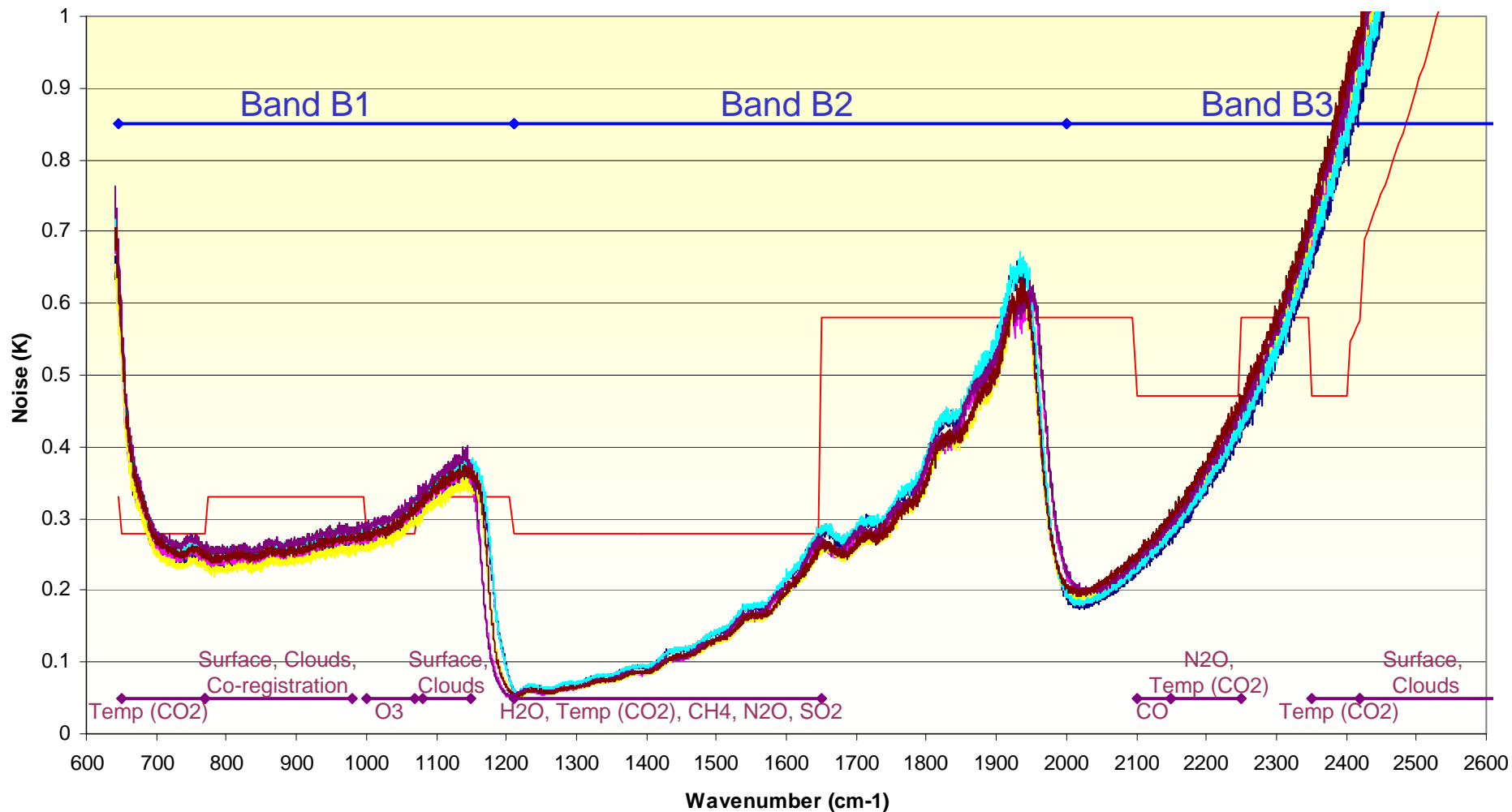
Data Set for Radiative Transfer Simulations

- Sub-sampled ECMWF ERA-40 re-analysis: 13495 globally distributed atmospheric situations (Chevallier, 2002)

- Perturbation of temperature profiles with randomly distributed 1K/1km double-Dirac dipoles

- Composition of land surface types:
  - Random selection of up to three types and corresponding fractions
  - Emissivity spectra: average weighted according to fractional coverage

- Sea surface emissivity from two-stream model accounting for wind speed, direction, and fetch

- Trace gas profiles: Perturbed climatological profiles (random variation between 50 and 150%)

- Aerosol particles: Random selection of aerosol types and perturbed climatological profiles (random variation between 0 and 200%)

EUMETSAT

Radiative Transfer Simulations and Super-Channel Clustering

- For the entire set of 53980 situations RTIASI-5 simulations have been carried out, assuming a random selection of possible scan angles

- The resulting radiance spectra (apodised) have been normalised with the corresponding noise

- A correlation analysis is done between spectral samples
  - All samples with a correlation higher than a threshold are retained in the respective cluster
  - The assumed thresholds vary between 0.95 and 0.999

- The lead sample is regressed against all other member samples of a cluster to obtain weighting coefficients and error estimates
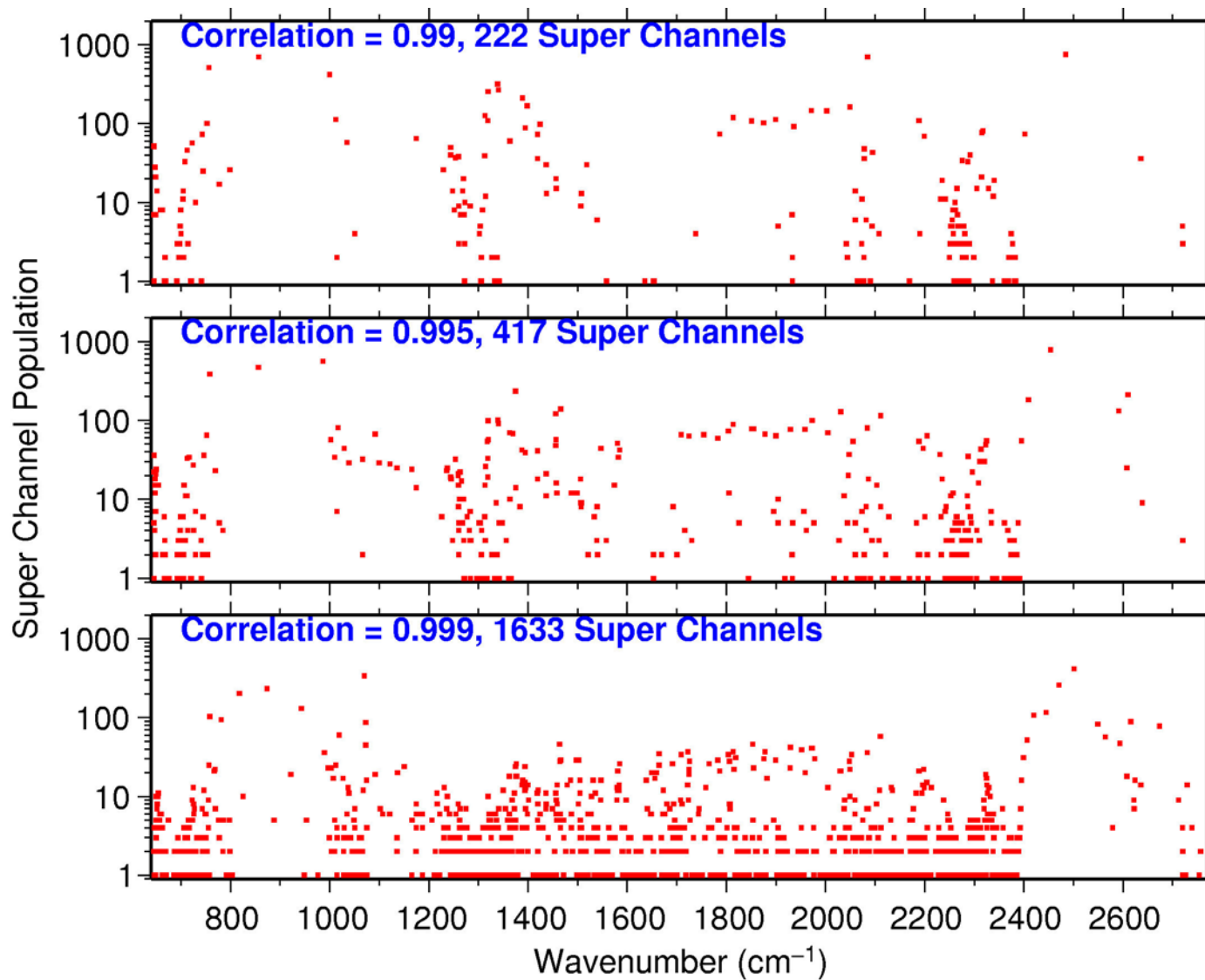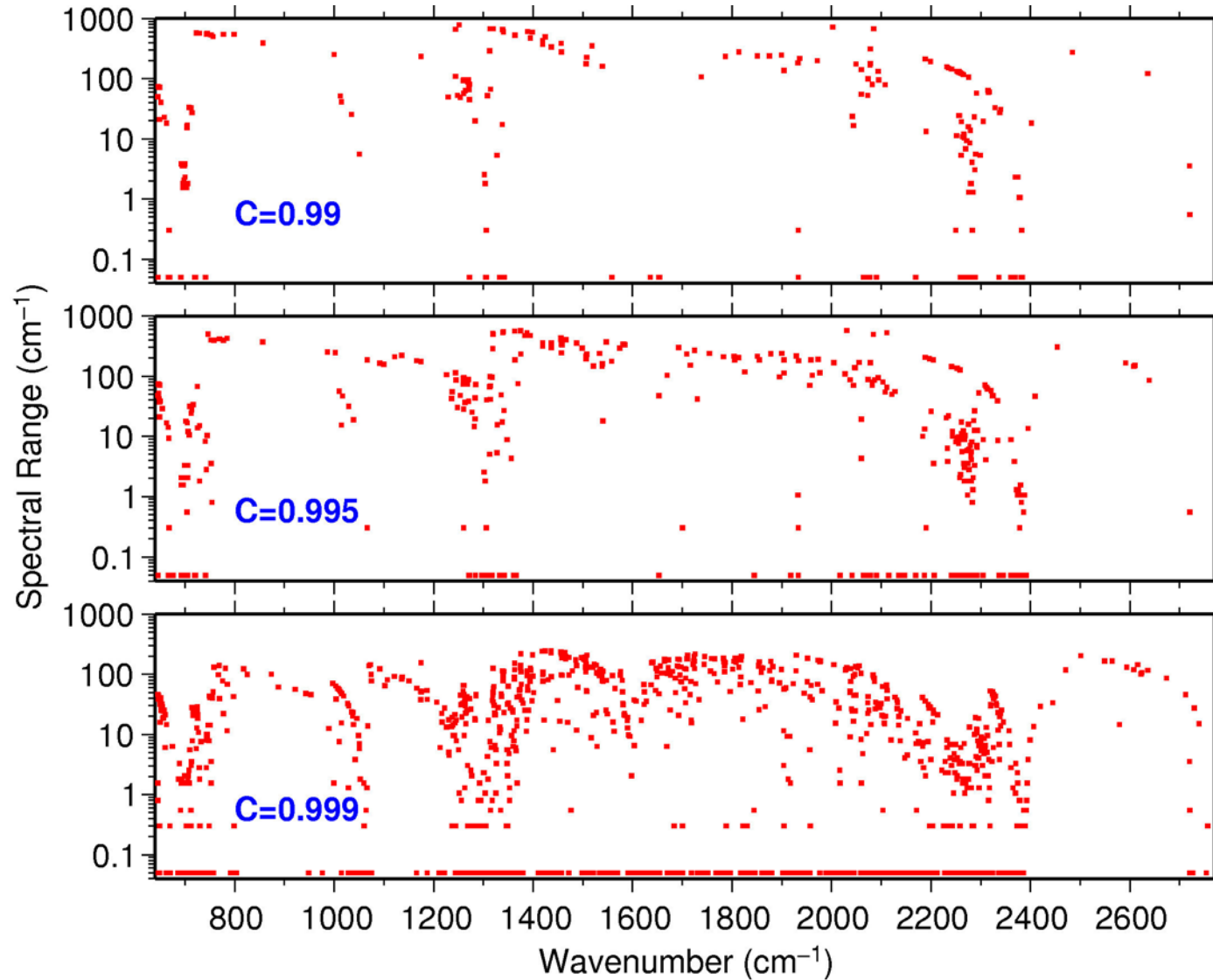
EUMETSAT

Super-Channel Properties

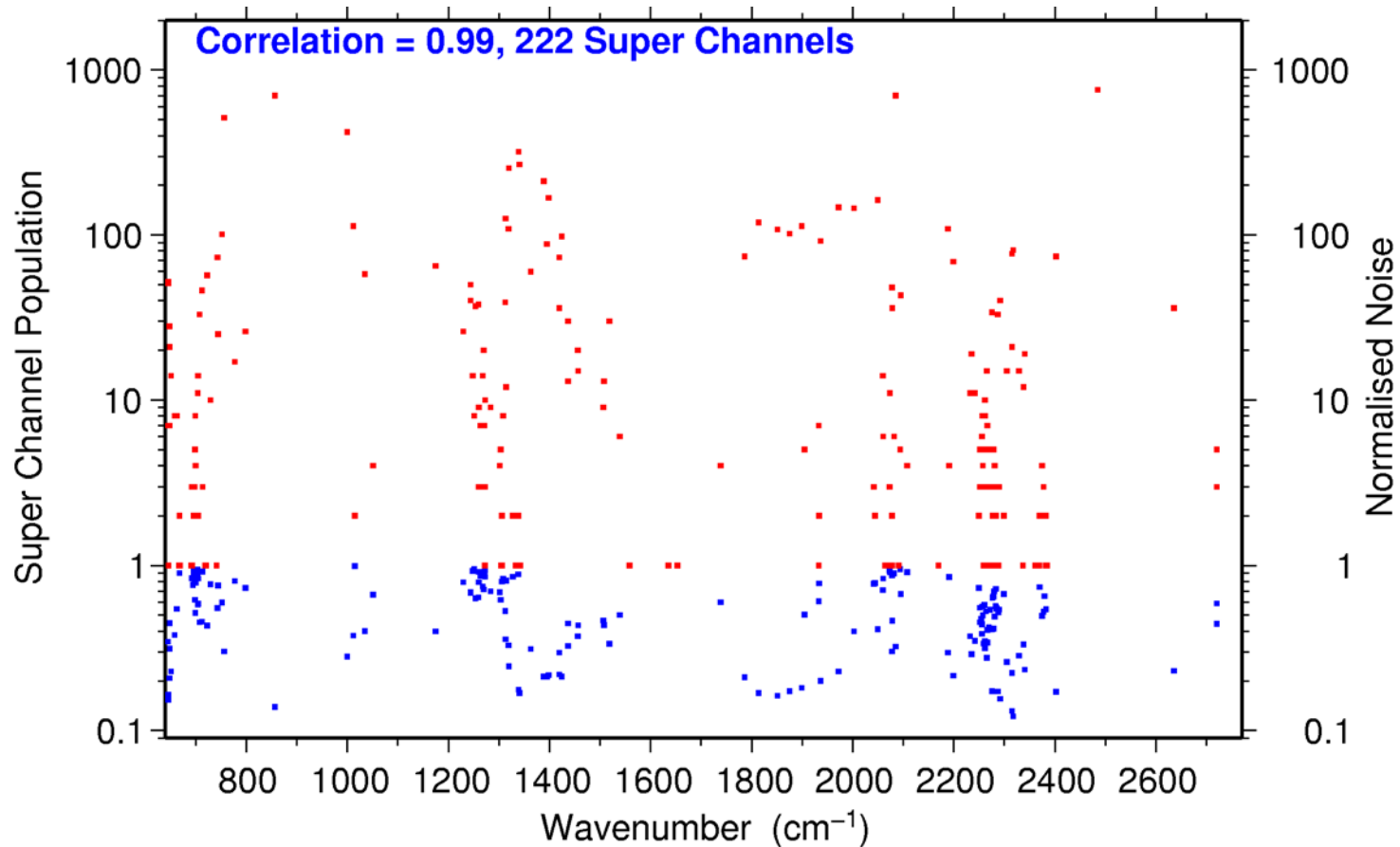| Correlation	Number of Super-Channel Clusters
0.95	47
0.98	113
0.99	222
0.995	417
0.999	1633

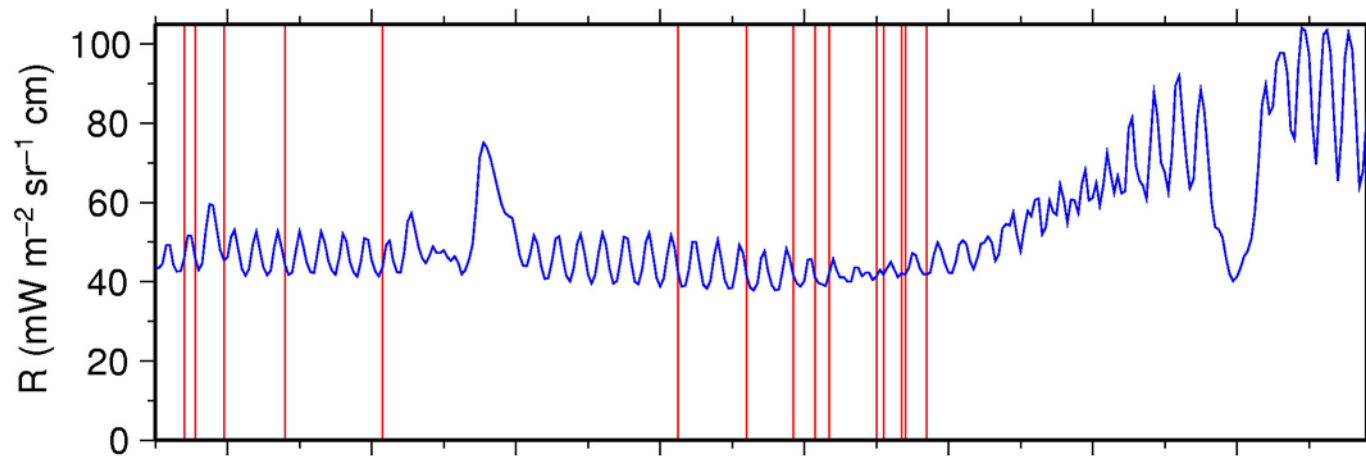# Population of Super-Channel Clusters

Spectral Range Covered by Super-Channel Clusters
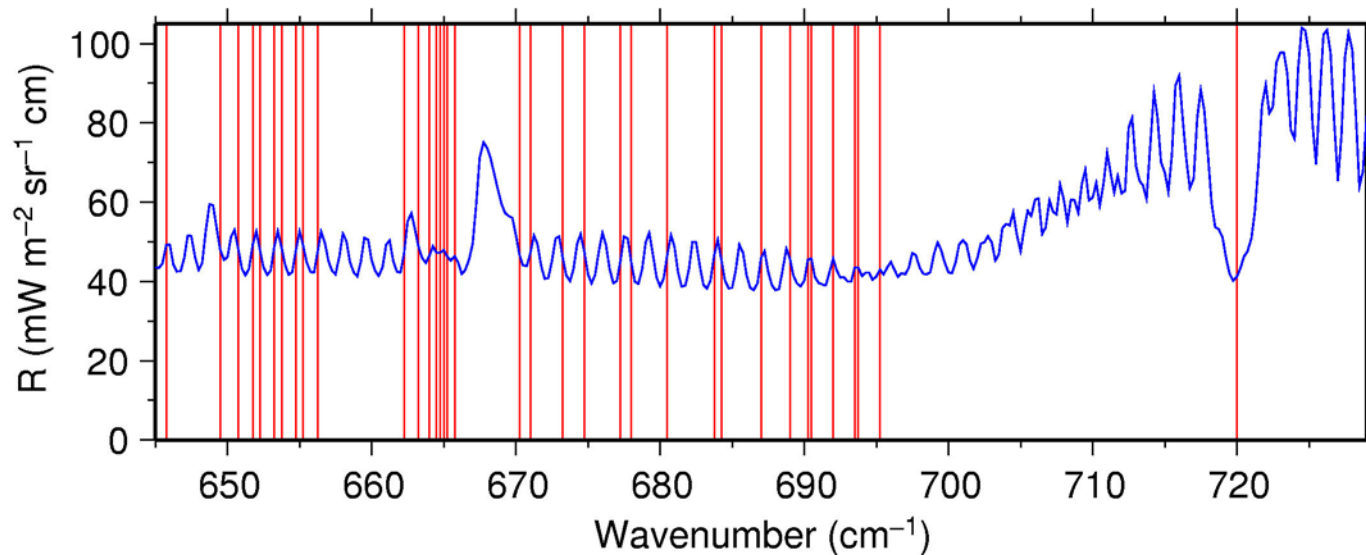
Population of Super-Channel Clusters and Noise Reduction

Examples of Super-Channel Clusters

Correlation: 0.995

Cluster 6:
15 samples
Lead: 647 cm^{-1}

Cluster 3:
36 samples
Lead: 645.75 cm^{-1}

Conclusion

- IASI spectral samples can be combined to super channel clusters to reduce the number of measurements entering a variational retrieval

- A "lead channel" is selected to represent the cluster in the radiative transfer simulations of the variational retrieval

- The number of clusters can range from as low as 47 to more than 1633, assuming sample correlations between 0.95 and 0.999

- Very high correlation implies that many clusters are populated with single samples only, which is desirable when information about minor contributions is needed, e.g. trace gases

- The spectral range covered by single clusters is not limited to narrow spectral regions, but can cover a major part of the IASI spectrum

- For a modest number of 222 super channel clusters the noise is reduced by more than a factor 5 for many clusters, compared to single samples

EUMETSAT