AOEI
Towards an Adaptive scene-dependent Optimal Estimation for IASI

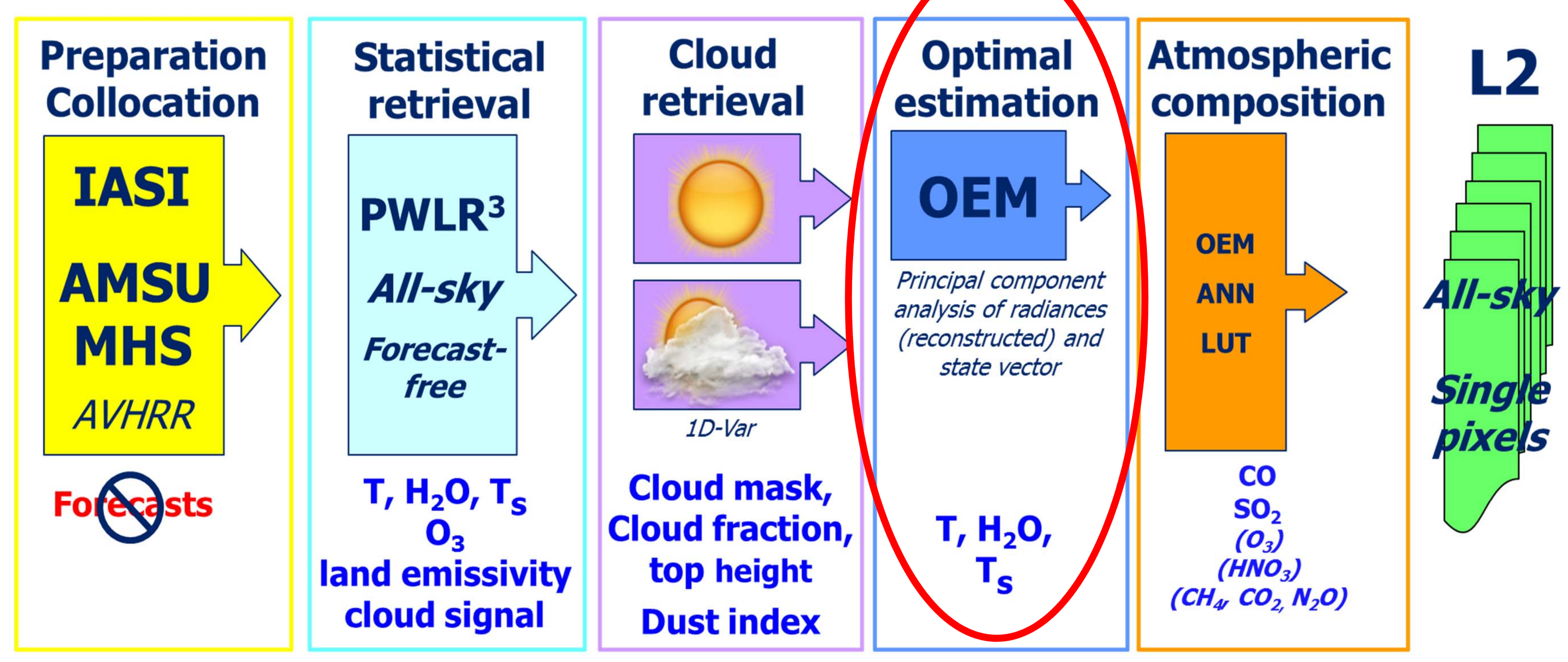
T. August and T. Hultberg, EUMETSAT

Optimal estimation, e.g. as formulated in Rodgers 2000, is popular inverse method for atmospheric sounding. It builds on Baye's theorem and requires accurate observation and a priori errors, assumed to be Gaussian. These errors are usually evaluated off-line of the retrieval and applied globally for a given observing system. In the operational IASI L2 processor at EUMETSAT, the observation error has been defined statically as the covariance of the difference between observed and calculated radiances, using the first guess state vector and RTTOV as a forward model. Two distinct observation error matrices are defined for land and maritime scenes since the release of IASI L2 v6.3, with the aim to account for usually more accurate knowledge of surface emissivity and temperature over oceans than over continental surfaces. This lead also to more accurate humidity retrieval in the low troposphere.

In this work, we study the feasibility and advantages of dynamically applying scene-dependent observation errors to each individual retrieval instance from IASI.

First, distinct observation classes are established off-line for IASI by application of unsupervised K-mean clustering to a representative set of observations, based on their leading principal components. A different observation error matrix is computed as the covariance of OBS-CALC in each of these classes. We analyse the geographical distribution of the observation classes, their seasonal variations and the amplitude of the corresponding observation errors.

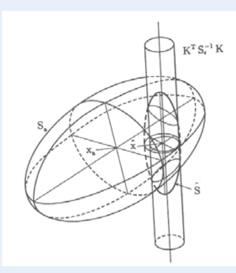
Second, to compensate for systematic differences between observations and forward modelling, a scene-dependent bias correction is regressed as a function of IASI observations and viewing geometry (e.g. satellite zenith angle, elevation).



Bayesian approach and assumptions

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{x})P(\mathbf{x})}{P(\mathbf{y})}. \qquad (2.19)$$

Rodgers, 2000



Linear Problems with Measurement Error 25

Gaussian statistics are usually a good approximation for the errors in real measurements, so express $P(\mathbf{y}|\mathbf{x})$ as

$$-2\ln P(\mathbf{y}|\mathbf{x}) = (\mathbf{y}-\mathbf{Kx})^T\mathbf{S}_\epsilon^{-1}(\mathbf{y}-\mathbf{Kx}) + c_1, \qquad (2.21)$$

where c_1 is a constant and $\mathbf{S}_\epsilon$ is the measurement error covariance. Less realistic, but convenient, is to describe prior knowledge of $\mathbf{x}$ by a Gaussian *pdf*:

$$-2\ln P(\mathbf{x}) = (\mathbf{x}-\mathbf{x}_a)^T\mathbf{S}_a^{-1}(\mathbf{x}-\mathbf{x}_a) + c_2, \qquad (2.22)$$

where $\mathbf{x}_a$ is the *a priori* value of $\mathbf{x}$, and $\mathbf{S}_a$ is the associated covariance matrix

$$\mathbf{S}_a = \mathcal{E}\{(\mathbf{x}-\mathbf{x}_a)(\mathbf{x}-\mathbf{x}_a)^T\}. \qquad (2.23)$$

Substituting Eqs. (2.21) and (2.22) in Eq. (2.19) we obtain for the posterior *pdf*:

$$-2\ln P(\mathbf{x}|\mathbf{y}) = (\mathbf{y}-\mathbf{Kx})^T\mathbf{S}_\epsilon^{-1}(\mathbf{y}-\mathbf{Kx}) + (\mathbf{x}-\mathbf{x}_a)^T\mathbf{S}_a^{-1}(\mathbf{x}-\mathbf{x}_a) + c_3, \qquad (2.24)$$

Error assumptions

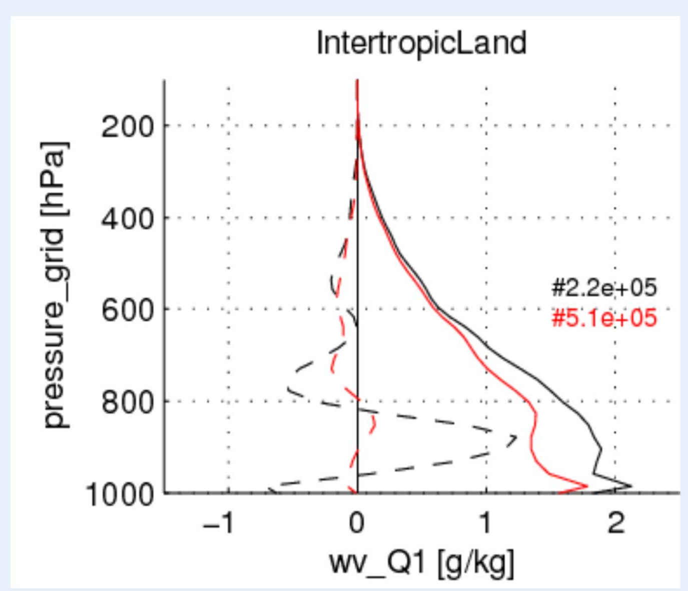
! Observation error contains more than the instrument noise: forward modelling errors
! If static and global: not Gaussian and not most suited to a given situation
? Bias OBS *vs* CALC ? Varying with viewing angle and scene
! Prior error usually not Gaussian if *a priori* is static global
! Errors (amplitudes and correlation) must be representative

Status of OEM configuration in the EUMETSAT IASI L2 processor

$$J = \underbrace{(x - x_a)^T S_x^{-1}(x - x_a)}_{} + \underbrace{(F(x) - y)^T S_y^{-1}(F(x) - y)}_{}$$

Current status (since IASI L2 v6, 2014)

- Scene-dependent x_a - from statistical method PWLR³ *(1)*
- Error more guaranteed to be Gaussian
- Background closer to true profile
- ✓More precise final retrievals



Current status (since IASI L2 v6, 2014)

- Reconstructed radiances in channels selected to contain as much information as in as many PCS *(2)*
- Bias correction = regression(sat. zenith, 1st PCS of Bands 1 & 2)
- S_y = covariance(OBS-CALC) over large clear-sky pixels sample
- S_y initially based on ocean pixels only
- Since v6.2 (2 June 2016), a different S_y for ocean and land pixels
- ✓ Larger yield and more accurate humidity retrievals with dedicated errors over land (larger e.g. due surface emissivity uncertainties) - see figure to the left

(1) *Hultberg et August, "Evolutions and self-organisation of the piece-wise linear regression for IASI", ITSC-XX 2015*
(2) *Hultberg et August, ITSC-18, "Hyperspectral retrieval and subspaces"*

Fig: IASI L2 q vs ECMWF analyses

- *|lat| < 30°*
- *Land pixels*
- *April-May 2016*
- ***v6.1** - **small** obs. error evaluated over **ocean***
- ***v6.2** - **larger** obs. error evaluated over **land***

Study objectives and principles

Explore benefits and operational potential of scene-dependent OEM configuration for the retrieval of T, q, O_3 and Ts.

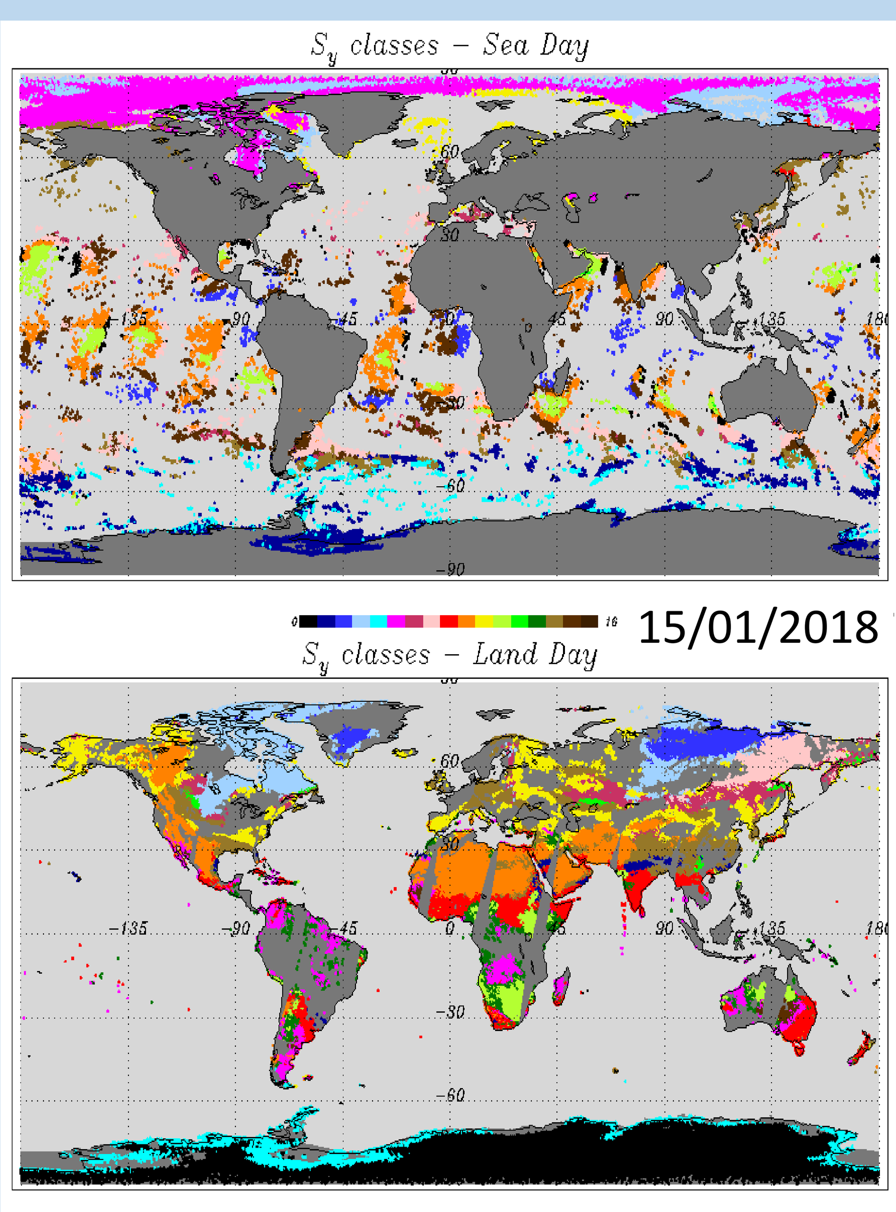
Offline statistical "training"

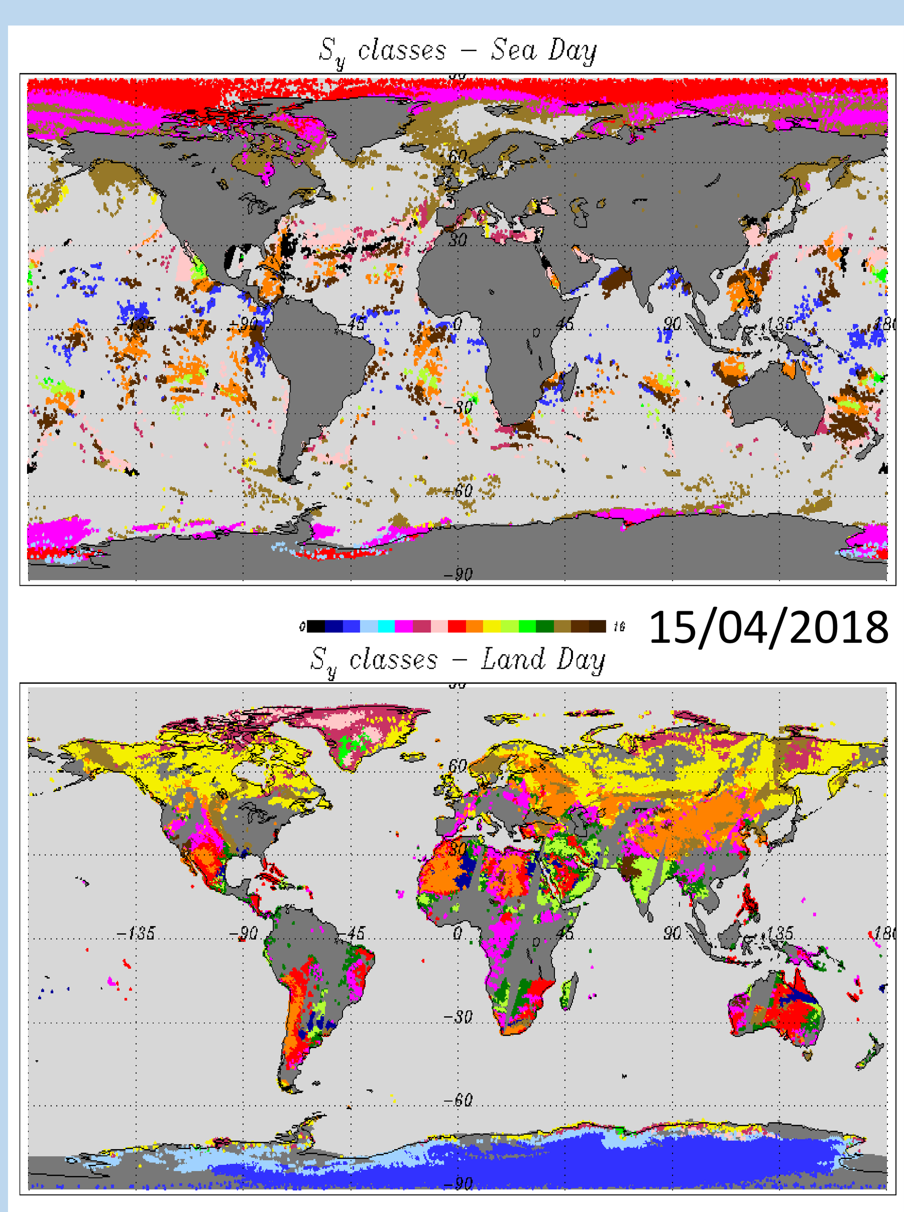
- Large set of OBServed and CALCulated radiances
  - using RTTOV and the state vector retrieved with the statistical method PWLR³
  - in cloud-free pixels as per IASI L2 cloud mask
- Observation space partitioned with K-mean clustering
- For each k-cluster
  - S_y^k= covariance(OBS-CALC)
  - Bias OBS - CALC = $\boldsymbol{regress^k}$ (sat. zenith $\hat{z}$, PCS)

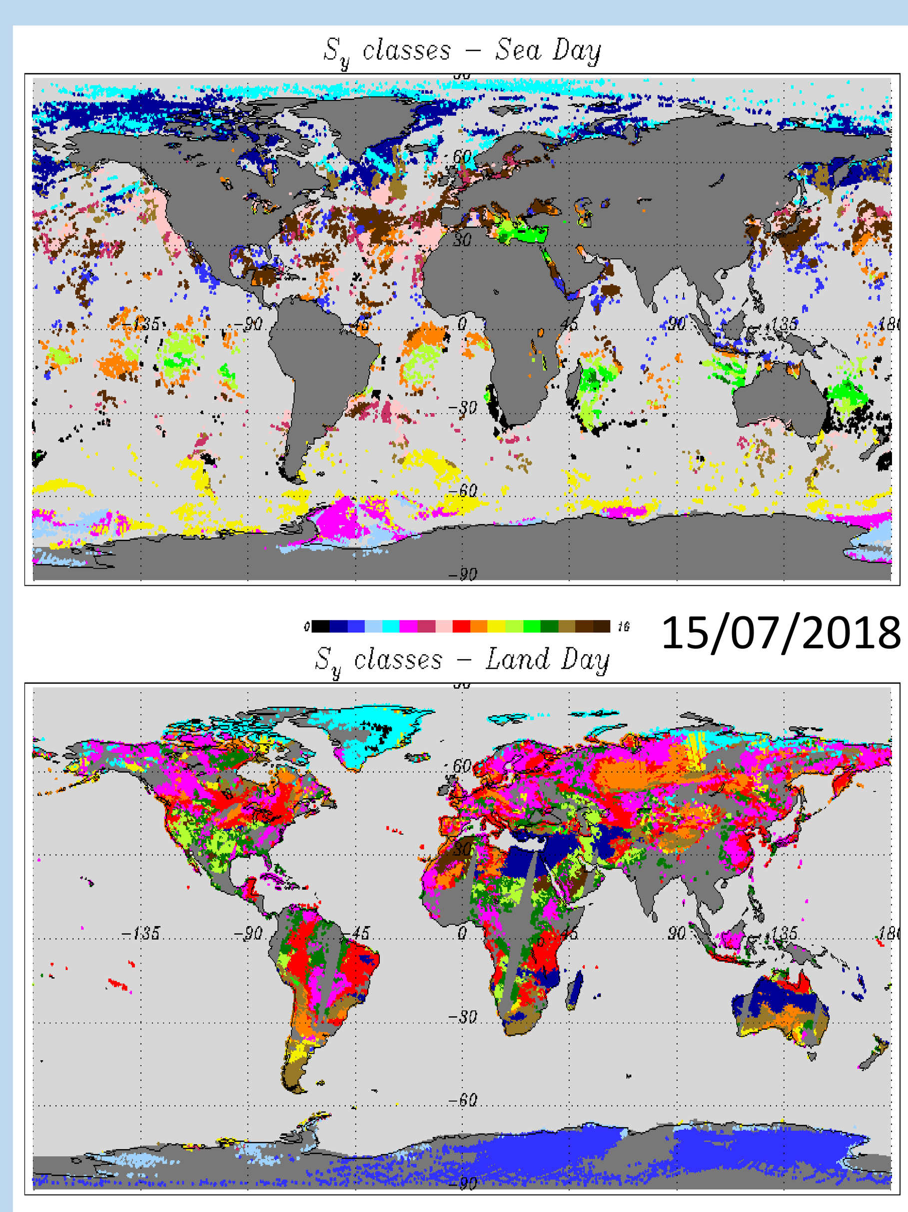
Online application – retrieval stage

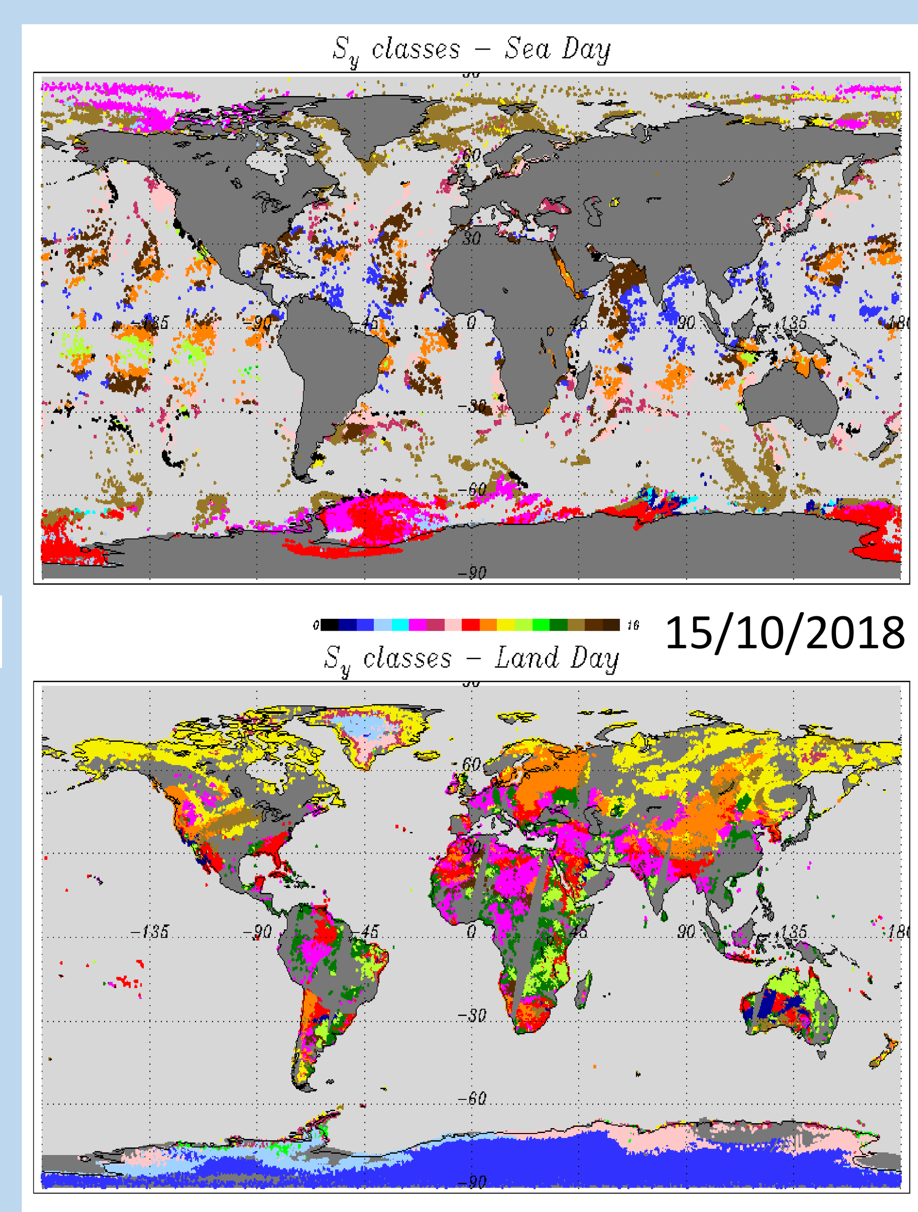
- Scene classification within the K clusters
- Compute scene-dependent
  BC =$\boldsymbol{regress^k}$ ($\hat{z}$, PCS)
- Load and apply scene-dependent observation error S_y^k

Scene classification based on IASI observations (in first 3 PCS of bands 1 and 2)

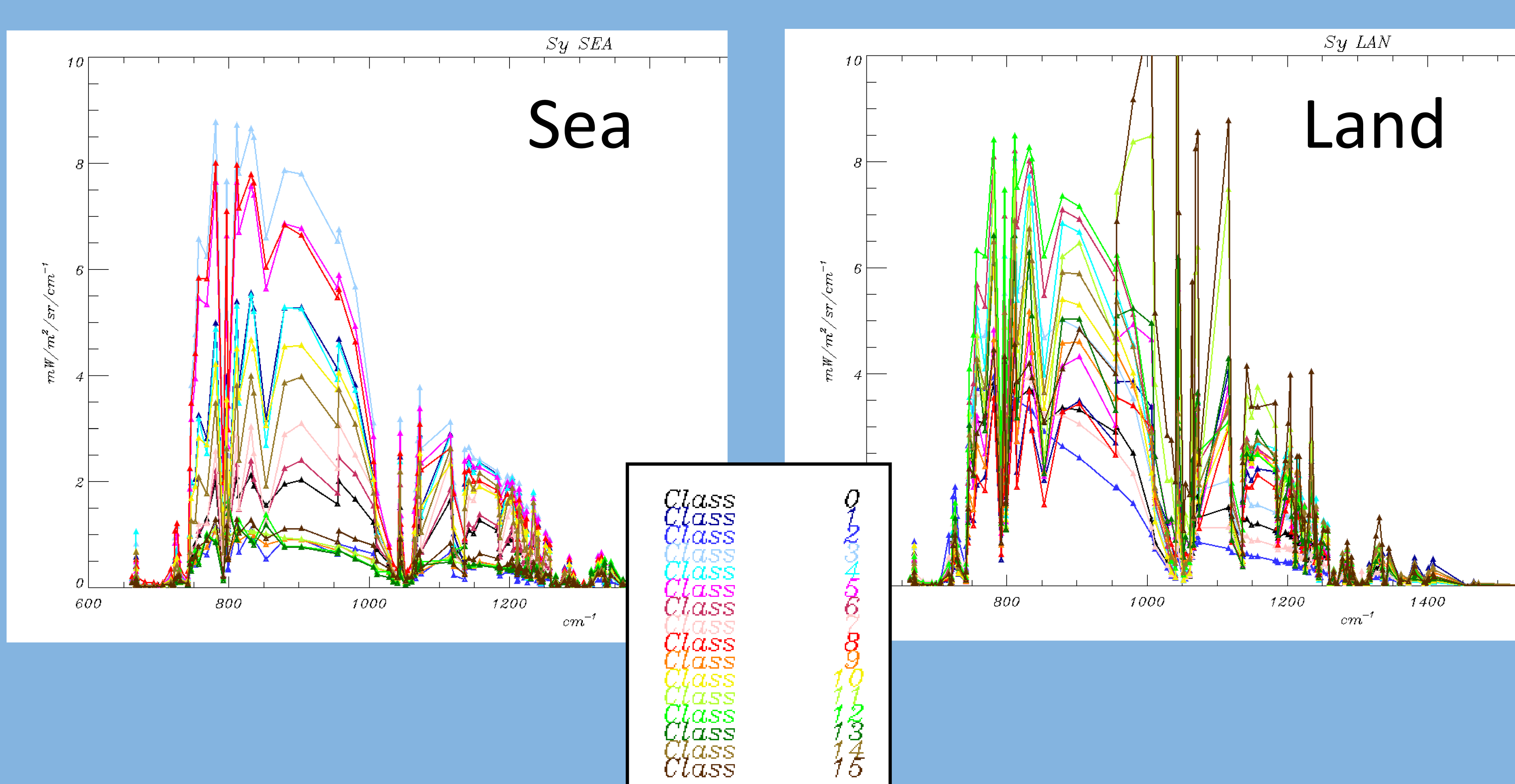






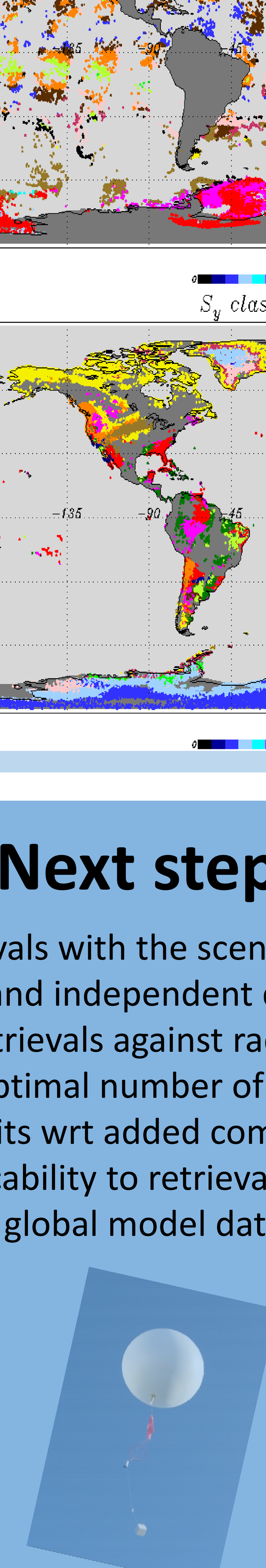


Comparisons of the observation errors in each cluster



Next steps

- Perform retrievals with the scene-dependent OEM configuration and independent datasets
- Analyse the retrievals against radio-sondes
- Evaluate the optimal number of classes and the possible benefits wrt added complexity
- Evaluate applicability to retrieval of species without representative global model data



- **Comments, ideas, suggestions … let's talk!**

Contact: thomas.august@eumetsat.int