The Community Cloud Retrieval for Climate (CC4CL): Retrieval System and Application to AVHRR and MODIS sensors

M. Jerg and the ESA Cloud CCI Team

ESA Cloud CCI Phase I

- Two global cloud property datasets with uncertainty estimates for 2007-2009:

  1) "AVHRR heritage"-product based on MODIS, AATSR, AVHRR, see also Session 2 C. Poulsen.

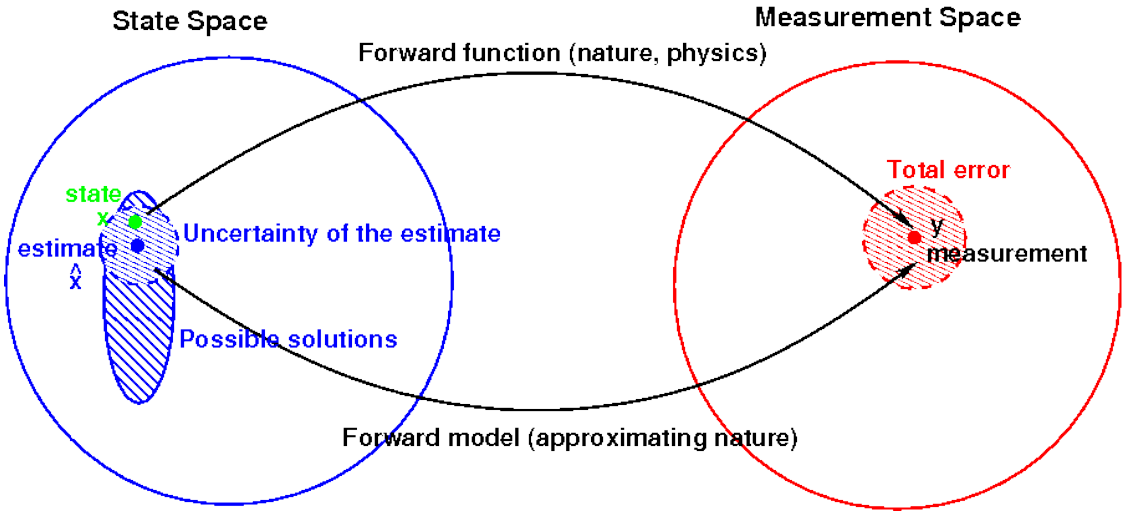
  2) Combined AATSR + MERIS product, see Session 4, C. Carbajal Henken

- Development of an "open-source" remote sensing system for cloud properties from passive imagers.
- Objective: Provide long-term coherent cloud property data sets exploiting the synergic capabilities of different Earth observation missions allowing for improved accuracies and enhanced temporal and spatial sampling better than those provided by the single sources.

Community Algorithm

- Selected through Round Robin algorithm comparison: CM SAF, ORAC, CLAVR-X for MODIS, AVHRR vs. CLOUDSAT, CALIPSO and AMSR-E.
- Advantages of optimal estimaton:
  - Consistency (sensors, channels)
  - Simultaneity (State vector)
  - Error estimates for state vector
  - Flexibility (future sensors, different channels, …)
- Development of ORAC into Community OE Cloud Retrieval for Climate (CC4CL) by DWD, RAL, UO.
- Available at http://proj.badc.rl.uk/orac via SVN.

Optimal Estimation Principle



$$y = F(\hat{x}) + \varepsilon$$

solved by maximizing

$$P(\hat{x} \mid y, x_a)$$

Through minimization of Cost-Function:

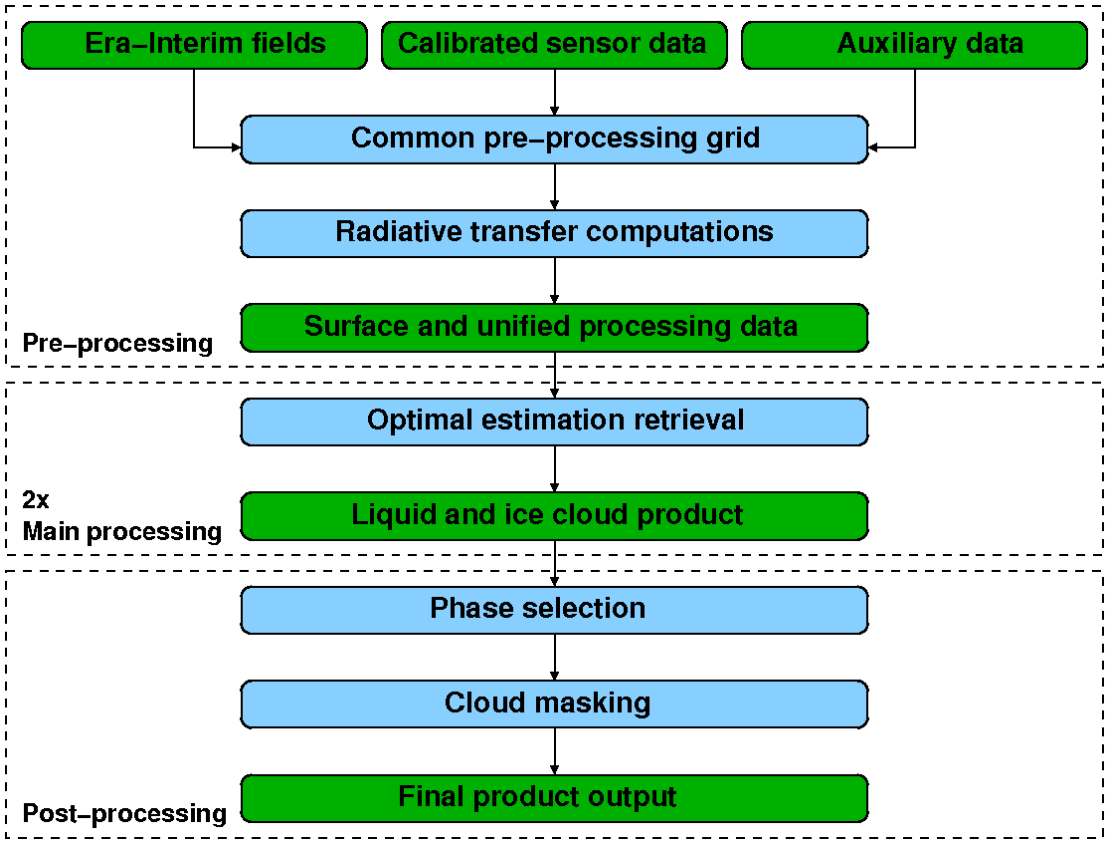
$$J(\hat{x}) = (y(\hat{x}) - y_m) S_y^{-1} (y(\hat{x}) - y_m)^T + (\hat{x} - x_a) S_a^{-1} (\hat{x} - x_a)^T$$

iteratively with the Levenberg-Marquardt method.

Covariance matrix holds uncertainty information:

$$S_x = (K^T S_y^{-1} K + S_a^{-1})^{-1} \quad ; \quad K_{i,j} = \frac{\partial y_i}{\partial x_j}$$

General processing layout



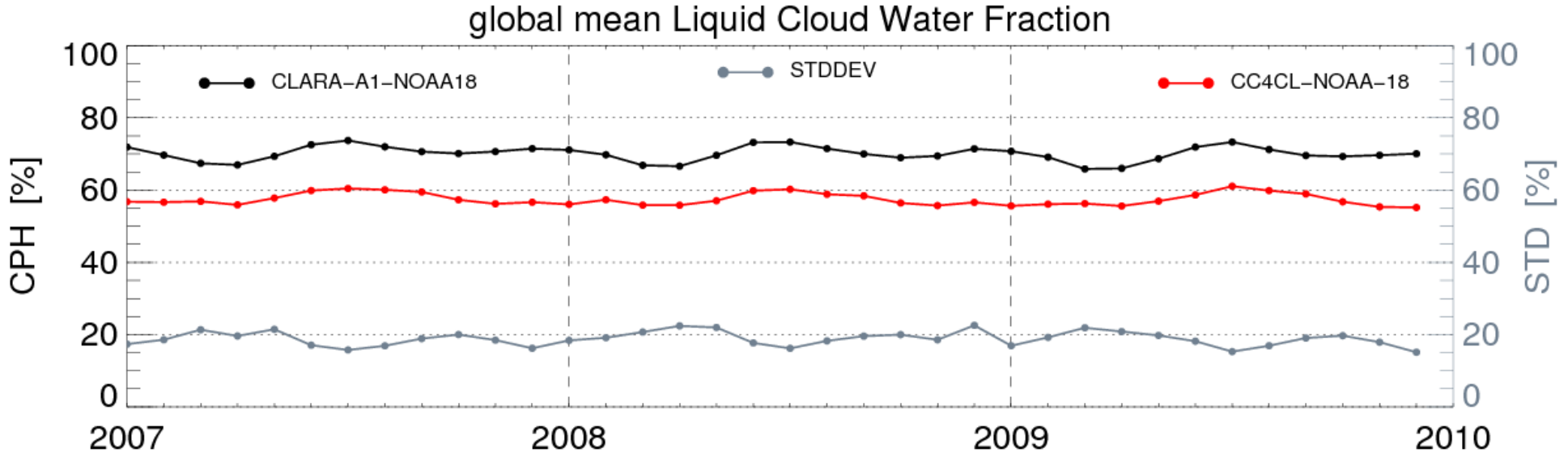
- Sensor measurements: R(650nm),R(850nm),R(1.6µm),R(3.7µm),B(11µm),B(12µm)
- COT,REF,CTP,CTT,CTH directly in state vector or derived.
- Cloud phase and cloud mask from postprocessing.

Cloud phase approach

- Apply retrieval for every pixel twice with different first-guess/a priori settings:

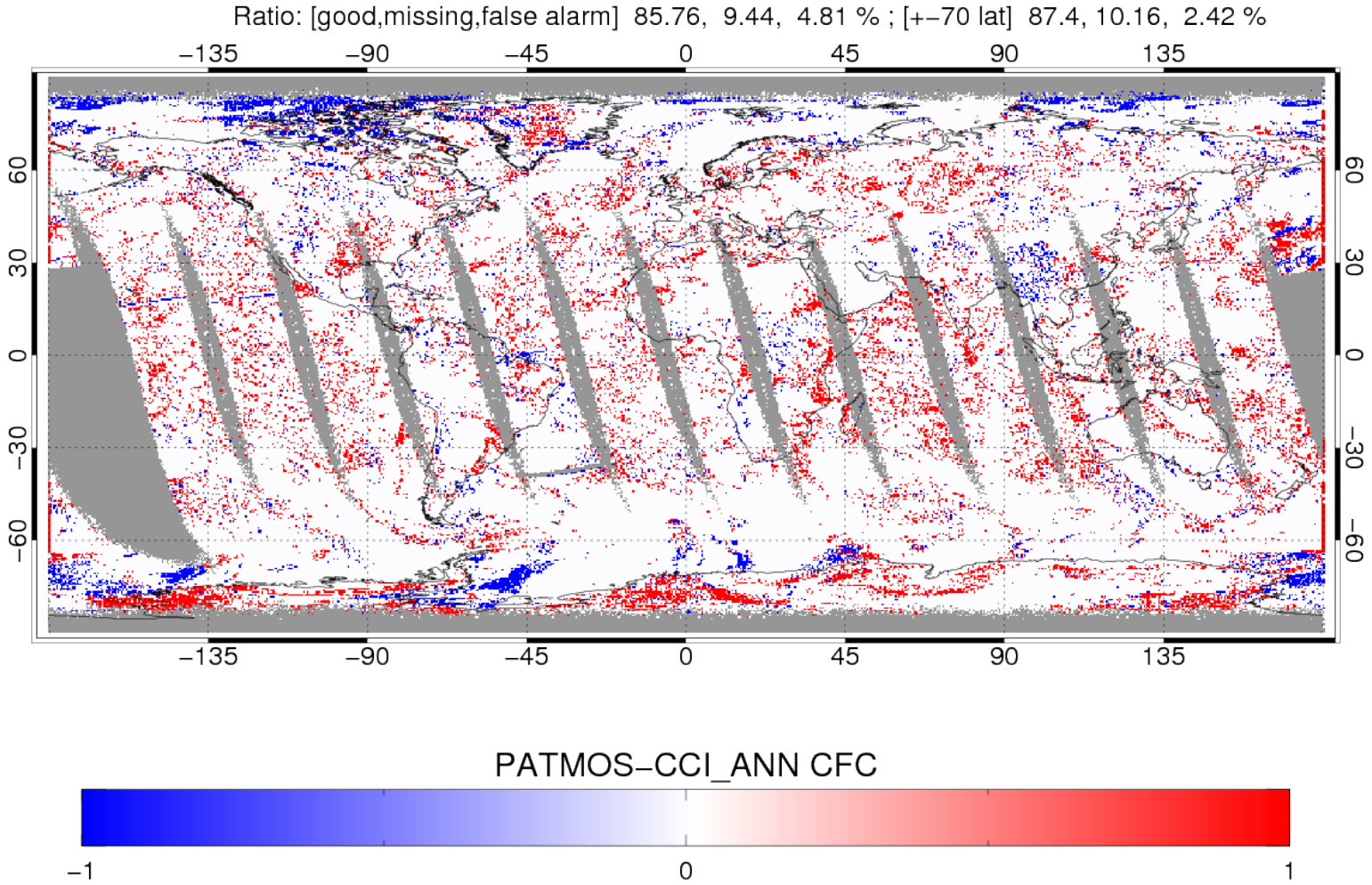
  Water(Ice): REF=12µm(30µm), CTP=700hPa(400hPa), COD=0.8, STEMP=300K

- Postprocessing: Both „raw" outputs are analyzed („Ice" selected as default)



Cloud mask approach

Two Artificial Neural Networks for day and night.
Training – dataset: 12 days of NOAA-18/Calipso collocations
Result between 0 and 1 representing a pseudo CALIPSO COT.
Comparisons show best results with thres. 0.2 over sea and 0.3 over land



Dataset processing

- Optimal estimation computationally heavy due to matrix operations, all pixels, both phases always processed.
- Sheer amount of data results in high I/O overhead.
- Retrieval installed on ECMWF and RAL HPCs, 100s of CPUs.
- Started to exploit parallel processing features: OpenMP, MPI etc.

One AVHRR orbit:

Threads	Runtime in min.
1	90
2	60
4	45

Per month:

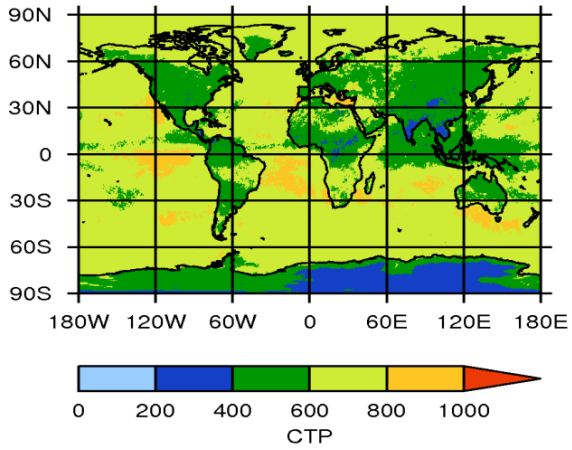
- Typical runtime for AVHRR:
  - 12-18h pure L2 processing.
- Typical runtime for MODIS:
  - 30h-36h L2 (thinned out outside Central Europe)

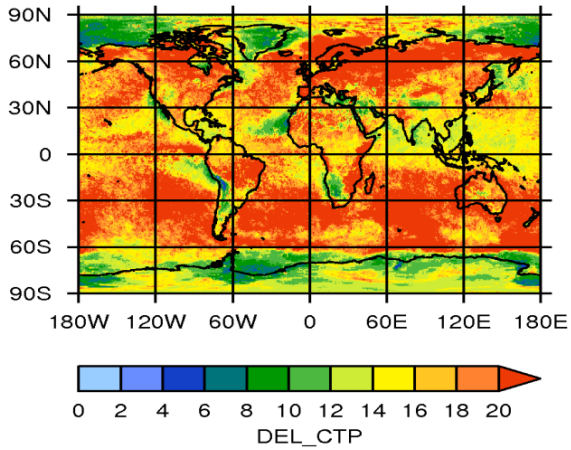
Global cloud property products

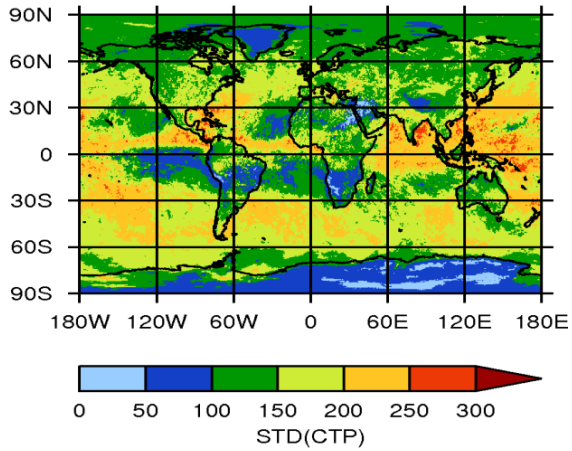
- Product suite COT,CTP,REF,CPH,CWP,CMa
  - Pixel based results including uncertainty estimates (L2)
  - 0.1 deg. daily L2 gridded composite (L3U) with smaller satellite viewing angle preferred.
  - Monthly 0.5 deg. averages, standard deviations.,Median incl. uncert. est., 2D COT-CTP Histograms (L3C).
- Present averaging approach:
  - Spatial and temporal averaging mixed.
  - Uncertainties treated as regular variables.
- L3S: Sensor families merged.

Error budget and propagation into L3

Error: Difference to the true value (unknown).
Uncertainty: Estimate of error, „doubt" of measurement.







But errors should add up according to their correlation:
(pers. comm. R. Bennartz)

Error of the mean:
(constant correlation)

$$\sigma^2_{\langle x \rangle} = c \cdot \langle \sigma_i \rangle^2 + (1-c) \cdot \frac{1}{N} \langle \sigma_i^2 \rangle$$

correlated

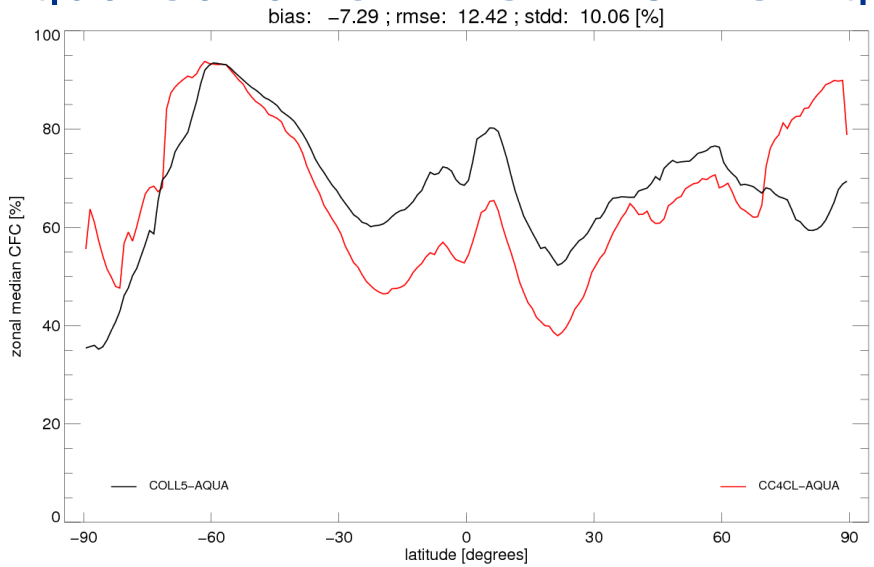
uncorrelated

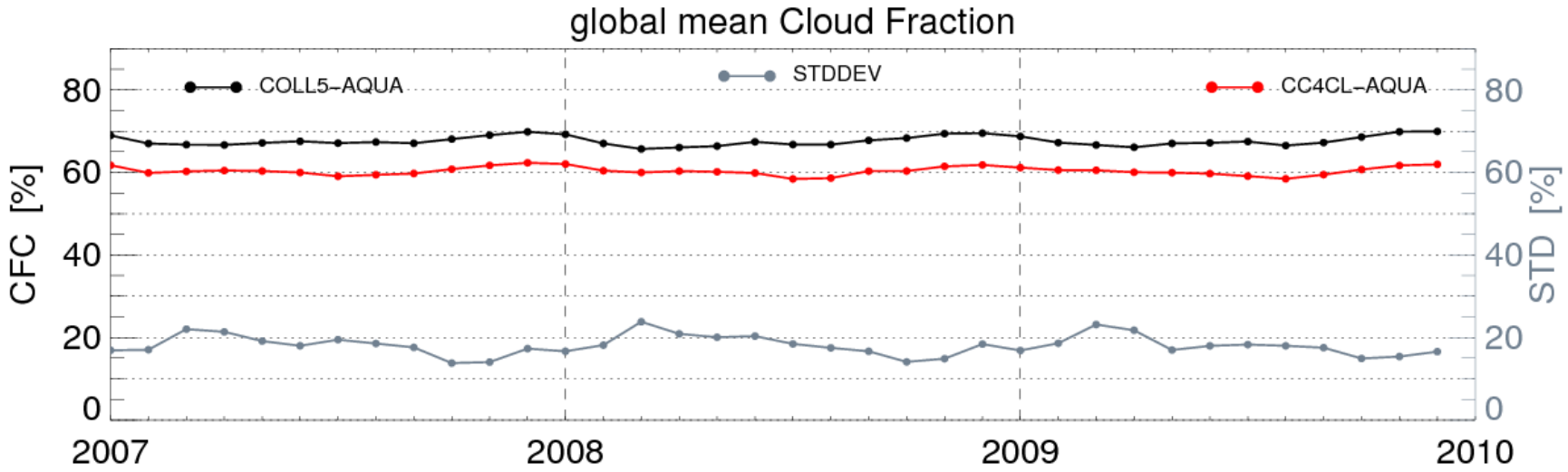
Correlation coefficient generally unknown…?
Seperate spatial and temporal averaging in the future.

Validation results

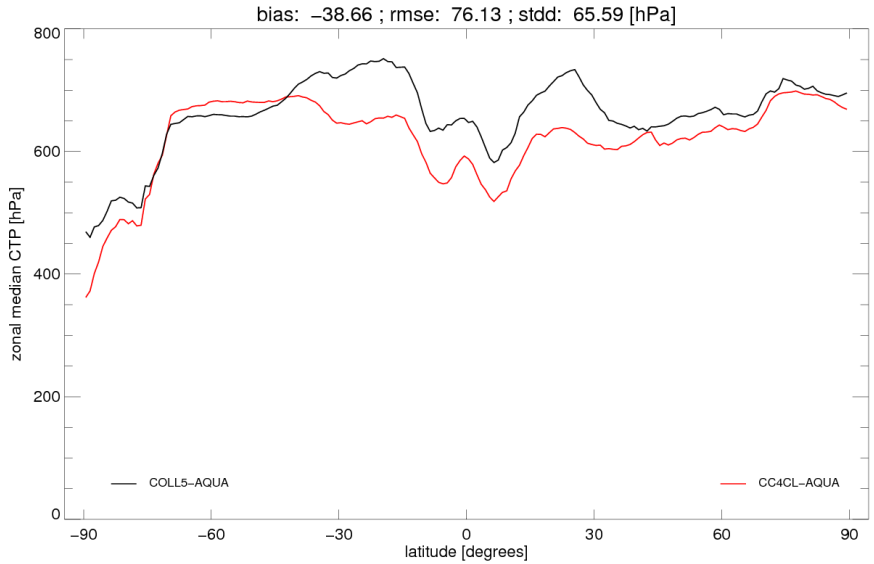
(see also poster by S. Stapelberg on Wednesday)

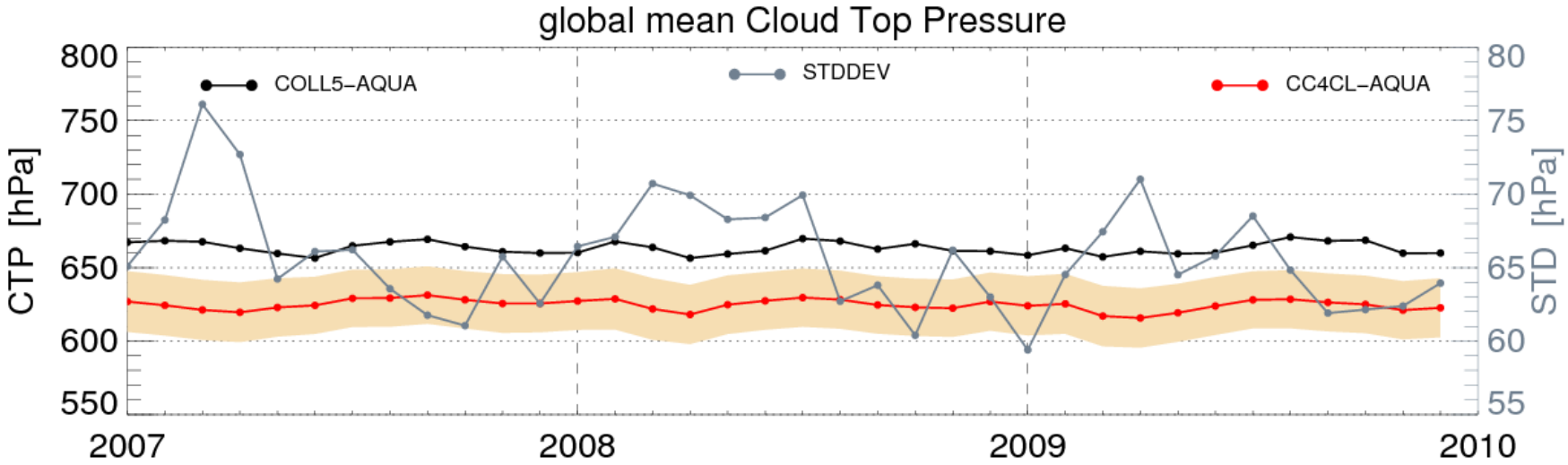
CFC CCI MODIS Aqua Col.6 vs. NASA MODIS Aqua Col. 5 2007-2009





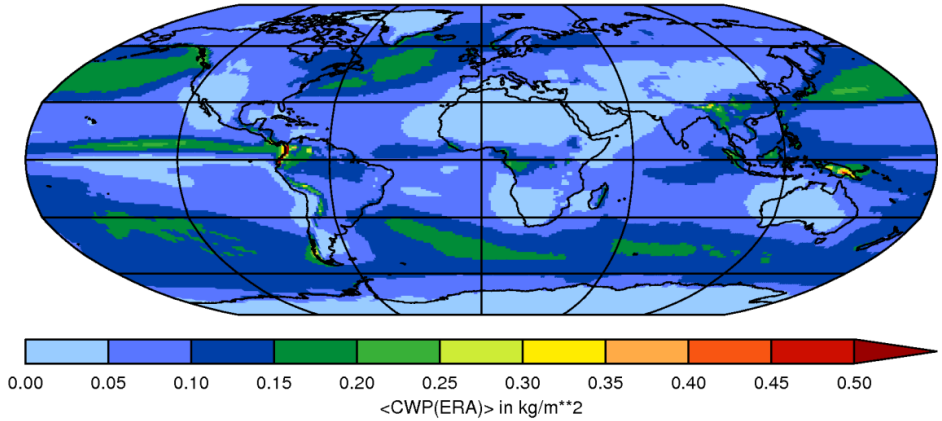
CTP CCI MODIS Aqua Col.6 vs. NASA MODIS Aqua Col. 5 2007-2009

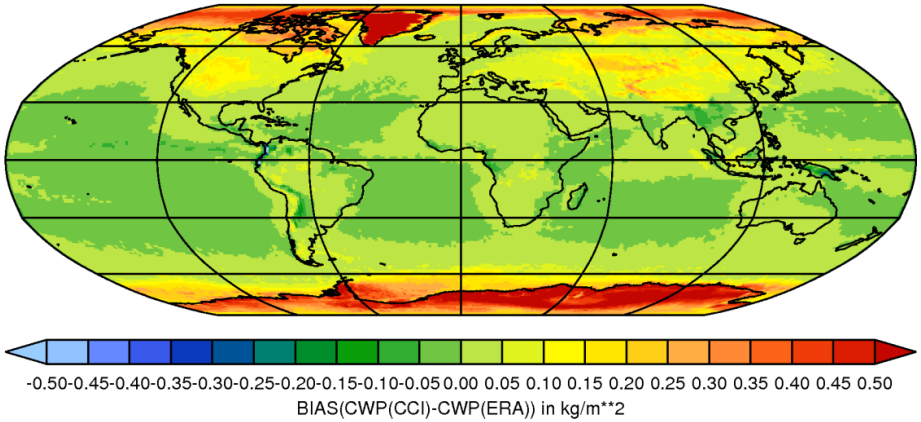


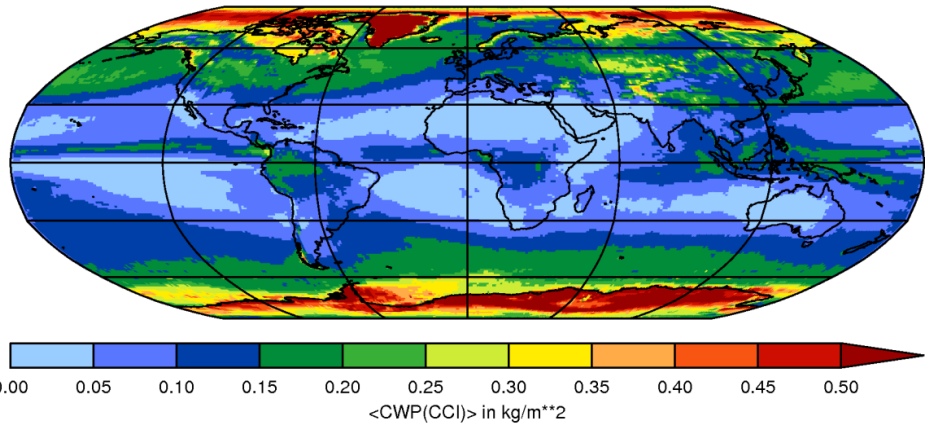


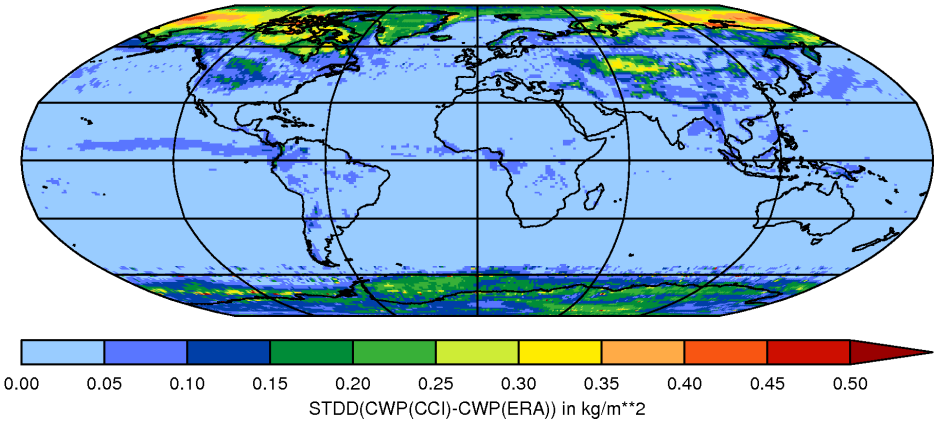
CWP Comparison with ERA-Interim
2007-2009 L3S (M+A)

200701- 200912







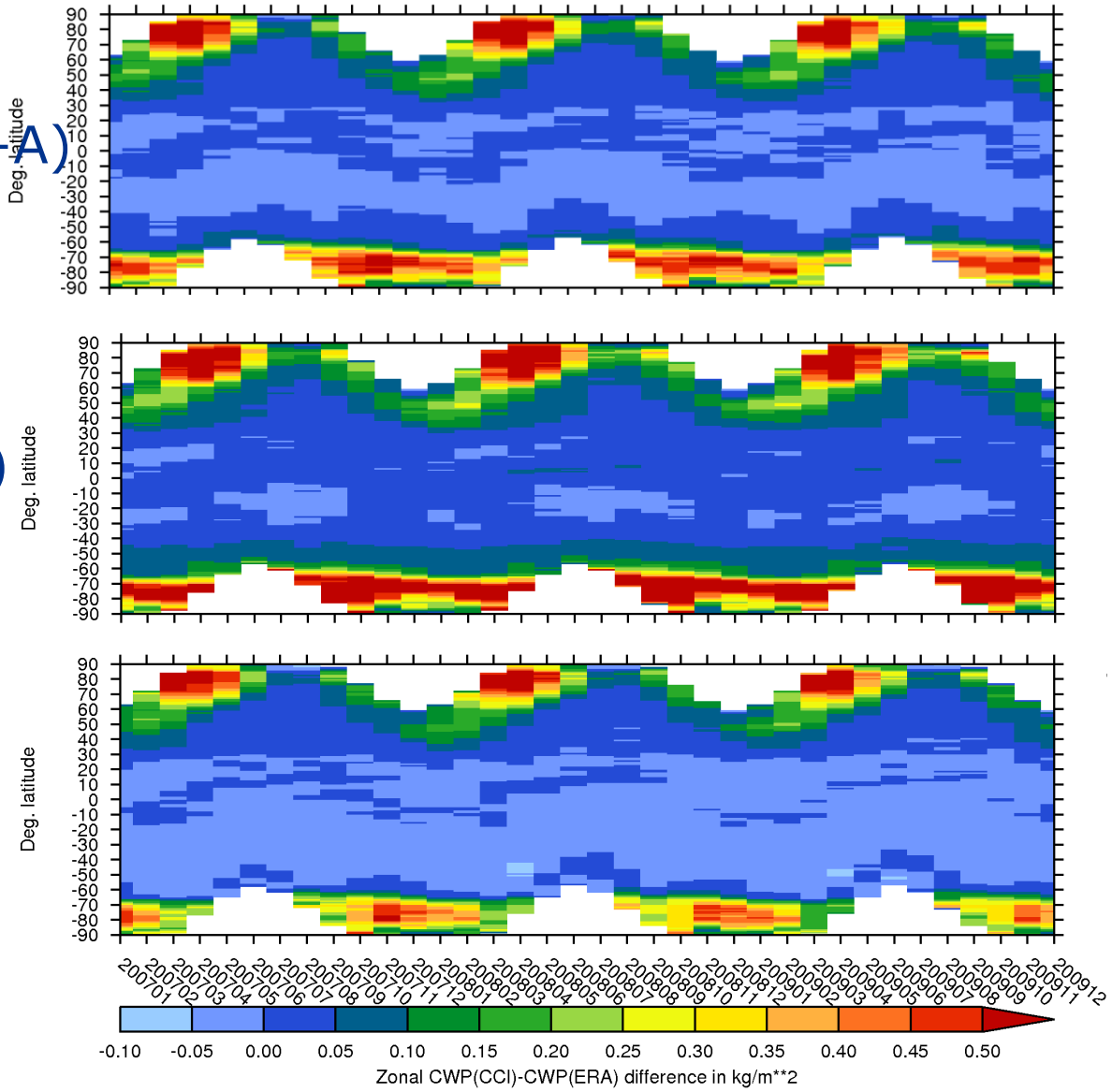


Different seasonal behavior in southern hemisphere

2007-2009 L3S (M+A)

2007-2009 L3S (M)

2007-2009 L3S (A)



For more results see poster by S. Stapelberg tomorrow.

Future work in Phase II

- Cloud phase according to Pavolonis & Heidinger.
- Multi-layer cloud treatment.
- Investigate and improve error budget and propagation.
- Processing performance.
- Processing of the complete time-series:
  - Aqua: 2002-present.
  - Terra: 2000-present.
  - AATSR: 2002-2012.
  - Prime AVHRRs:1982-present.