9th SPARC DA workshop, NMT, Socorro, NM; 11-13 June 2012

Parametrized observation operators and their application on illustrative 3D atmospheric chemical fields

Tijl Verhoelst, Jean-Christopher Lambert, and Sophie Vandenbussche

Belgian Institute for Space Aeronomy

Acknowledgments: funded by BELSPO/ProDEx SECPEA + A3C, and by EC FP6 GEOmon + FP7 NORS

General context: assimilation of observations of chemical atmospheric components

In their report on last year's SPARC DA workshop, Jackson & Polavarapu write:

> "Ozone in reanalyses remains a challenging area. While total column O_3 values may be useful, the vertical distribution is generally unreliable because of the dearth of vertical profile measurements."

> "... As a result, 6 target areas for future activity were identified...
> The second goal follows largely from the talks of Lambert and Yudin, and this is to develop a greater interaction between the satellite retrieval and data assimilation communities."

Outline

- The ideal observation operator
- Pragmatic observation operators
- Illustrations:
  - smoothing errors (global maps, time series)
  - smoothing differences (between remote sensing datasets)
- Applications:
  - interpretation of error budget
  - improved co-location criteria
- Conclusions and prospects for data assimilation

Observation operators in assimilation

Purpose and implementation

low the tropopause.

2.2 The 4D-Var system

Data assimilation is done using 4D-Var (Talagrand and Courtier, 1987). This method optimizes the model initial conditions to reproduce a set of observations over a time window. This is done by minimizing the following objective function, $J(\mathbf{x})$ (also denoted cost function) (Talagrand and Courtier, 1987) using the standard notation of Ide et al. (1997):

$$J(\mathbf{x}) = \frac{1}{2}[\mathbf{x}(t_0) - \mathbf{x}^{\mathrm{b}}(t_0)]^T \mathbf{B}_0^{-1}[\mathbf{x}(t_0) - \mathbf{x}^{\mathrm{b}}(t_0)] + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}^{\mathrm{o}}(t_i) - H(\mathbf{x}(t_i))])^T \mathbf{R}_i^{-1}(\mathbf{y}^{\mathrm{o}}(t_i) - H(\mathbf{x}(t_i)]) \quad (1)$$

given the model evolution equation

$$\mathbf{x}(t_i) = M_{i-1}[\mathbf{x}(t_i - 1)] \quad (2)$$

where $\mathbf{x}(t_i)$ represents the model state vector at time t_i, $\mathbf{x}^{\mathrm{b}}(t_0)$ is the first guess and $\mathbf{B}_0$ is the background error covariance matrix of $\mathbf{x}^{\mathrm{b}}(t_0)$. Vectors $\mathbf{y}^{\mathrm{o}}(t_i)$ and matrix $\mathbf{R}_i$ are, respectively, the observation state vector and the error covariance matrix associated with the observations at time t_i. The observation operator H maps the model state into the observation space and M is the model operator that calculates the time evolution of the model state. Minimization of Eq. (1)

Errera et al., Atmos. Chem. Phys.,8, 2008

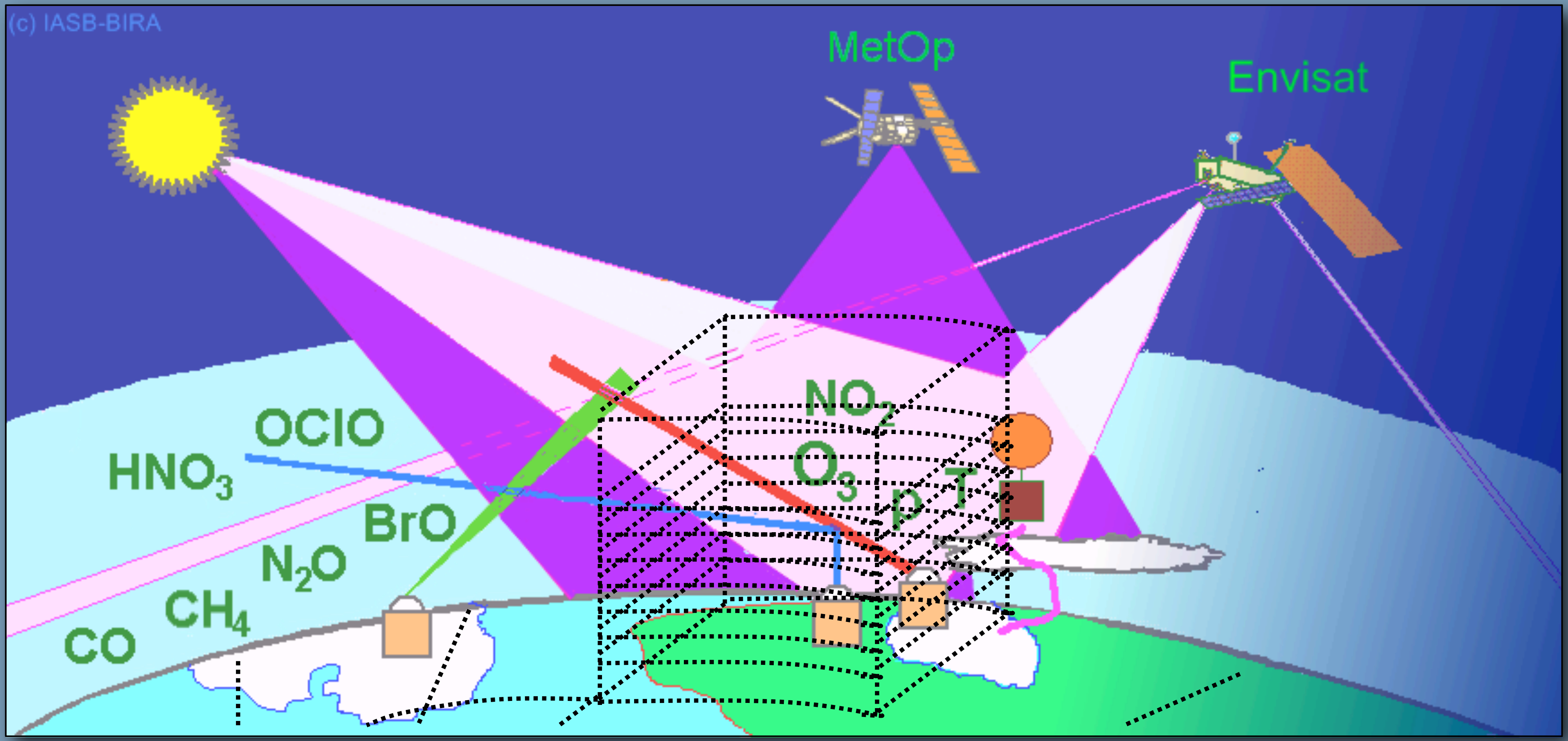
are assimilated by the BASCOE system to constrain its CTM outputs, while HALOE and POAM-III data are monitored by the system and used for a posteriori evaluation of the BASCOE analyses. Assimilated data are volume mixing ratio at the location of the tangent point and the observation operator interpolates linearly the model values at the eight grid points to the surrounding tangent point of any available observation at ±15 min of the model time step. In the monitoring procedure, the BASCOE observation operator is used only to map the analyses at HALOE and POAM-III locations. Finally, BASCOE analyses interpolated at observation locations (as

Ideally, *H(x(t$_i$))* mimics perfectly the observation, i.e. including all smoothing and sampling properties of the observation system in both the vertical and horizontal dimensions.

Observation operators

Smoothing and sampling issues

The airmass contributing to the retrieved value is determined by (1) the forward model, that is, the radiative transfer through the atmosphere following the entire optical path of photons detected by the observational set up, and by (2) the retrieval model.



At the spatial resolution of current Data Assimilation systems, this airmass may spread over multiple model grid cells, and is often not centered on the nominal location of the observation (e.g. tangent point or footprint).
»» smoothing and sampling issues in the horizontal and vertical directions!

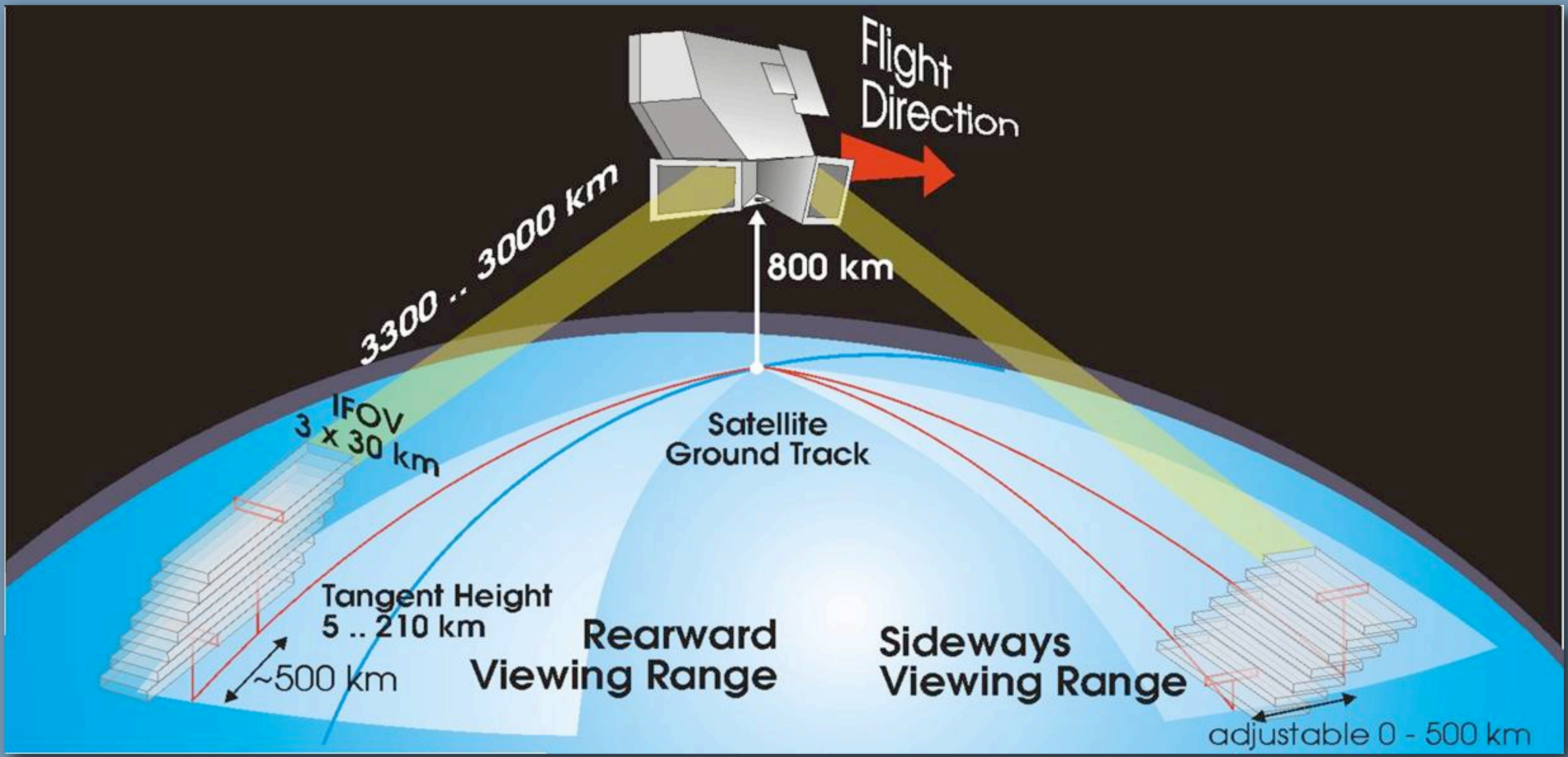
Pragmatic observation operators

Objectives

- Reduce apparent discrepancies generated by differences in smoothing and sampling of the atmospheric field
  - ➡ match the characteristics of the retrieved information
  - ➡ depending on species, atmospheric state, and type of use
- Are not just a way to “adapt the resolution”, but should also address issues of sensitivity and information content (e.g. no measured information below clouds, or drifting vertical sensitivity due to instrumental degradation…)
- Must be affordable in terms of design (e.g. built with existing RT tools or with minor extensions/modifications) and use (no heavy RT calculation for every single ingestion)

Pragmatic observation operators

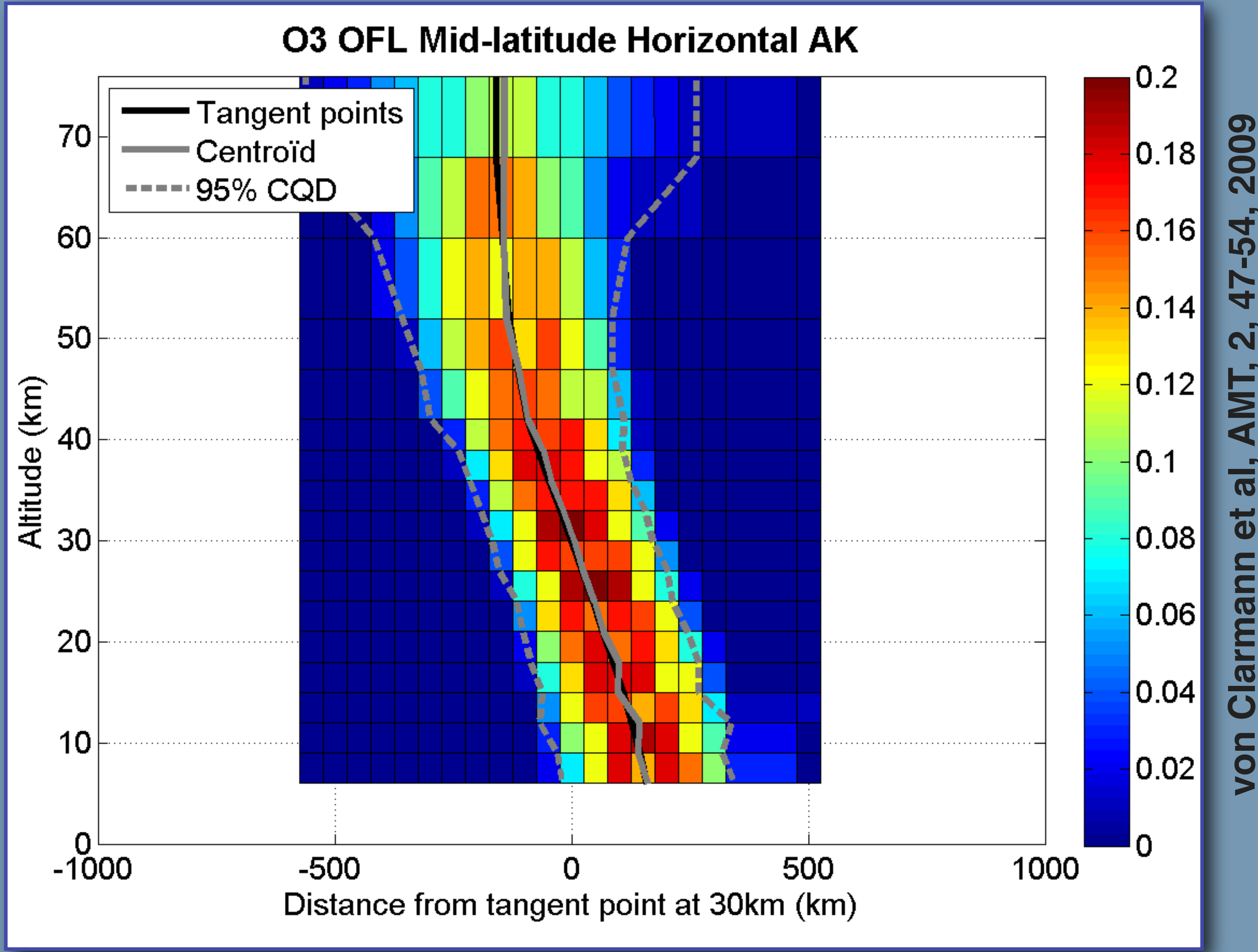
An example: O_3 with MIPAS



Pragmatic observation operators

An example: O_3 with MIPAS

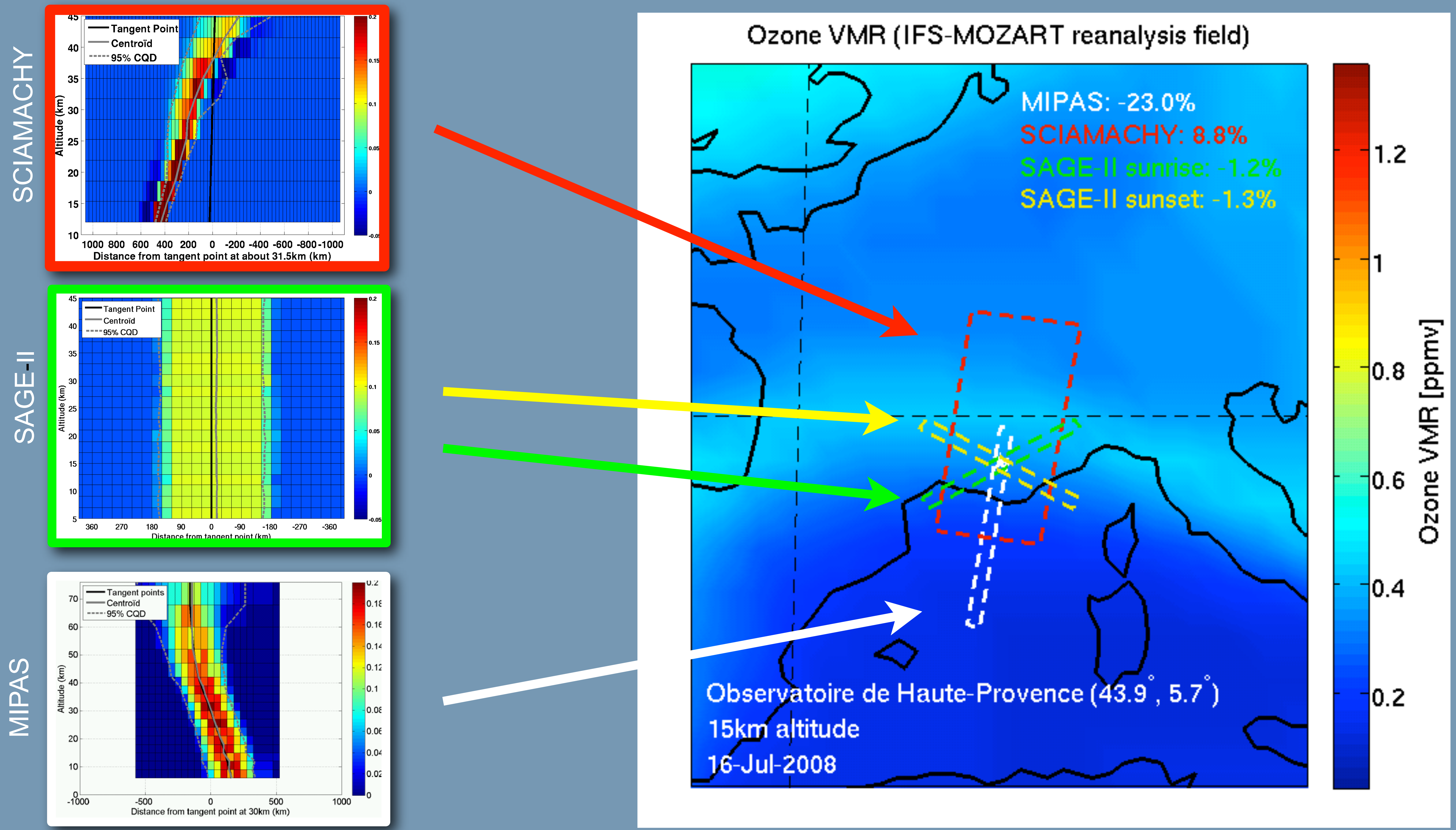
Horizontal Averaging Kernel catches the information spread/offset of the 1D profile



made at IMK, based on settings of the ESA IPF 4.61 processor, nominal mode

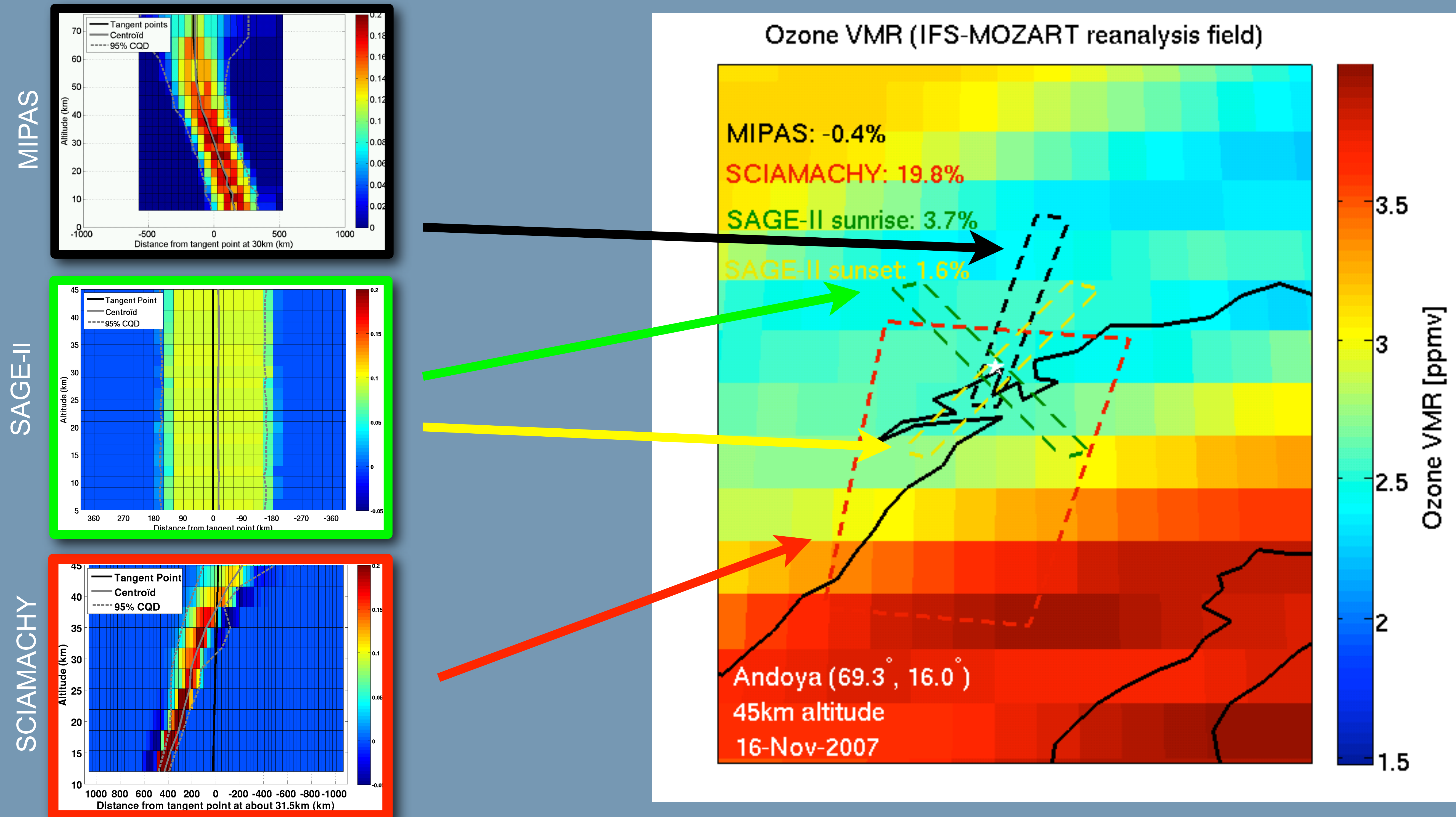
Illustration

O_3 horizontal smoothing errors



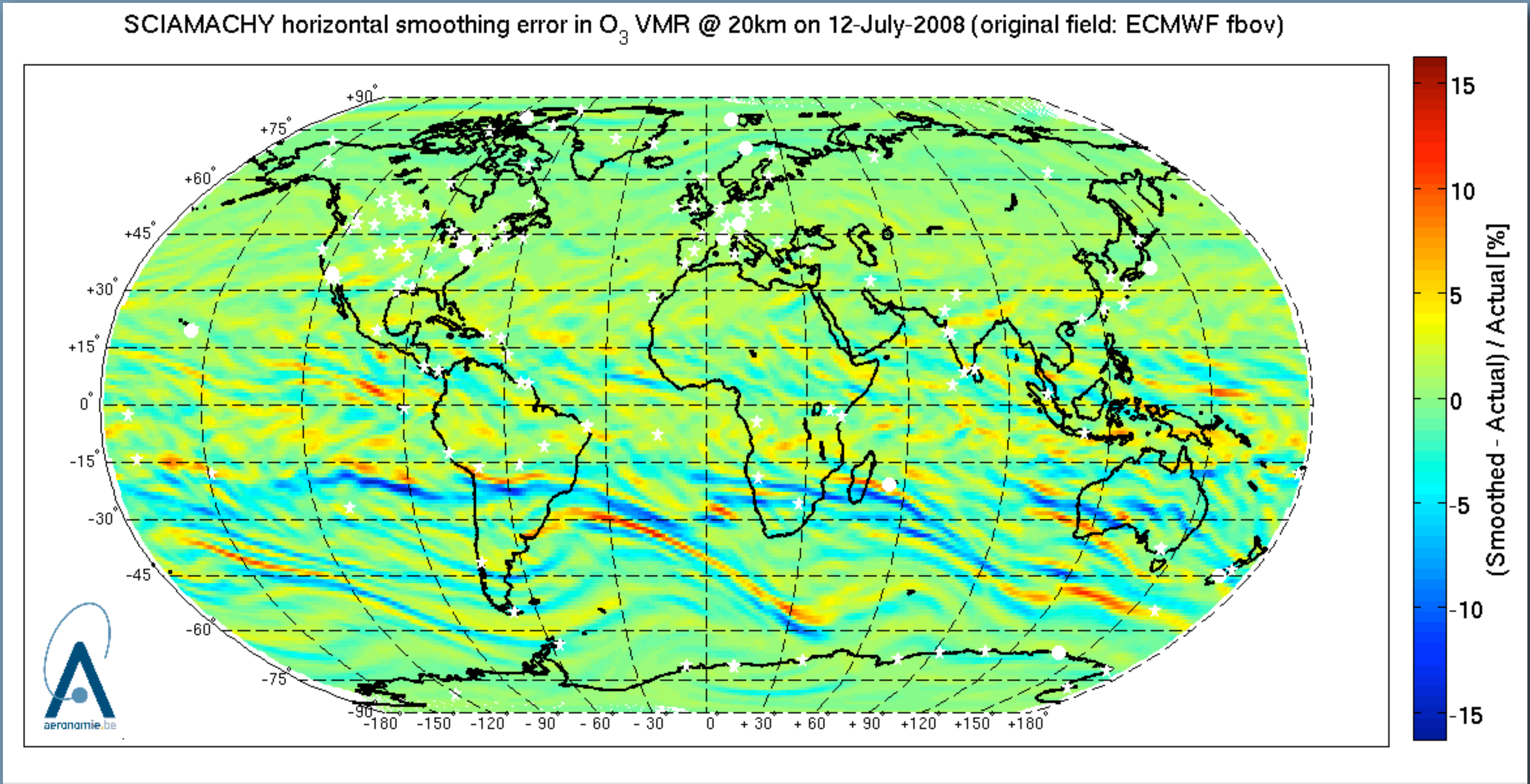
Illustration

O_3 horizontal smoothing errors



Illustration

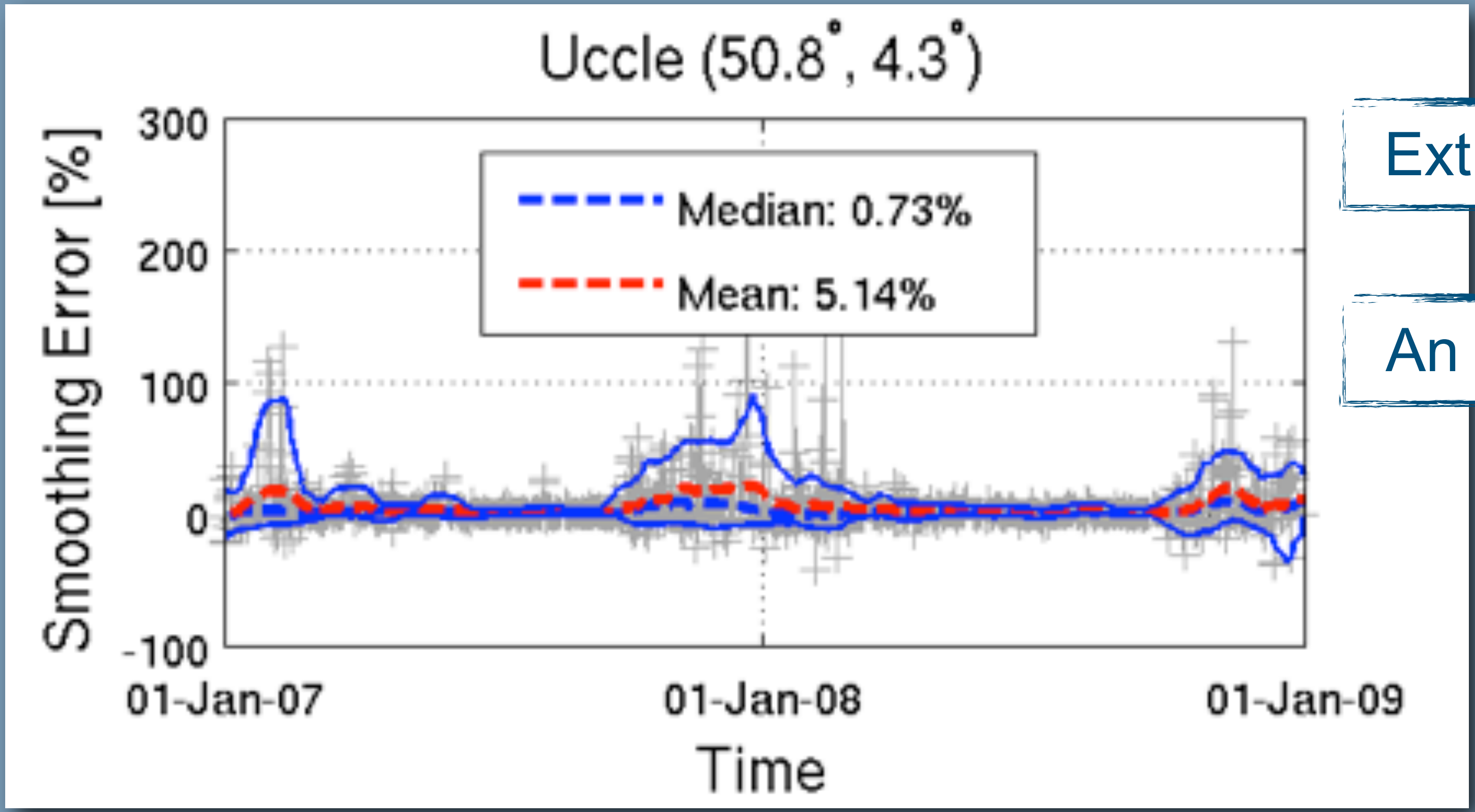
O_3 smoothing error global snaphot



Illustration

O_3 smoothing error @ selected stations from 2007-2008

SCIAMACHY limb-scattered UV-Vis sunlight observations at 15km altitude above Uccle, B



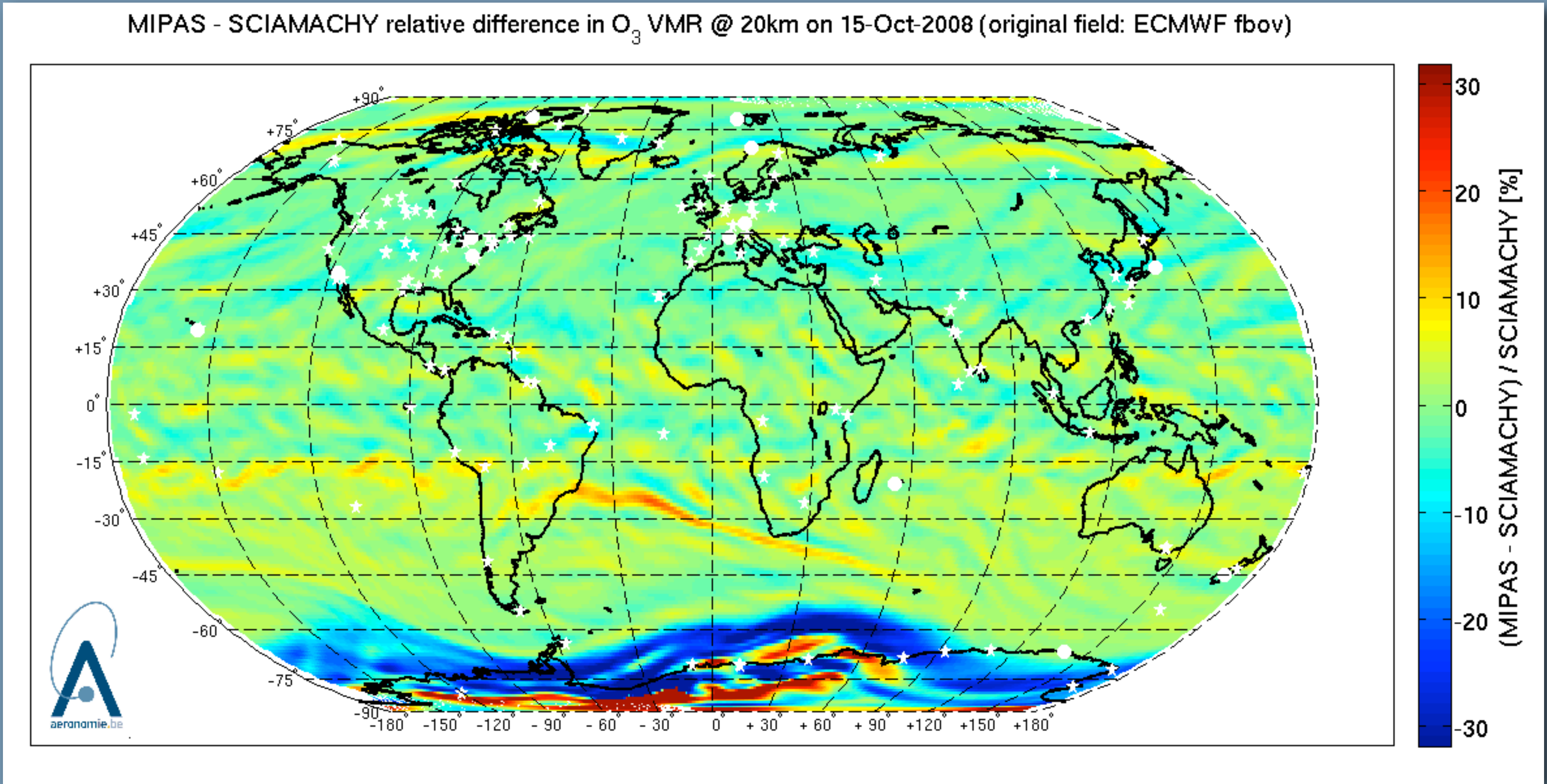
Extreme events may cause “data loss”

An additional noise term is introduced.

Positive biases in winter - due solely to the smoothing properties - may introduce unphysical seasonal signals in comparisons!

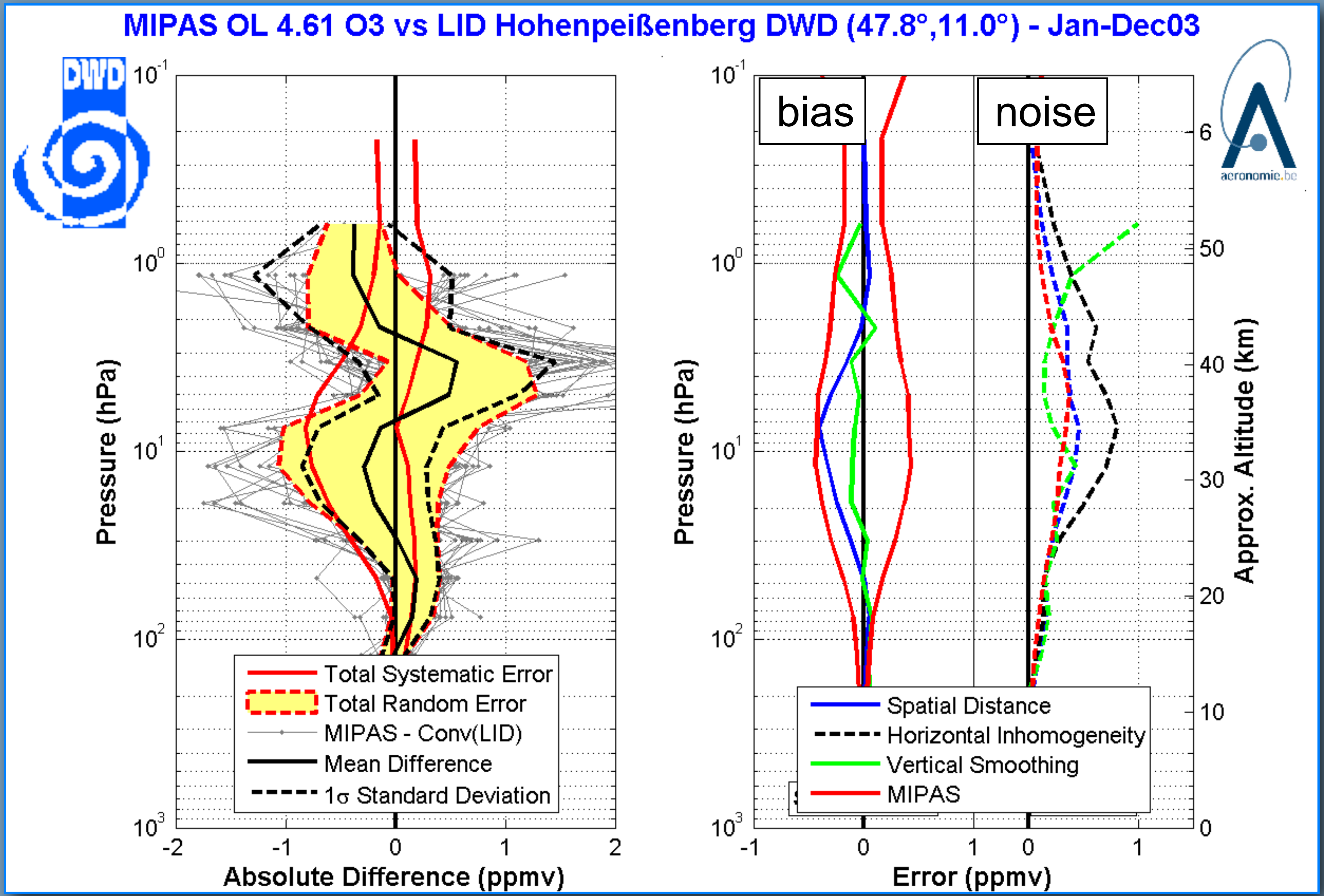
Illustration

Apparent discrepancies due to O_3 smoothing differences



Application

Error budget of ground-based satellite validation

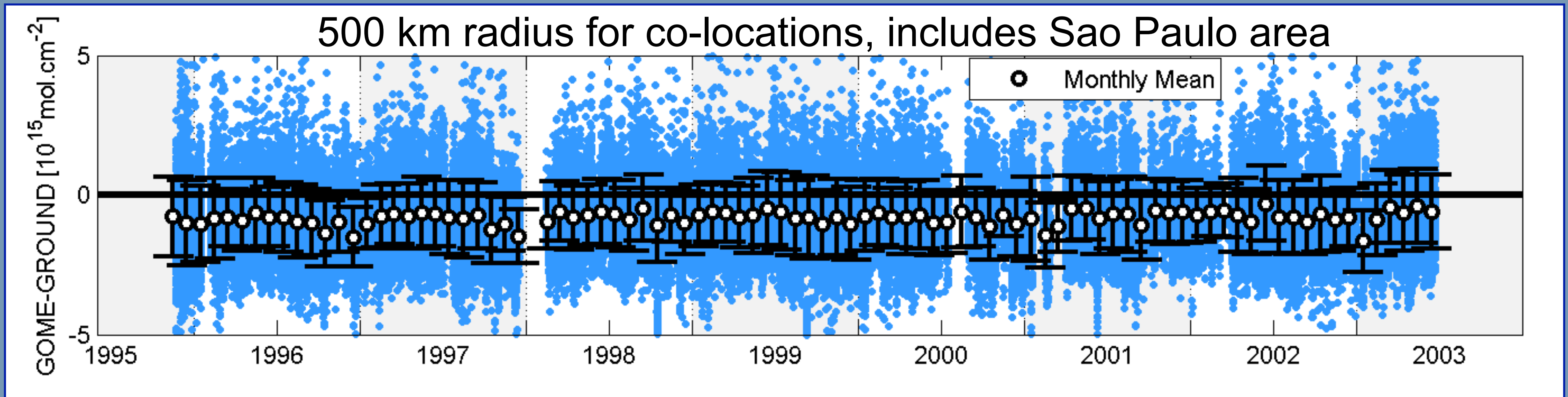


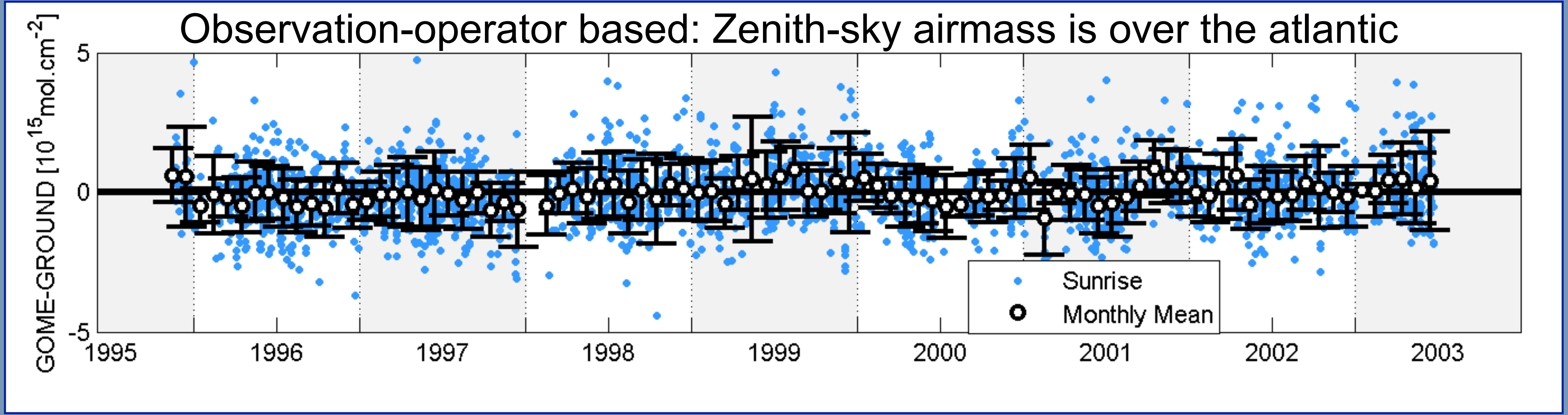
Method described in Section 4.1 of Cortesi *et al.*, ACP 2007

Application

Improved co-location criteria in ground-based satellite validation

NO2 columns: satellite versus ground-based over Bauru, Brasil





Clear reduction in noise **and** bias

Ground-based data courtesy of CNRS/LATMOS/UNESP

Conclusions and prospects

- Pragmatic observation operators allow quantification of vertical and horizontal spread and offset of remotely sensed data, at relatively "low cost"
- They can be used to improve co-location criteria, comparison strategies, and interpretation of the error budget in validation work.
- How can/should this be translated in chemical data assimilation? Improved ingestion, improved correlations?
- How do we quantify the improvement that follows from the implementation of such observation operators in the ingestion scheme? Convergence criteria? Cost function criteria? Impact on noise parameters or correlations? Ratio of contributing/rejected observations? …
- Support from DA community is requested to define diagnostics of improvement and appropriate case studies.