Let's start talking about uncertainties, part 2: Comparisons and measurement mismatches

Tijl Verhoelst, J.-C. Lambert, G. de Leeuw, A. Fassò, T. Gardiner, P. Green, R. Kivi, F. Madonna, and K. Rannat
and the WP3 team

Coherence between theoretical measurement uncertainty and actual comparison results?

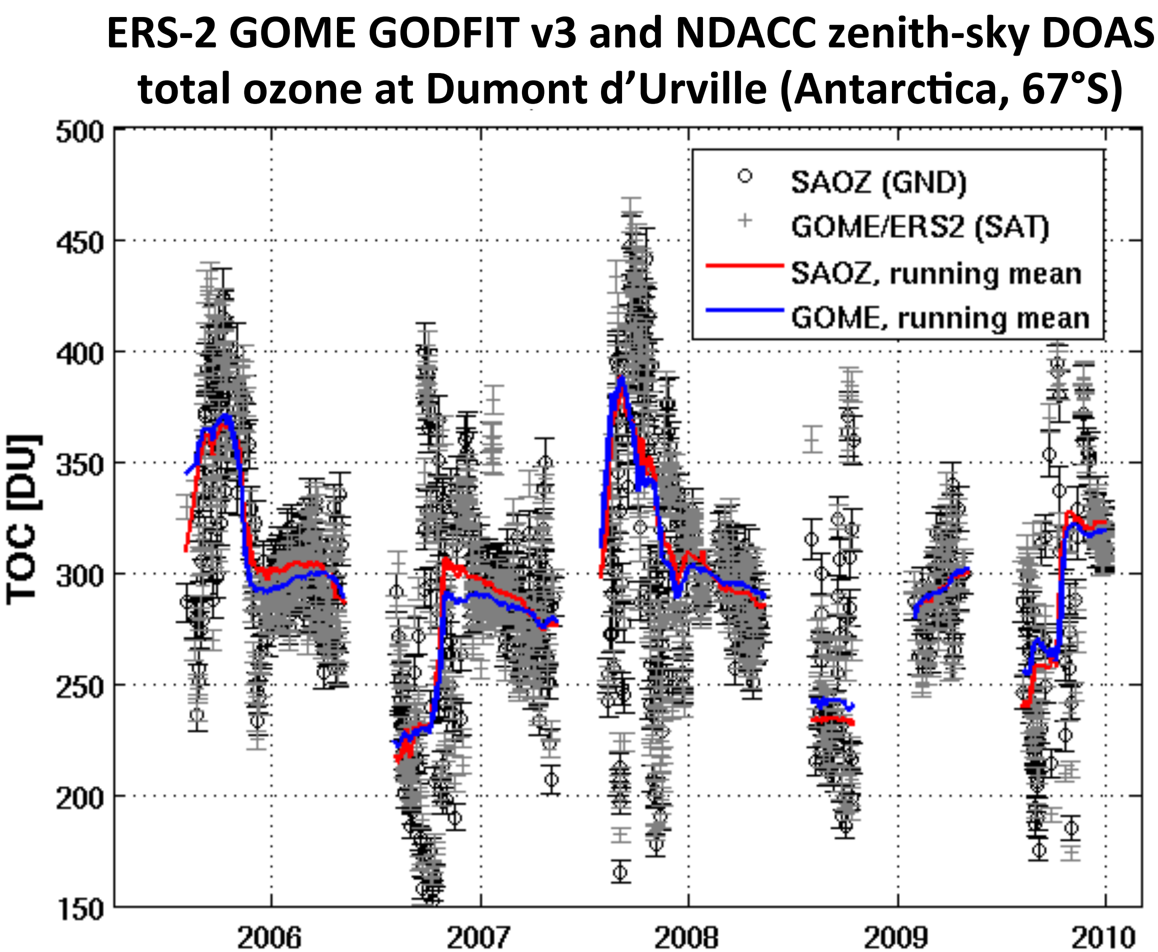


Figure from Verhoelst et al., AMTD, 2015

Coherence between theoretical measurement uncertainty and actual comparison results?

Relative difference between
ERS-2 GOME GODFIT v3 and NDACC zenith-sky DOAS
total ozone at Dumont d'Urville (Antarctica, 67°S)

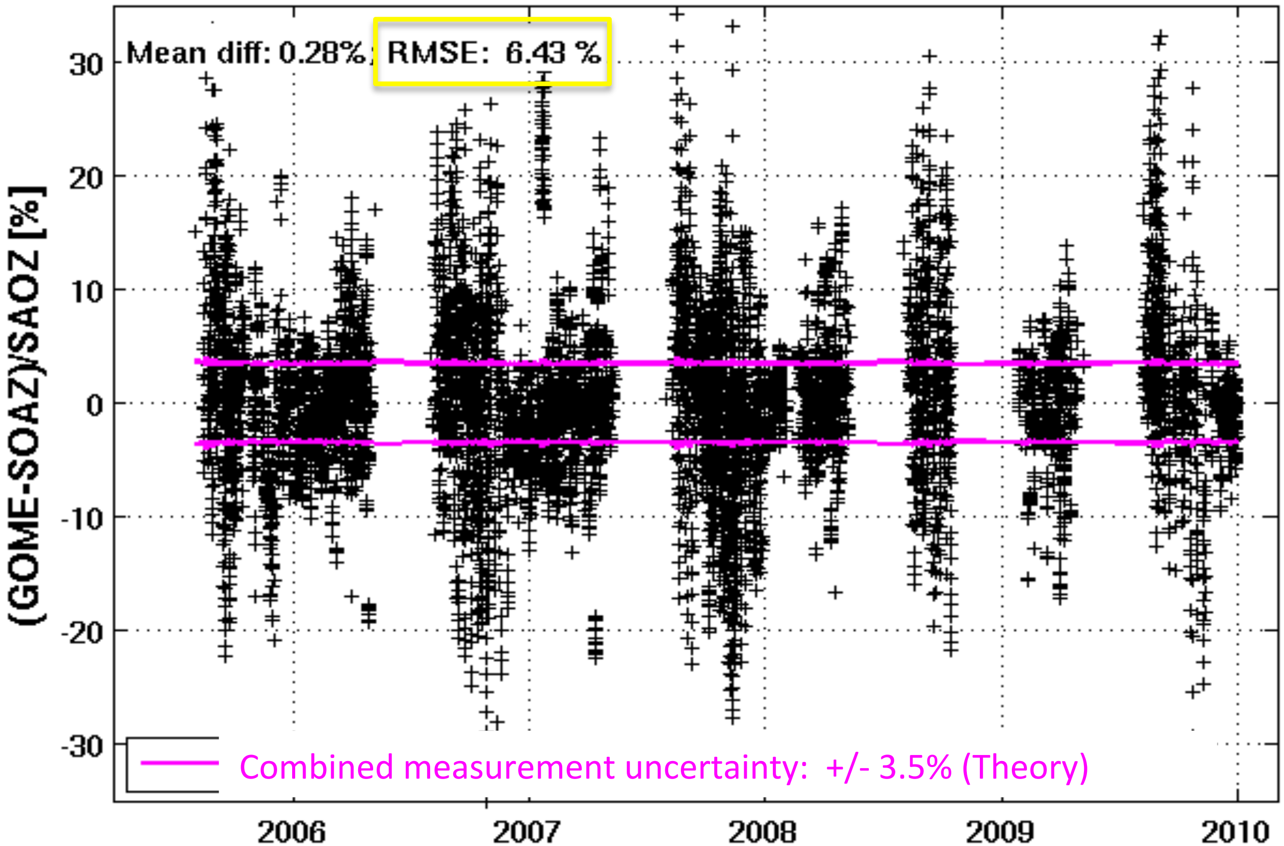


Figure from Verhoelst et al., AMTD, 2015

Combined measurement uncertainty versus actual median and spread of the relative differences

Relative difference between ERS-2 GOME GODFIT v3 and NDACC zenith-sky DOAS total ozone at Dumont d'Urville (Antarctica, 67°S)

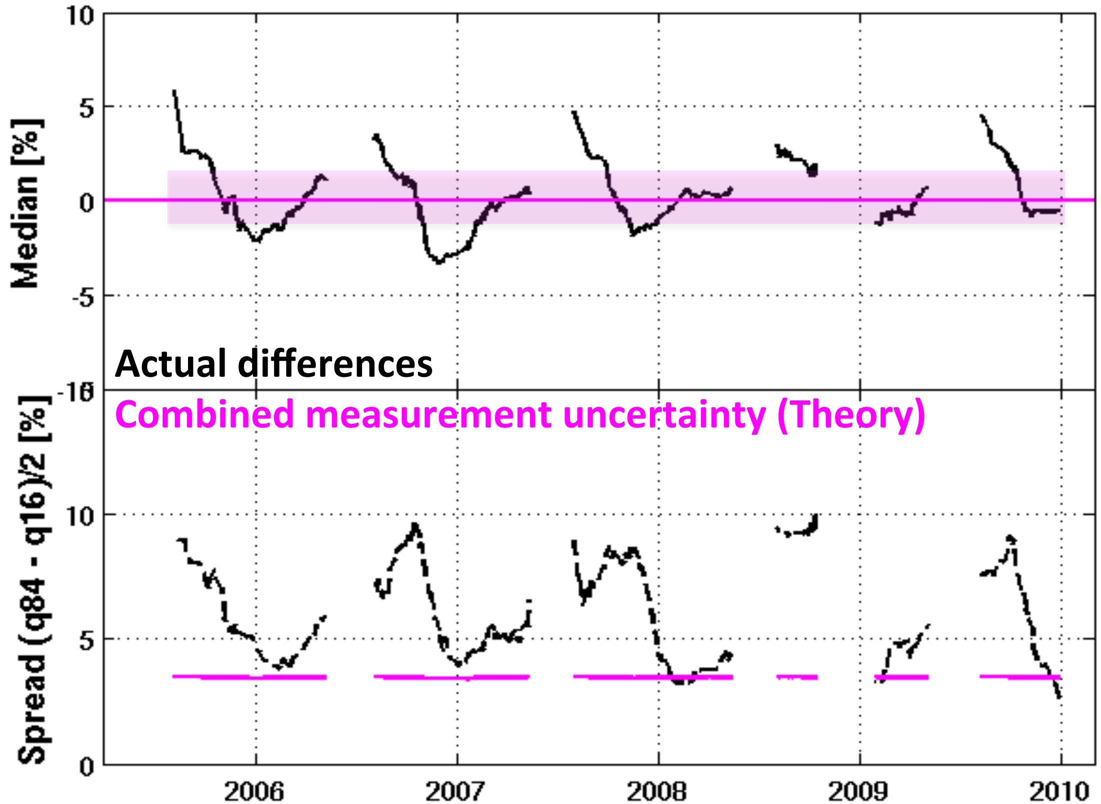
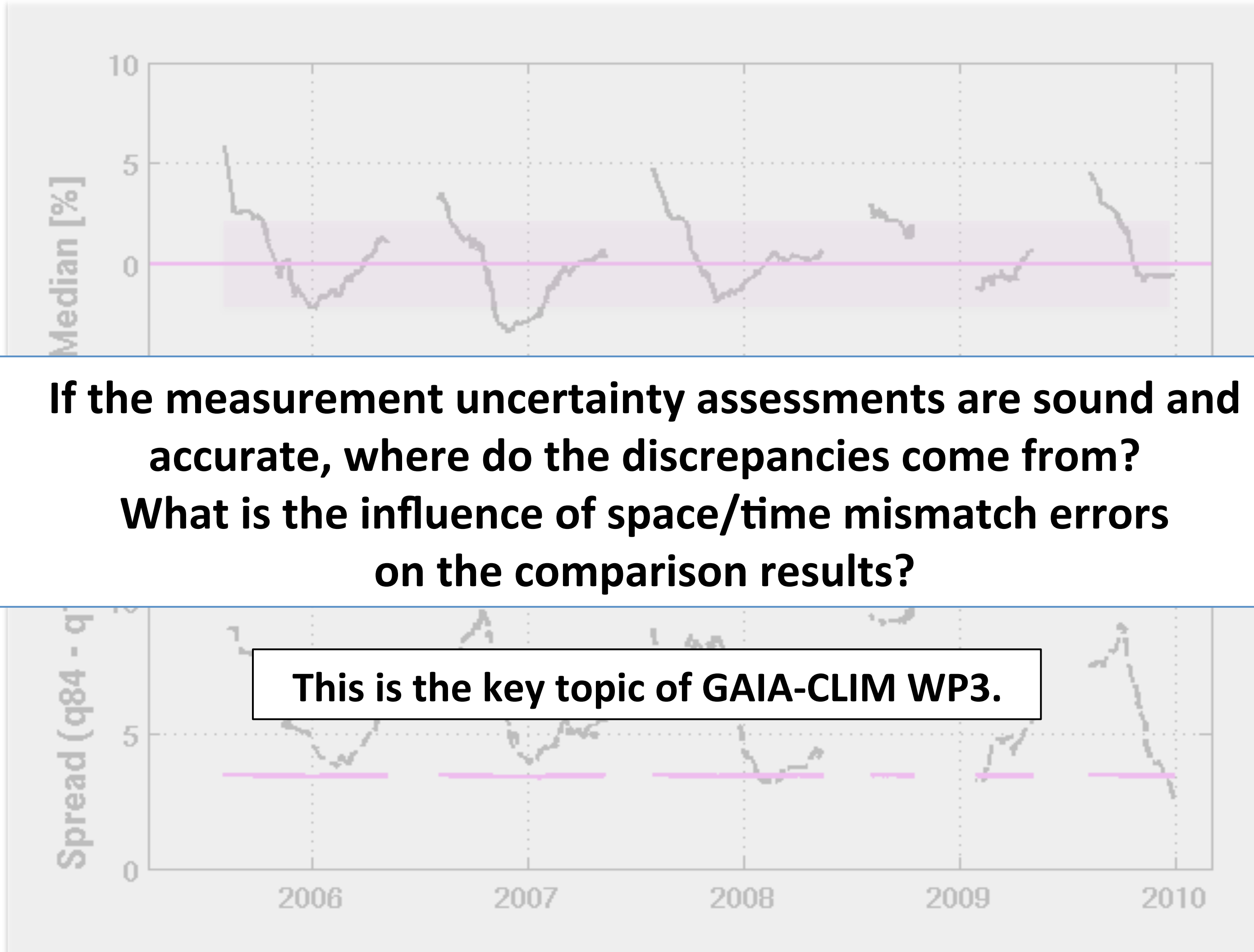


Figure from Verhoelst et al., AMTD, 2015

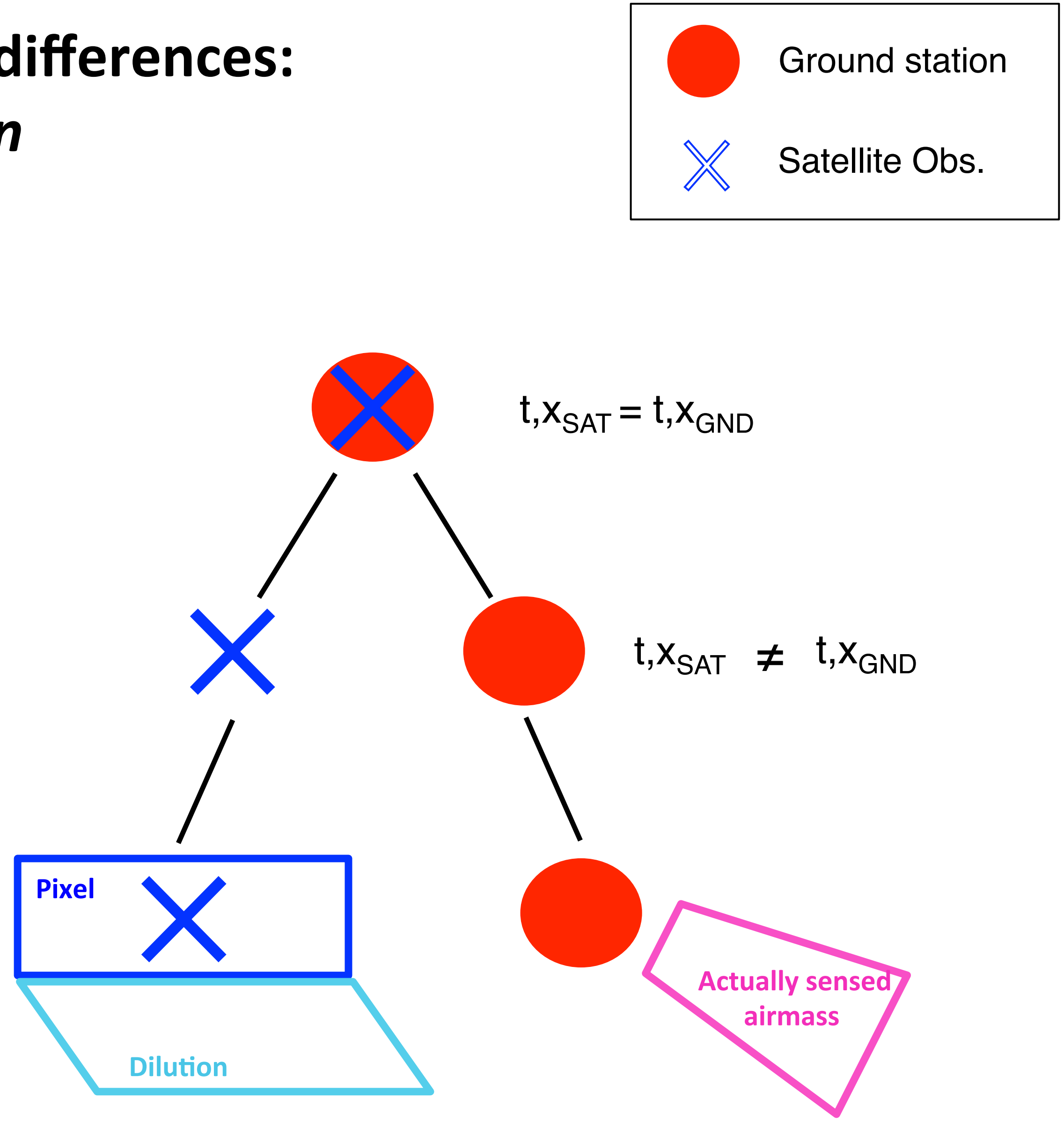
Combined measurement uncertainty versus actual median and spread of relative difference



If the measurement uncertainty assessments are sound and accurate, where do the discrepancies come from?
What is the influence of space/time mismatch errors on the comparison results?

This is the key topic of GAIA-CLIM WP3.

Sampling and smoothing differences: ***Conceptual representation***



Ideal case: coinciding measurements sensing the same airmasses

2D co-location mismatch: temporal and spatial offset

4D sampling and smoothing differences

Typical scale: ± 100 km

Sampling and smoothing differences:
Conceptual representation

WP2

$$\left| m_1 - m_2 \right| < \sqrt{u_1^2 + u_2^2}$$

$$\left| m_1 - m_2 \right| < \sqrt{\sigma^2 + u_1^2 + u_2^2}$$

Where σ includes:

- differences in vertical sampling
- differences in vertical smoothing
- differences in prior information

Hopefully mitigated by using X_a and vertical AKs (*)

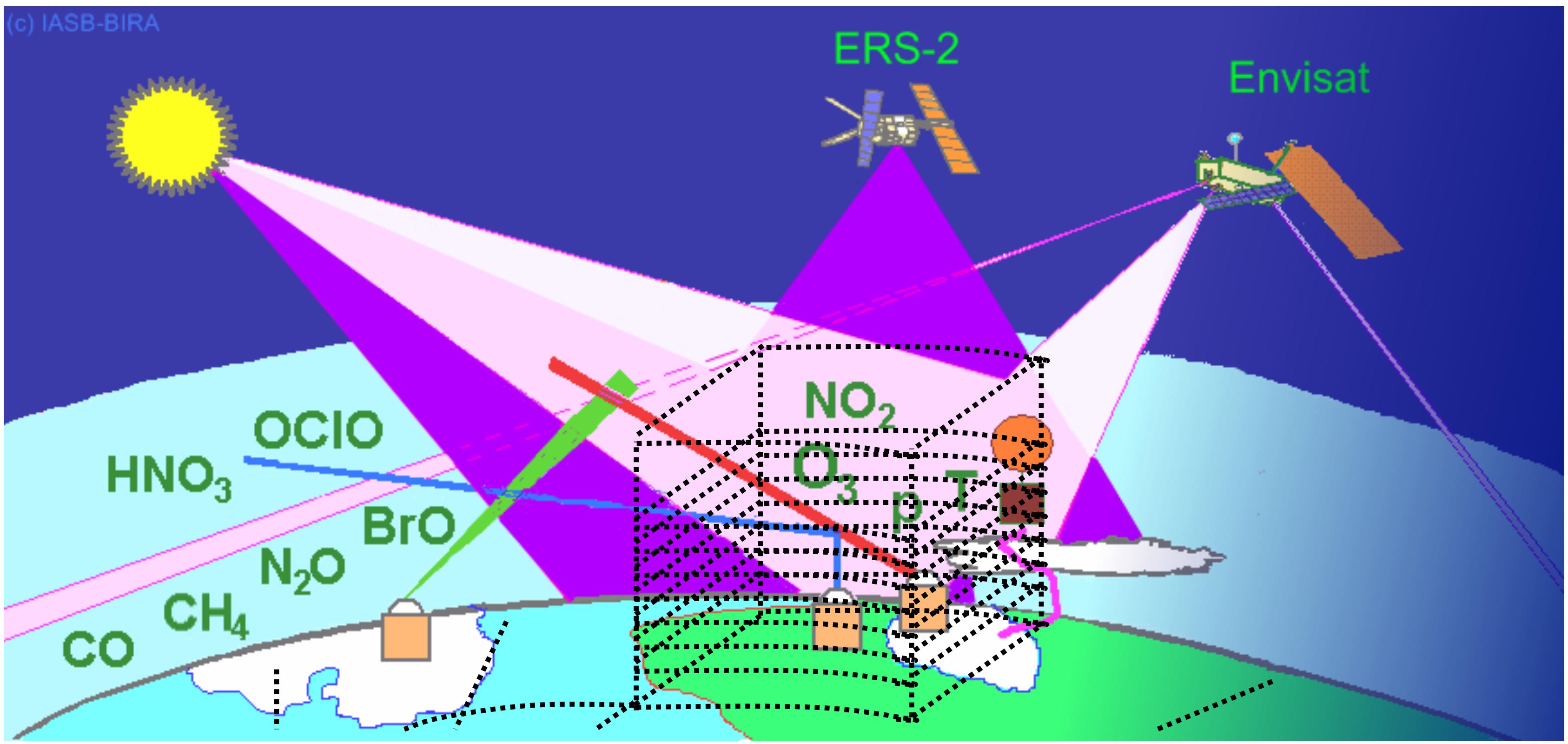
- differences in temporal sampling
- differences in temporal smoothing
- differences in horizontal sampling
- differences in horizontal smoothing

WP3

(*) **See Keppens et al., AMT, 2015 or the QA4ECV DPM for protocols**

Sampling and smoothing differences: *Variety in measurement techniques and platforms*

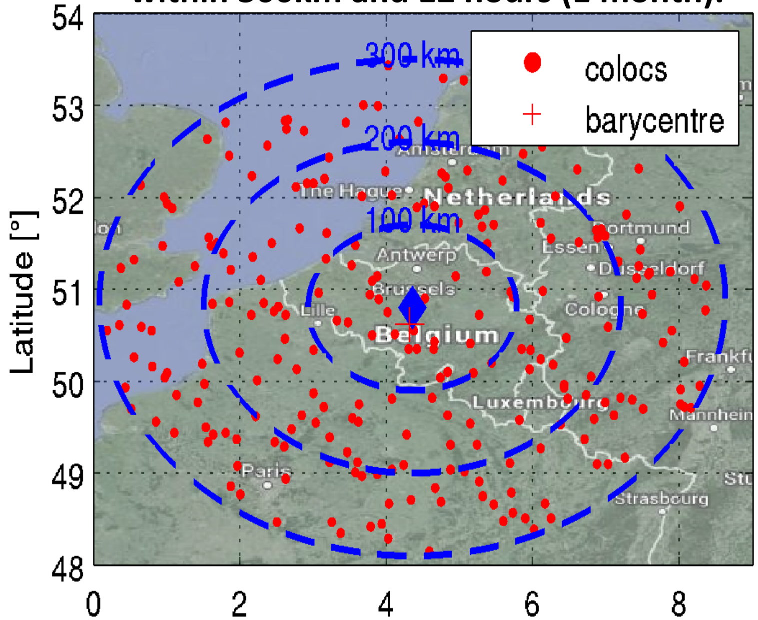
- *Sensing methods:* passive remote (e.g. backscatter UV satellite), active remote (e.g. lidar), in situ (e.g. sonde)
- *Viewing geometry:* limb, nadir, zenith sky, direct sun,…
- *Measurement time:* sunrise/sunset, sun-synchronous orbits, launch time,…



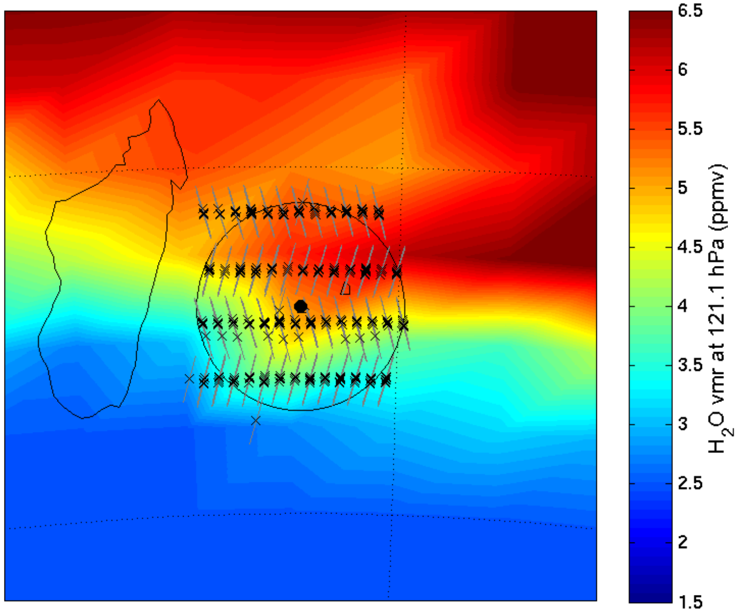
Sampling and smoothing differences:
co-location requirements and sampling properties

Co-location criteria must represent an acceptable compromise between:
(1) closest spatio-temporal coincidence, to reduce mismatch erros, and
(2) largest amount of comparison pairs, to enable meaningful statistical analysis.

Co-locations between GOME-2/MetOp-A and ozone sondes launched from Uccle, Belgium, within 300km and 12 hours (1 month).



MIPAS daily sampling around Reunion Island

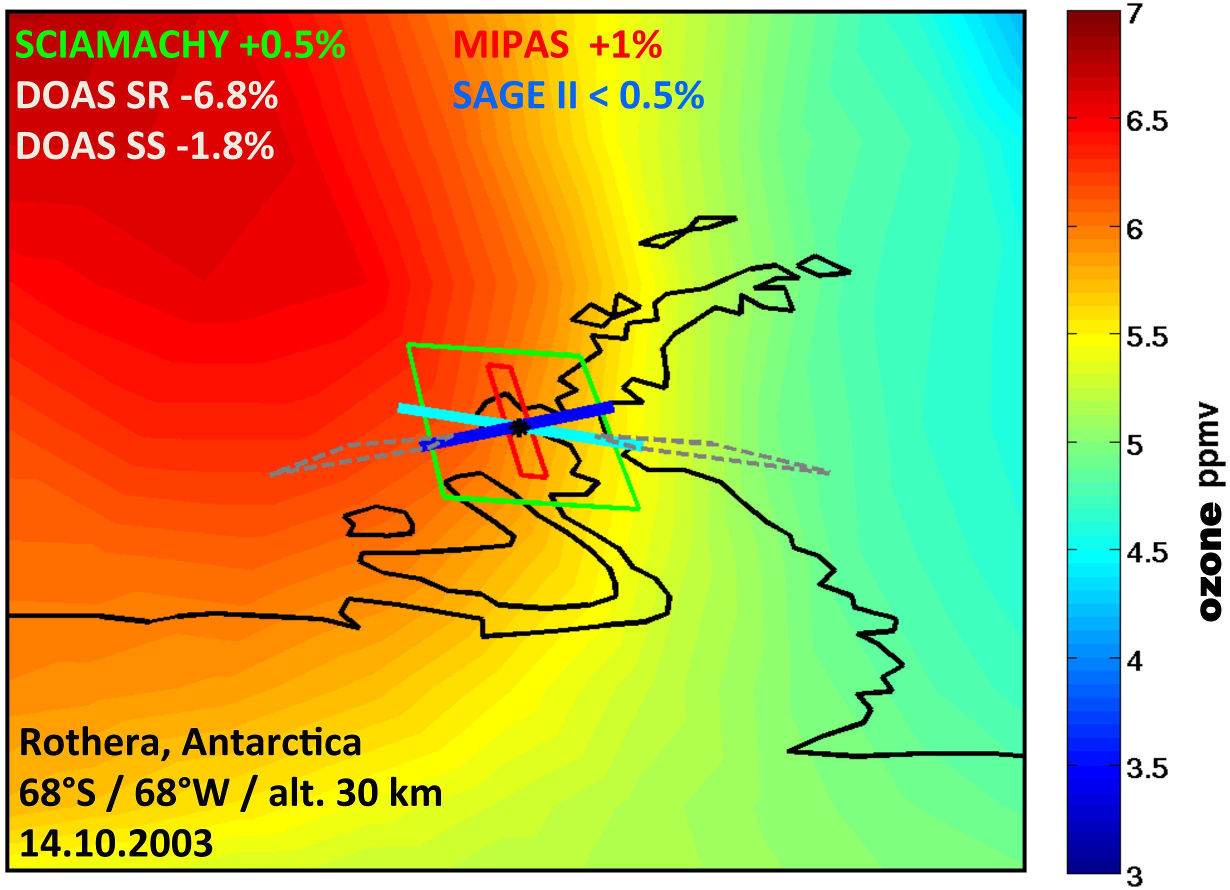


Lambert et al, ISSI Book on Atmospheric Water Vapour, Chapter 10, 2012

Sampling and smoothing differences: *smoothing properties*

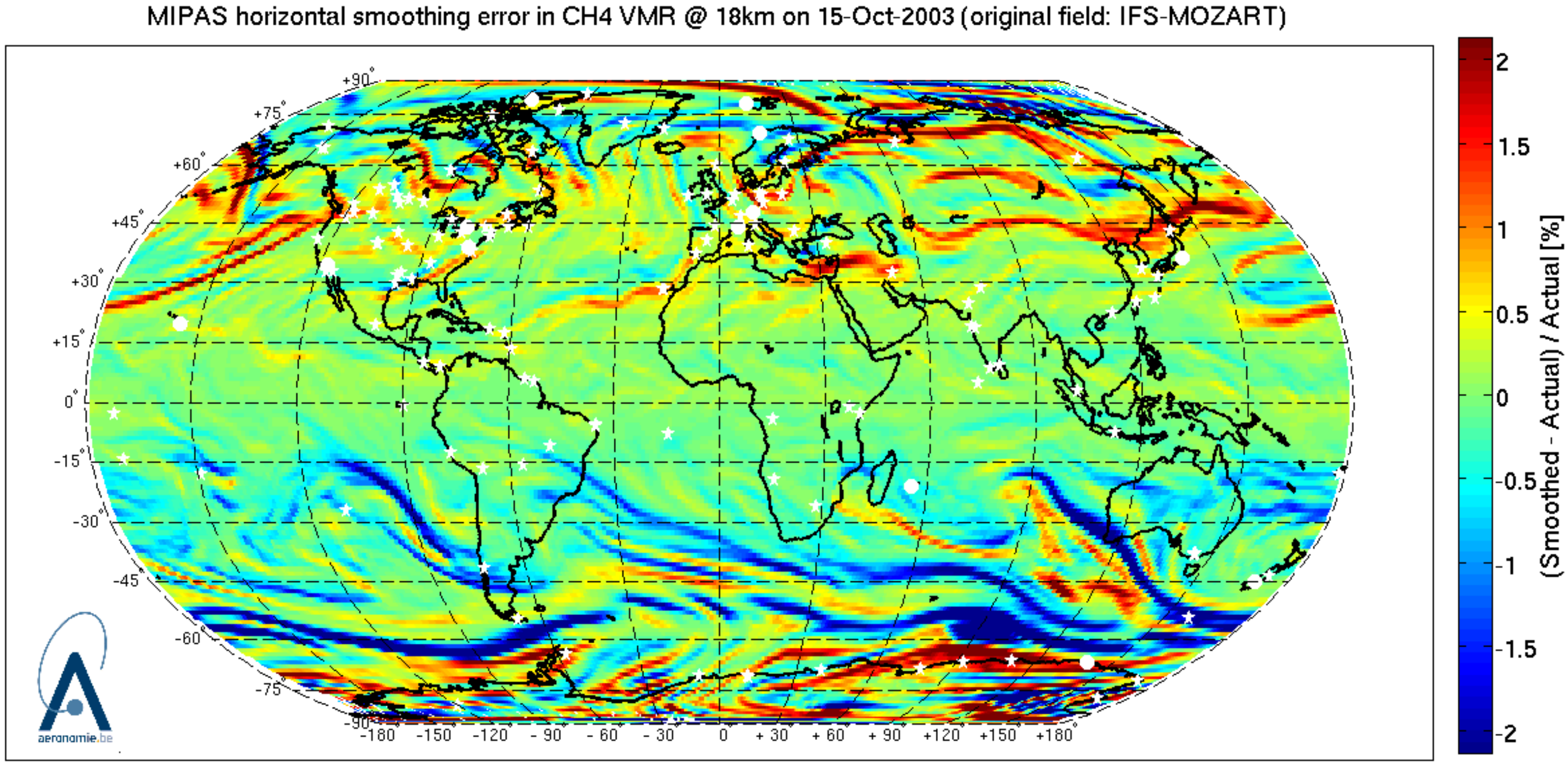
Even where pixel centre and station location do coincide perfectly, and when measurement times are identical,

different measurement techniques still sample different airmasses.

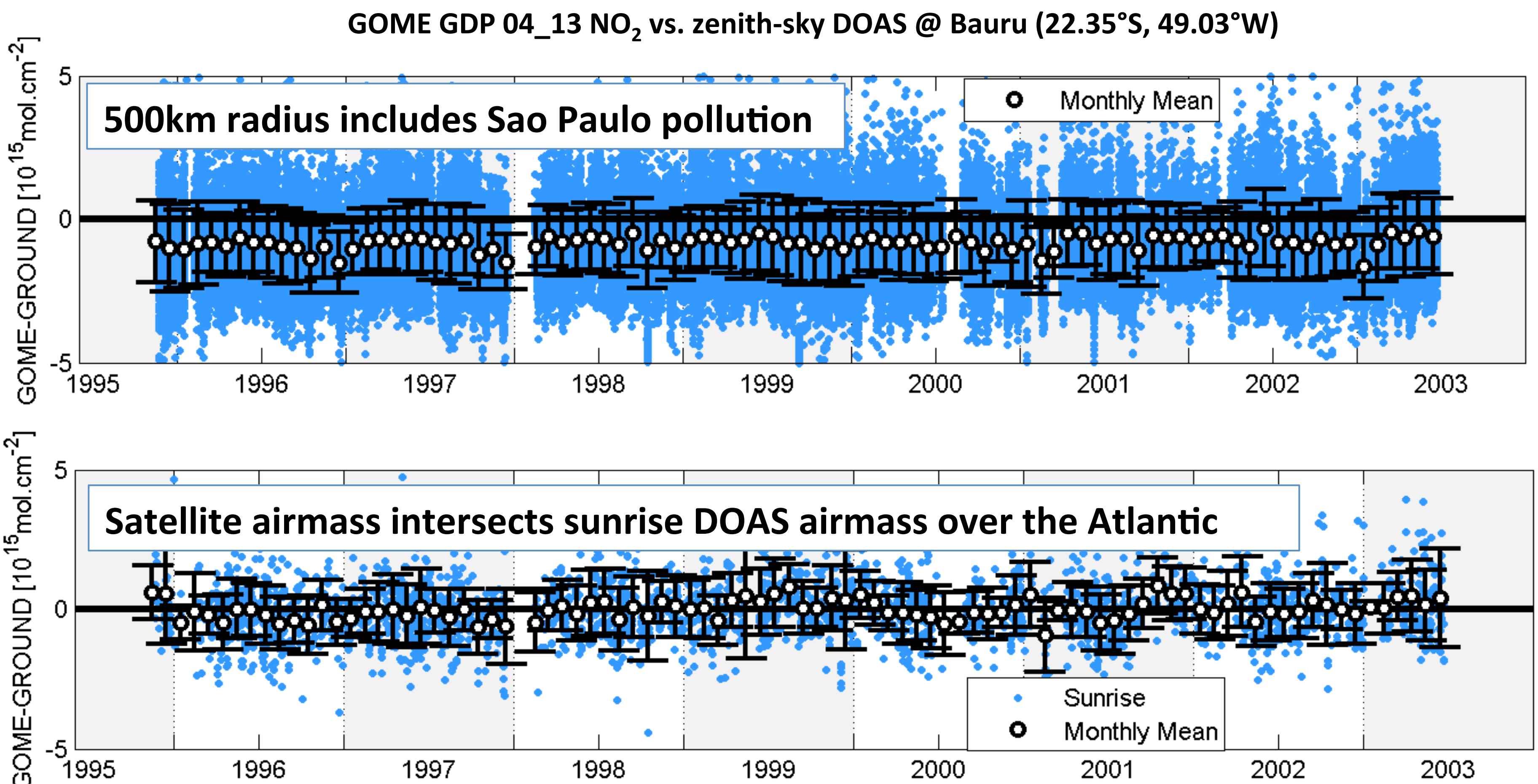


Vandenbussche et al., Lambert et al., GEOmon TNs, 2011

Sampling and smoothing differences:
smoothing properties



Mitigation: optimised co-location criteria



WP3: Four approaches to quantify mismatch errors

Empirical: Data driven

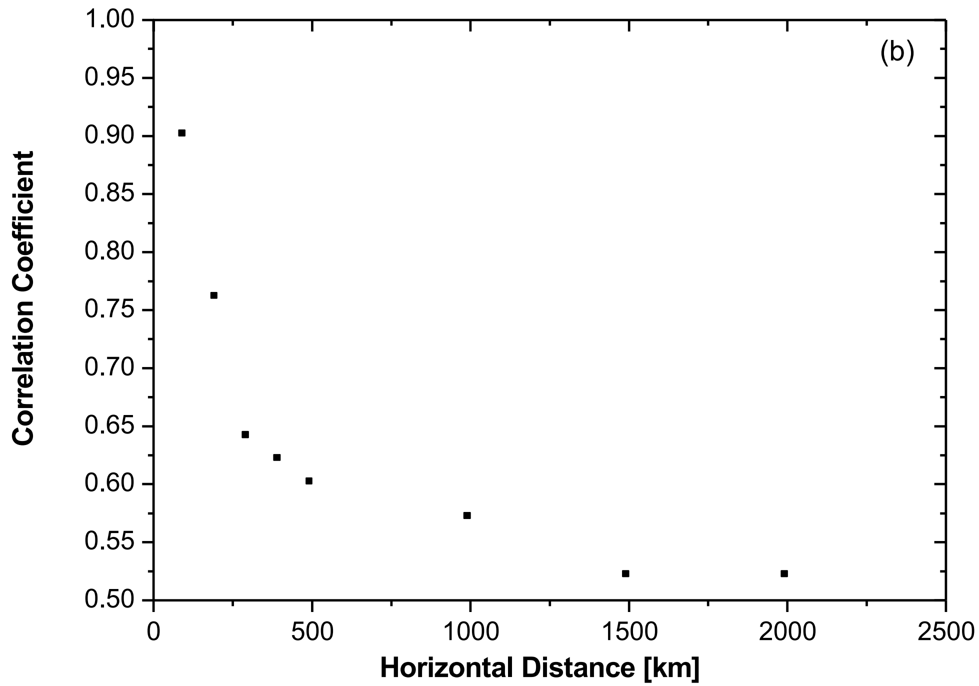
Explicit: Physics & Model driven

- Purely empirical: Data driven study of co-location dependences. E.g.: Aerosols from Calipso and Earlinet.
- HFRM: Heteroskedastic Functional Regression Model, using meaningful covariates to construct a statistical model at the scale of the comparisons. E.g. meteorological parameters from radiosondes.
- Parameterizations based on modelled fields and satellite sampling properties. E.g. sampling uncertainty of zonal means of O_3 profiles.
- OSSSMOSE: Explicit simulation of Observing System (of Systems) properties and errors, applying (1) multi-dimensional observation operators (2) set up with observation system(s) metadata (3) onto modelled atmospheric fields . E.g. trace gases from satellite and ground-based instruments.

Complementary + consistency checks

1. Empirical co-location dependence

Correlation coefficient between CALIPSO and EARLINET backscatter counts distributions as a function of the maximum considered horizontal distances between the 2 observations. From ESA project "Aerosols and Clouds: Long-term Database from Spaceborne Lidar Measurements", courtesy of F. Madonna.



+ Independent of our (sometimes limited) understanding of underlying physics

- Requires "large" data sets

- Difficult to separate sampling, smoothing and measurement errors

2. Heteroskedastic Functional Regression Model (HRFM)

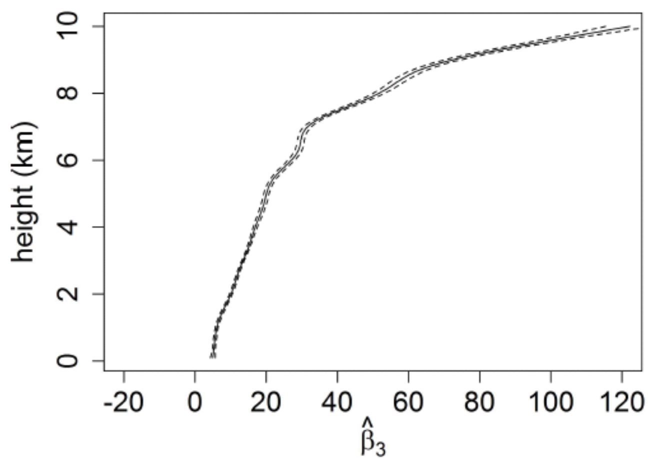
Modelling the differences of inter-comparisons with a locally linear model, which is a function of measured covariates and allows for varying measurement and environmental uncertainties: $\Delta\mu = \beta(h)'\, x(h) + \omega(h)$.

Case study: radiosonde relative humidity comparisons (Fassò et al., 2013).

$$\Delta_{rh}(h) = 3.40 + \underline{\beta_1(h) rh^0(h) + \beta_2(h) mr^0(h) + \beta_3(h)\Delta_{mr}(h)} + \underline{\omega(h)} + \underline{\Delta_\varepsilon(h)}$$

86 % 11% 0.2%

(Δmr)



- \+ More or less independent on our understanding of underlying physics
- \+ Allows separation of metrological and measurement errors, on per pair basis
- \- Requires "large" data sets to limit model sampling error
- \- Requires meaningful covariate observations

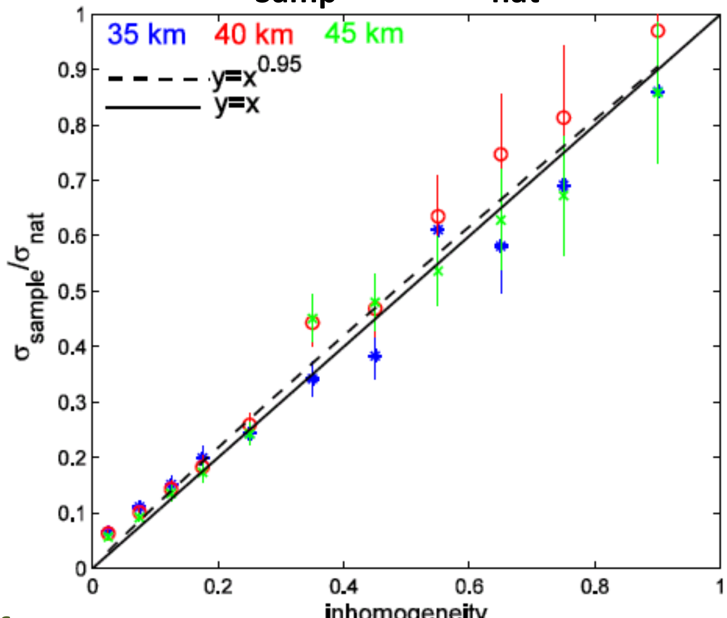
3. Parameterization using modelled fields

Sampling uncertainty of gridded ozone profile monthly means (Sofieva et al. 2014):

•Parametrization of the measurement inhomogeneity ***H*** as a combination of asymmetry ***A*** and entropy ***E***: $H = \frac{1}{2}(A + (1 - E))$

•Parametrization of natural variability as a climatological variance: σ_{nat}

$$\sigma_{Samp} = H \cdot \sigma_{nat}$$



\+ Computationally efficient

\+ Linked to physical properties of the observing system and atmosphere

\- Depends on knowledge of atmospheric variability

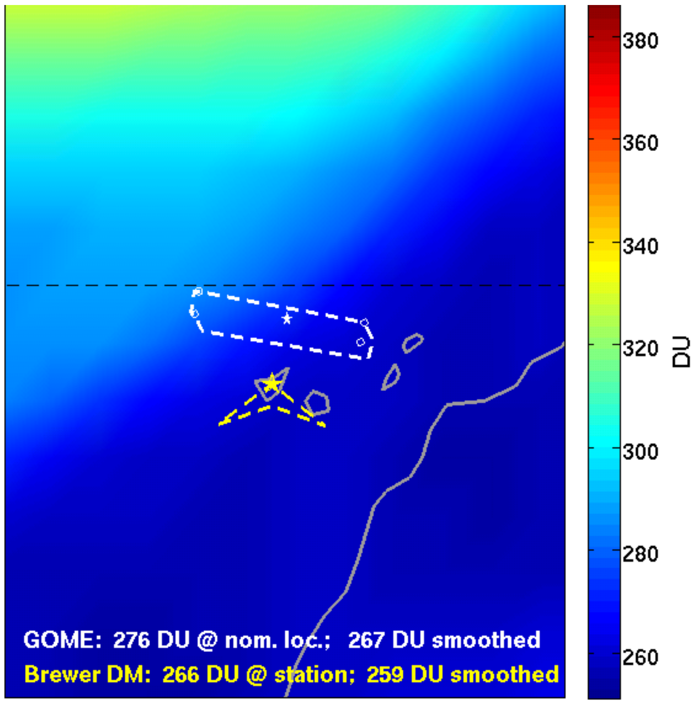
\- Valid only in terms of variances

4. Explicit simulation of the observing system (OSSSMOSE)

Error budget closure of total ozone column validation (Verhoelst et al., AMTD, 2015):

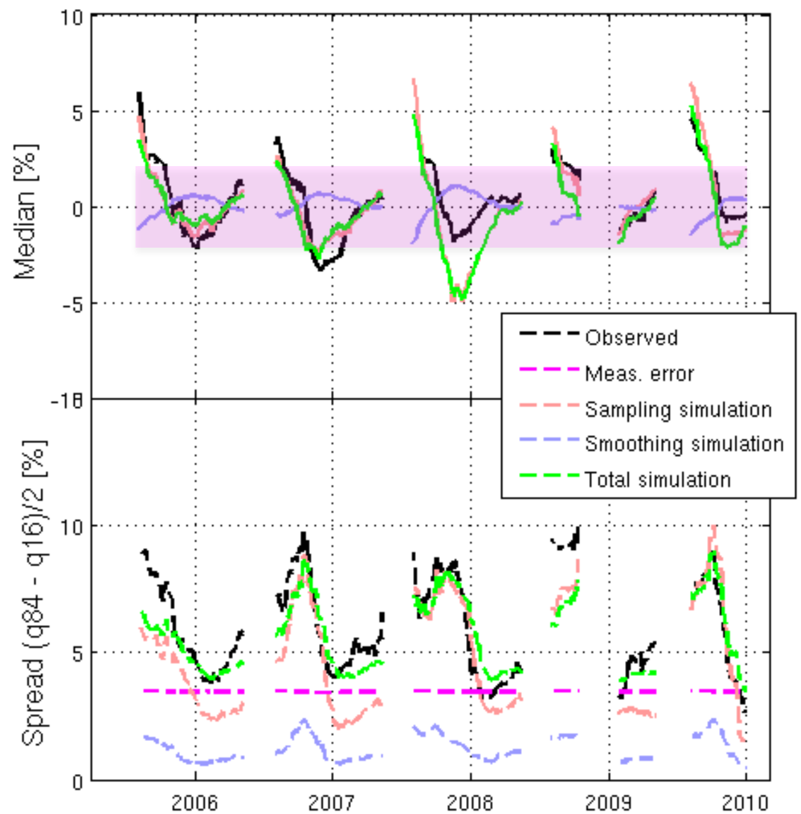
•Simulate individual measurements and their differences by applying observation operators, set up by observation metadata, onto modelled fields.
•Decomposition of the error budget in terms of smoothing, sampling, and measurement errors.

Measurement simulation on IFS-MOZART reanalysis fields



4. Explicit simulation of the observing system (OSSSMOSE)

Relative difference between ERS-2 GOME GODFIT v3 and NDACC zenith-sky DOAS total ozone at Dumont d'Urville (Antarctica, 67°S)



- \+ Detailed results, at the level of individual pairs
- \+ Allows detailed physical interpretation and understanding
- \- Computationally demanding
- \- Requires high-resolution and accurate models

4. Explicit simulation of the observing system (OSSSMOSE)

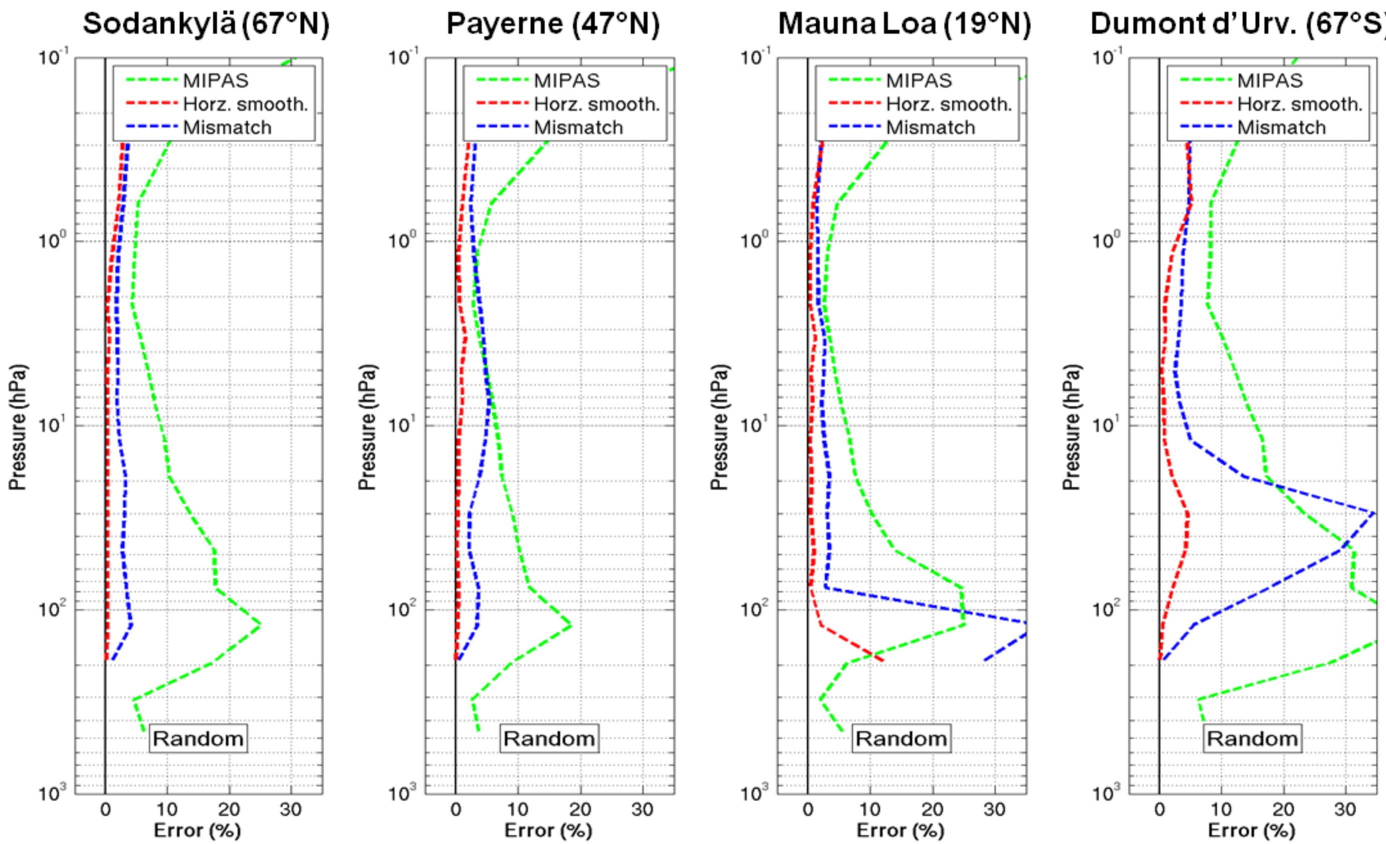


Figure 10.7 – Vertical profile of the total random error estimate for MIPAS water vapour retrievals (in green), and estimates of the random component of comparison errors (e.g. between MIPAS and sondes) due to differences in horizontal smoothing (in red) and to horizontal co-location mismatch (in blue). From left to right: Sodankylä year-round (Finland, 67°N, 11°E), Payerne (Swiss Alps, 48°, 7°E) in October-December, Mauna Loa (Hawaii, 19°N, 155°W) in April-June, and Dumont d'Urville (Antarctica, 67°S, 140°E) in July-September 2003.

From Lambert et al, ISSI Book on Atmospheric Water Vapour, Chapter 10, 2012

Aims within GAIA-CLIM

Development:

- *Identify gaps in the state-of-the-art regarding measurement mismatch analyses (GAID input).*
- *Study the smoothing properties of the remote sensing measurements addressed by GAIA-CLIM.*
- *Study the co-location uncertainties in demonstrator comparison experiments and derive implications on suitable co-location criteria (T, q, ozone and aerosols are first on the list)*
- *Provide reference-quality documentation on measurement mismatch issues, how to mitigate them, and how to quantify them.*
- *Develop tools for implementation in the virtual observatory.*

Offered through the Virtual Observatory:

- *Educational material to raise attention of users on multi-dimensional issues*
- *Recommendations for appropriate co-location criteria*
- *Look-up tables, (time series) graphs, and (global) maps to serve as guidance for the user in (1) adjusting co-location criteria for his/her specific analysis aims, and (2) interpreting the comparison results.*
- *Data-driven tools to assess the impact of measurement mismatch on the datasets co-located using the VO.*