Ozone-CCI: on uncertainty estimates and their validation

The Ozone_cci team

Outlines

- **Uncertainties of Level 2 data**
  - geophysical validation of precision estimates
- **Uncertainties of Level 3 data**
  - Importance of characterization of sampling uncertainties
- **Experiences and outlook**

Geophysical validation of precision estimates in Level 2 data

- **It is important**
  - Error estimates are based on propagation of instrumental and other uncertainties through the inversion algorithm can be not perfect (due to different approximations)
- **It is difficult:**
  - Always changing atmosphere
- **Main strategy in classical methods:**
  - Statistics of differences in a set of collocated measurements x_1 and x_2 :

$$\underbrace{s_{12}^2 = \left\langle \left(x_1 - x_2 \right)^2 \right\rangle - \left\langle x_1 - x_2 \right\rangle^2}_{\text{Sample variance of differences}} = \underbrace{\sigma_{0,nat}^2}_{\text{Natural variability}} + \underbrace{\sigma_1^2 + \sigma_2^2}_{\text{Data uncertainties}}$$

Advanced precision validation through evaluation of the structure function

$$s_{12}^2(\mathrm{r}) = \sigma_{0,nat}^2(\mathrm{r}) + 2\sigma^2$$

Sample variance of differences

Natural variability

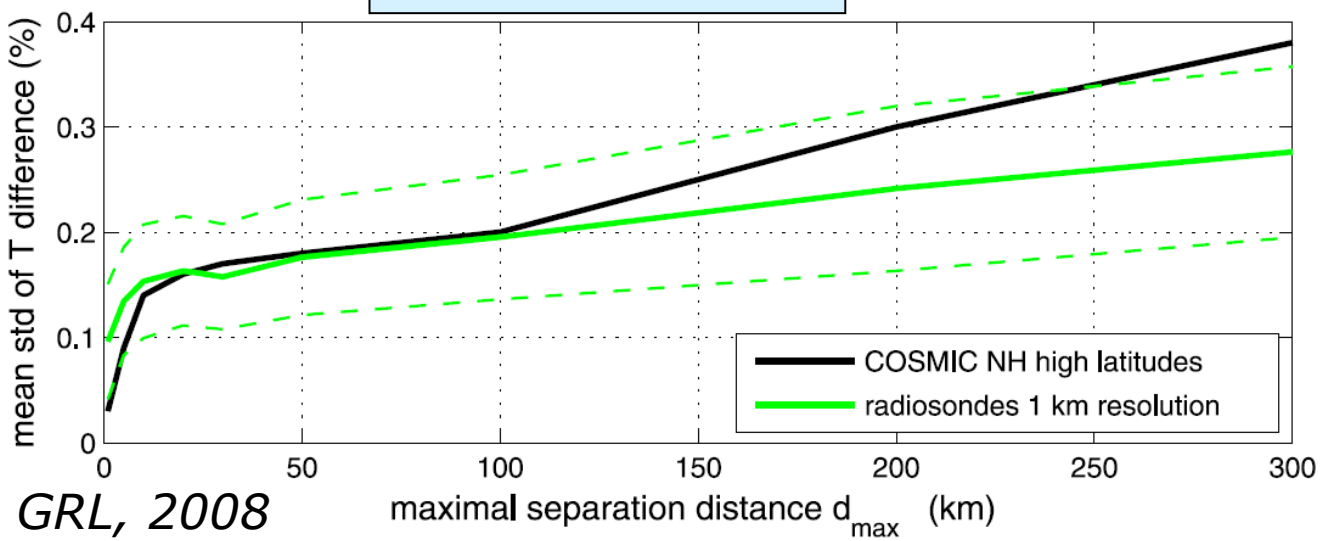
Data uncertainties

"experimental" precision estimate

$$\frac{s_{12}}{\sqrt{2}} \rightarrow \sigma \text{ as } r \rightarrow 0$$

Error estimate

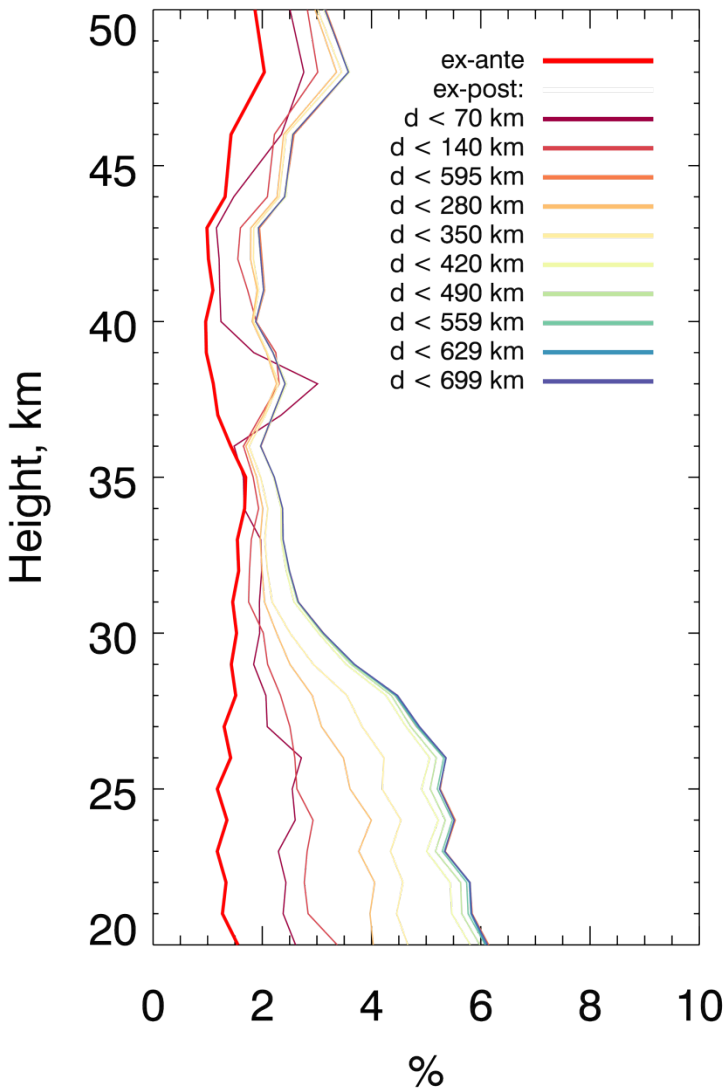
This is the classical definition of the structure function used in the theory of random processes



Sofieva et al., GRL, 2008

Application to limb ozone profiles

MIPAS

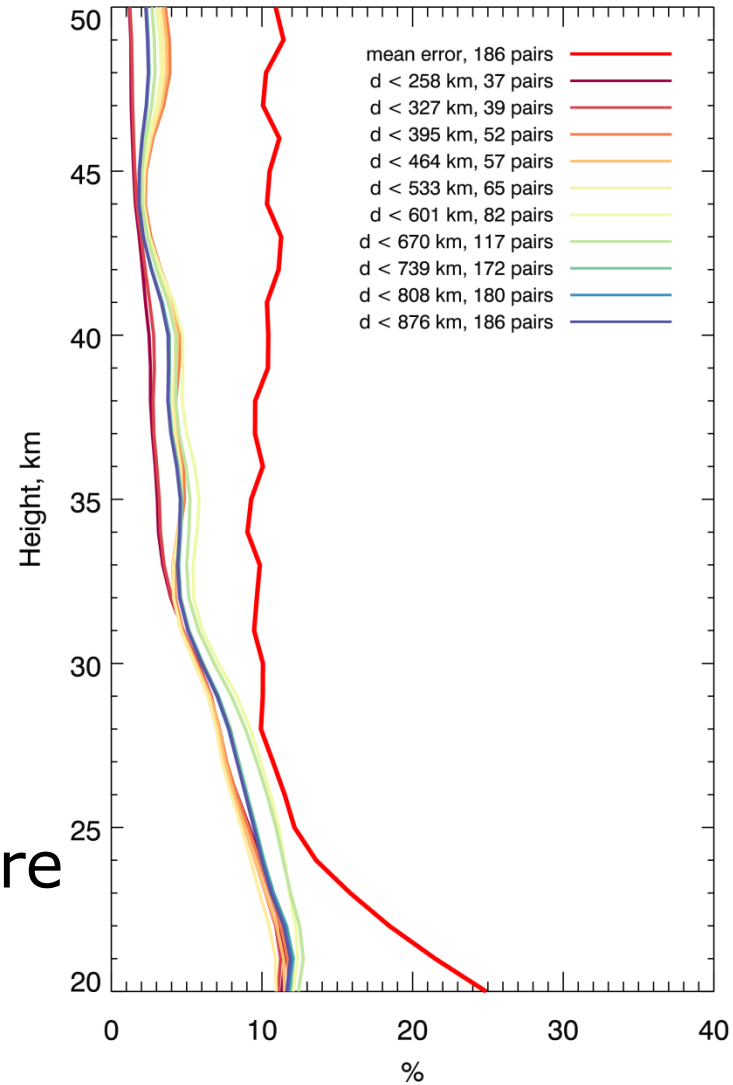


South Pole,
local summer

MIPAS: convergence to uncertainty estimates

SCIAMACHY: uncertainties are clearly overestimated

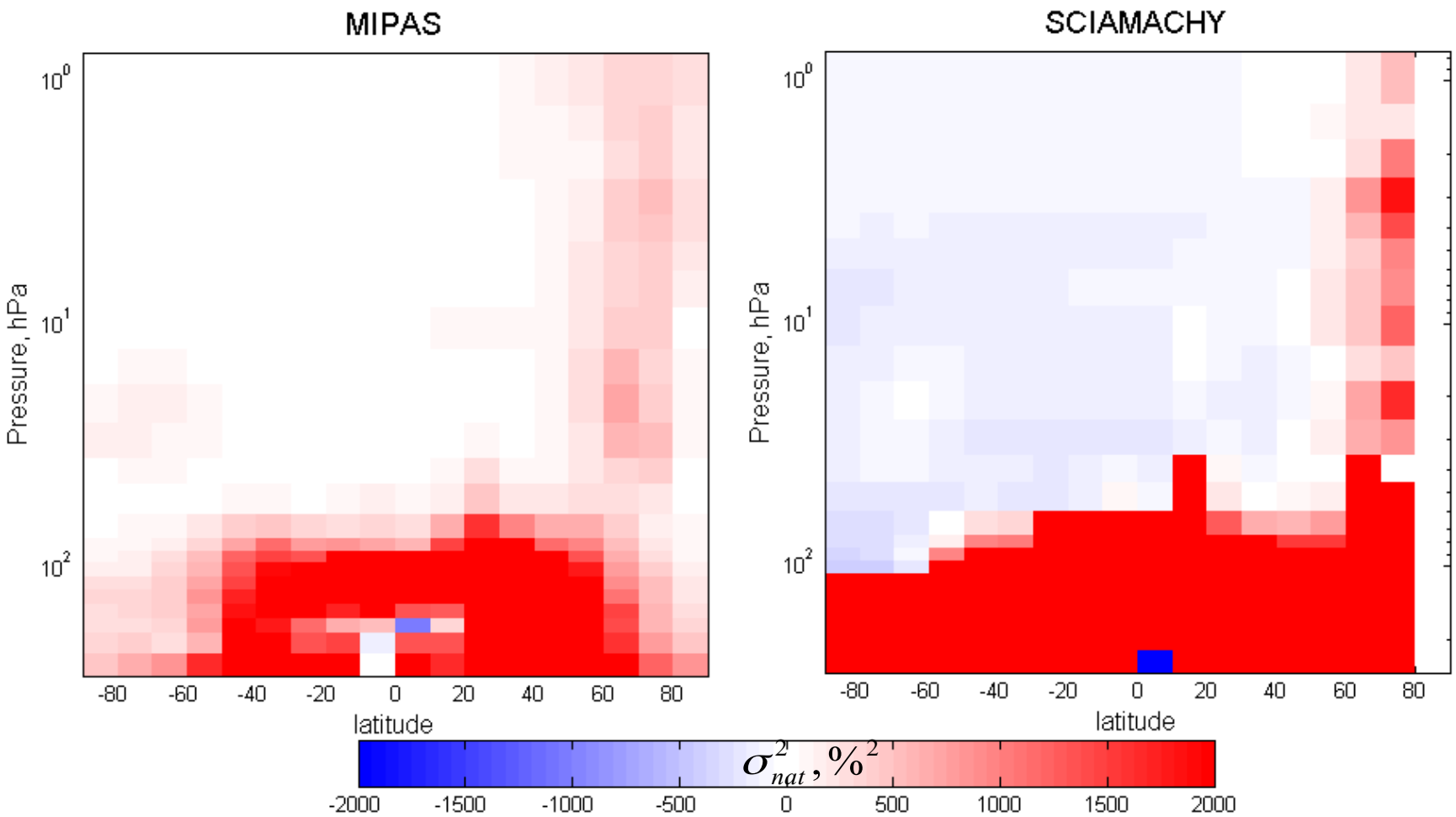
SCIAMACHY



A simple verification of overestimated uncertainties

$$\mathrm{var} = \left\langle (x - \langle x \rangle)^2 \right\rangle \approx \sigma_{nat}^2 + \sigma_{noise}^2 \longrightarrow \sigma_{nat}^2 = \mathrm{var}(x) - \sigma_{noise}^2$$

sample variance — Natural — precision



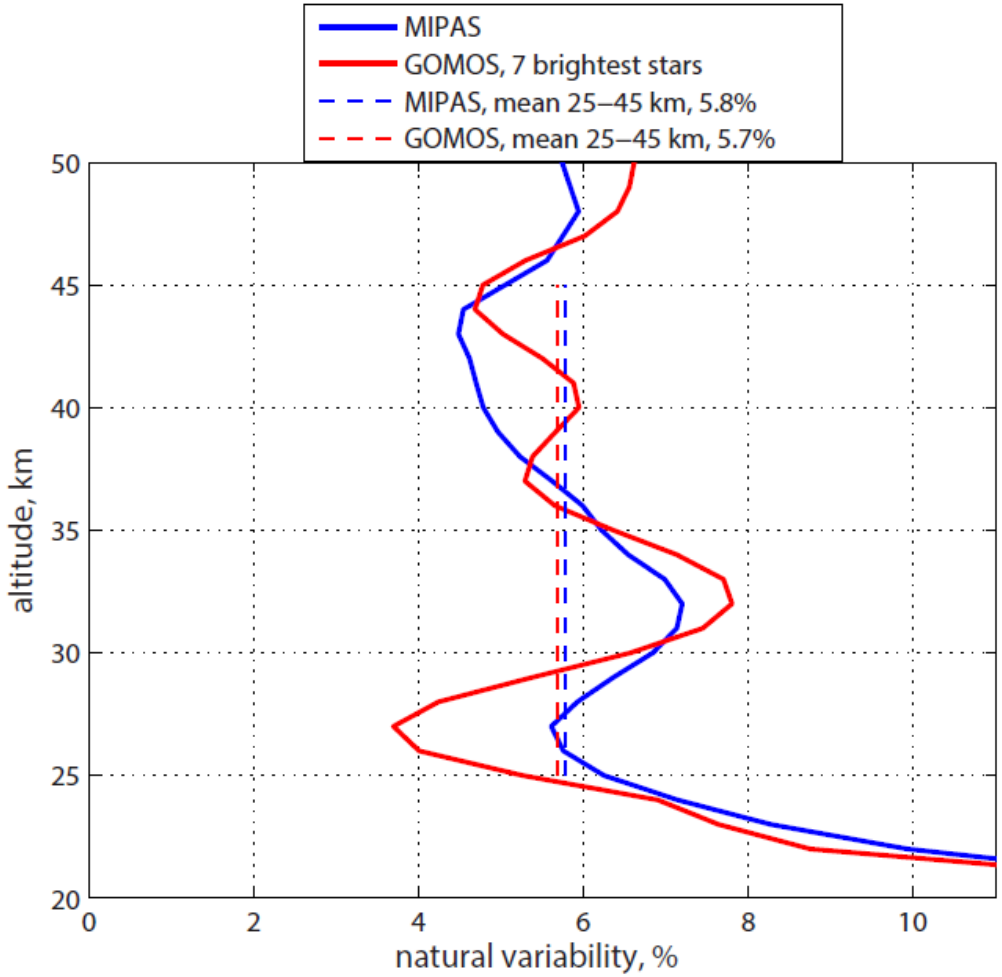
Multi-instrument approach

- **Based on comparison of natural variability estimates**

$$\sigma_{nat}^2 = \mathrm{var}(x) - \sigma_{noise}^2$$

Perfect agreement in ozone natural variability estimates in tropics 20S-20N in 2008 for GOMOS and MIPAS/IMK

Mean at altitudes 25-40 km:
GOMOS: 5.7%
MIPAS: 5.8 %



Sofieva et al., AMT, 2014a

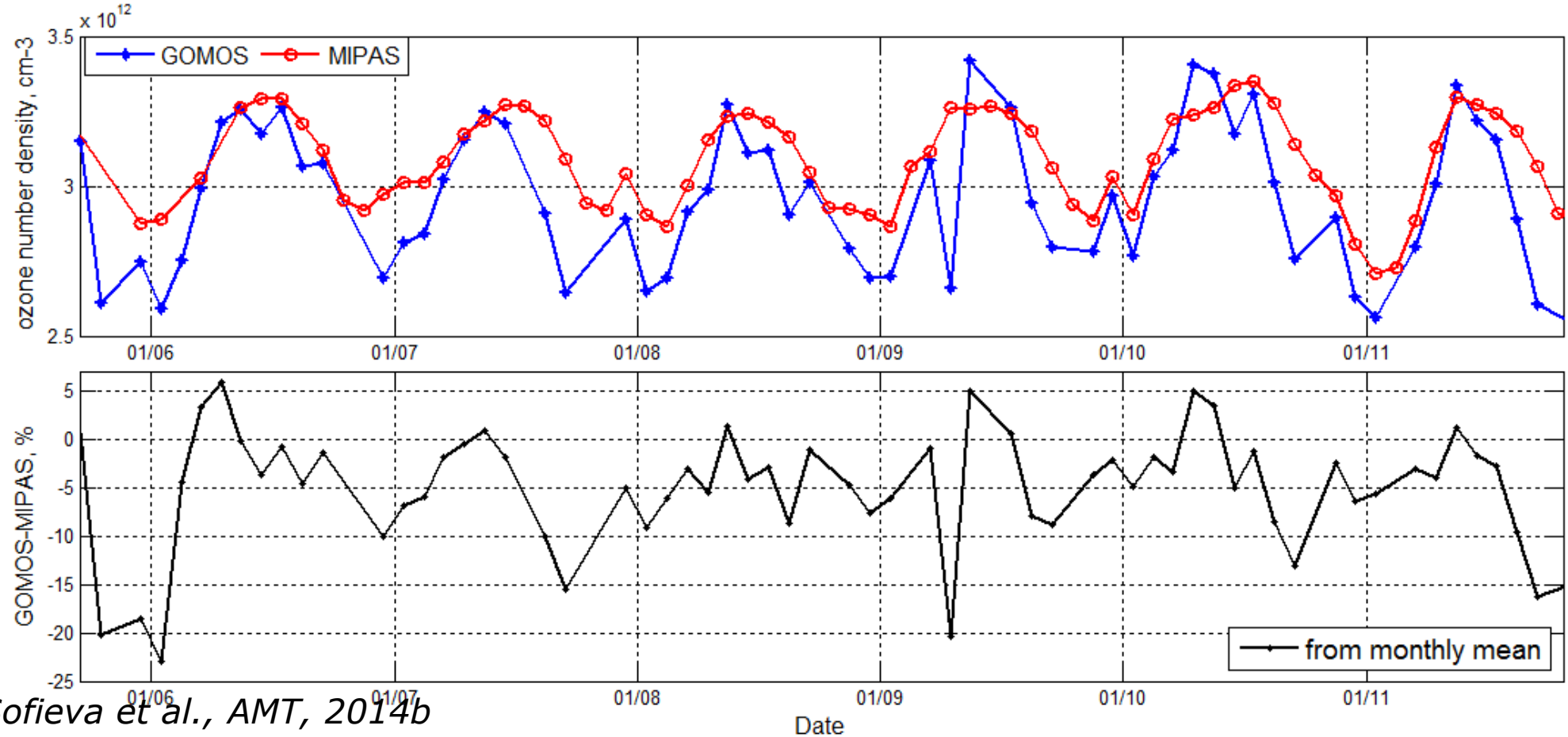
Level 3

- **Standard error of the mean does not fully characterize the Level 3 data uncertainty**
- **Hidden uncertainty due to non-uniform sampling**
  - Insufficient or inhomogeneous sampling, if not taken into account, can result in
    - inaccurate average estimates
    - and even induce spurious features…

Representation of annual cycles

Ozone profiles from MIPAS and GOMOS instruments on board Envisat

Monthly zonal mean ozone number density at 15 hPa (~ 30 km), latitudes 40°N-60°N, years 2005-2011

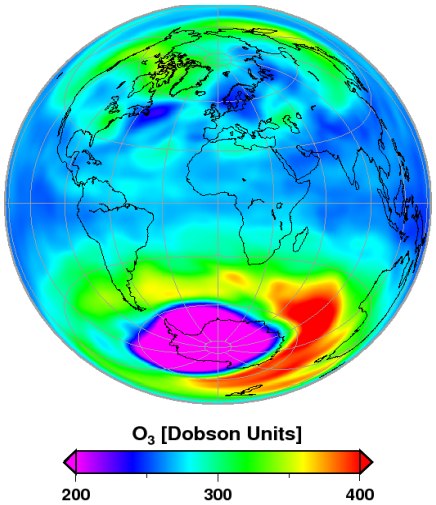


Sofieva et al., AMT, 2014b

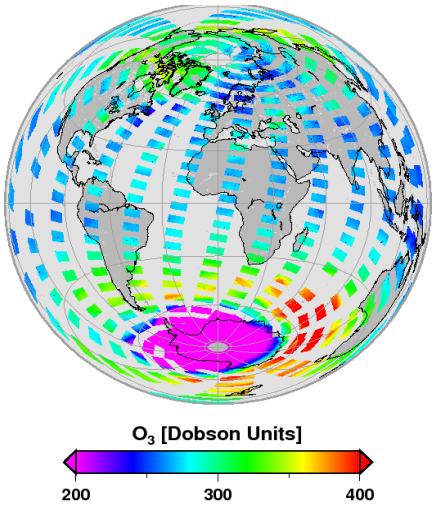
Total Ozone Uncertainty

GTO-ECV sampling errors are characterized using an Observing System Simulation Experiment (OSSE)

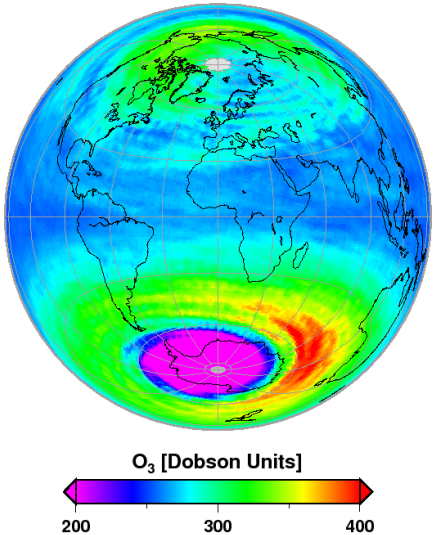
High-resolution ECMWF data



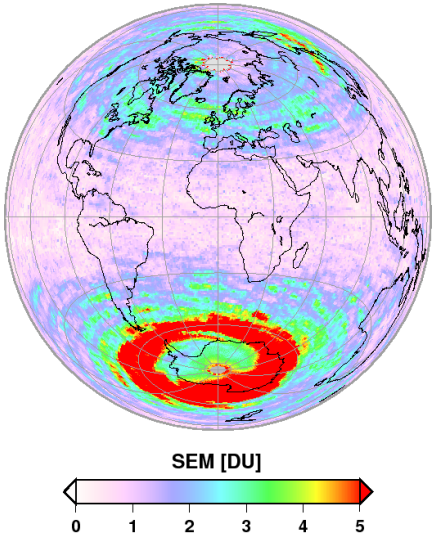
SCIAMACHY sampling



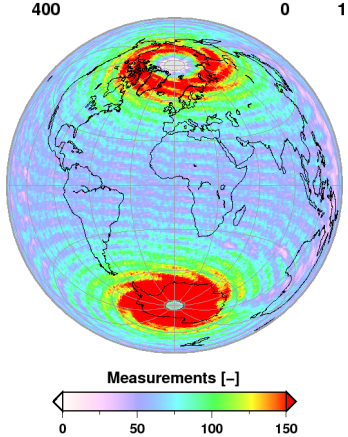
SCIAMACHY monthly mean



Standard error of the mean



SCIAMACHY number of measurements per month



DLR

How to handle sampling uncertainties?

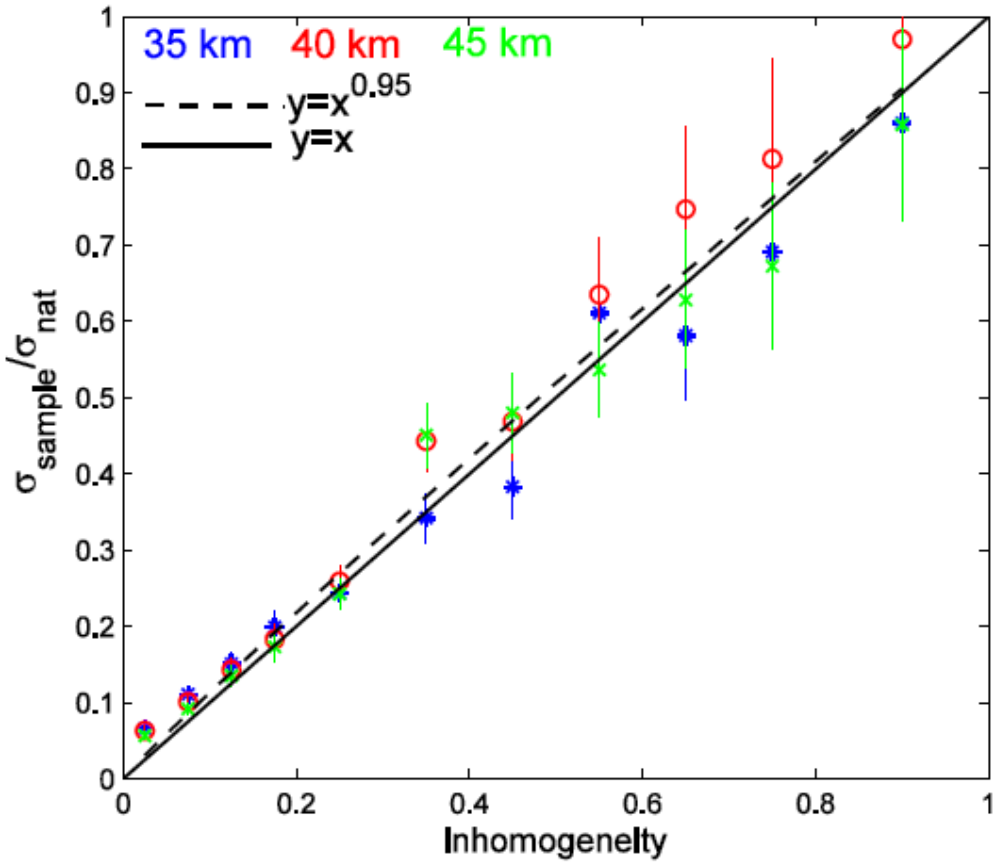
- **Estimate using atmospheric model and correct in a deterministic way?**
  - For ozone, such approach is not possible due to the absence of a sufficiently accurate CTM
  - furthermore, such correction would be incomplete and might induce some additional problems
- **Do not use data with insufficiently dense sampling?**
  - Related issue: inhomogeneity measure
- **Characterize the sampling uncertainty?**
  - Recently published paper in AMT
  - Sampling uncertainty is a random variable with std σ_{sample}

Parameterization of sampling error in monthly zonal mean data

- **Inhomogeneity measure**

$$H = \tfrac{1}{2}(A + (1 - E))$$

  - A is asymmetry, E is Shannon entropy
  - *H* ranges from 0 (homogeneous) to 1 (highly inhomogeneous)
- **Simulations with a realistic chemistry-transport model**
  - Comparison of "full" and sub-sampled data field
  - Observation: enhancements of sampling uncertainty for coarse sampling patterns and for regions of high ozone variability.
- **Found dependence** $\sigma_{sample} = \sigma_{nat} \cdot H_{tot}$



Sofieva et al., AMT, 2014b

Outlook

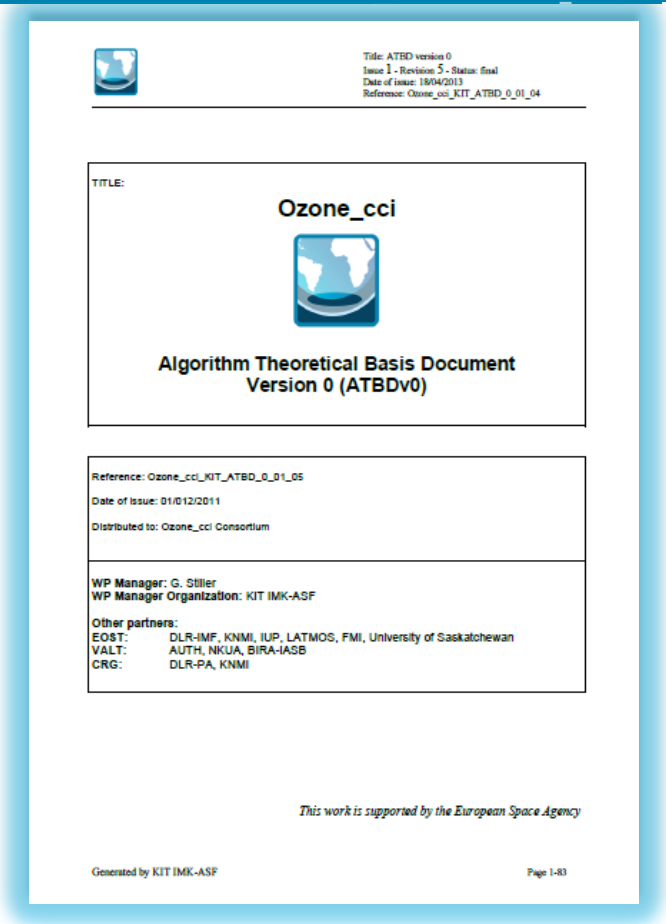
- **It is possible to validate the precision estimates of Level 2 data**
  - Geophysical validation of error bars has been performed for several instruments
  - Continuation with other instruments
- **Data representativeness should be taken into account in creating Level 3 data**
  - sampling uncertainty can be not negligible
  - Ozone_cci dataset include characterization of data inhomogeneity
  - Sampling uncertainty has been taken into account in Level 3 data merging
  - The proposed methods can be equally applied to the analyses of other (not only satellite) measurements

Back-up

A starting point: a common terminology

- **Data and Error Characterization Questionnaire was distributed to the team in the very beginning of the project**
- **Objective:**
  - compiling definitions and recipes with respect to data characterization and error budget
  - to find an optimal compromise between stringency of definitions and their practicability
- **Contents**
  - comprehensive range of the questions, e.g.
    - Retrieval grid
    - A priori
    - Regularization
    - Averaging kernel and characteristics of vertical resolution
    - Uncertainties (random, systematic), their characterization (covariance matrix, error estimates)
    - Etc.
  - initially designed for the retrievals of vertical profiles of atmospheric constituents

Practical output



Title: ATBD version 0
Issue 1 - Revision 5 - Status: final
Date of issue: 18/04/2013
Reference: Ozone_cci_KIT_ATBD_0_01_04

TITLE:

Ozone_cci

Algorithm Theoretical Basis Document
Version 0 (ATBDv0)

Reference: Ozone_cci_KIT_ATBD_0_01_05

Date of issue: 01/012/2011

Distributed to: Ozone_cci Consortium

WP Manager: G. Stiller
WP Manager Organization: KIT IMK-ASF

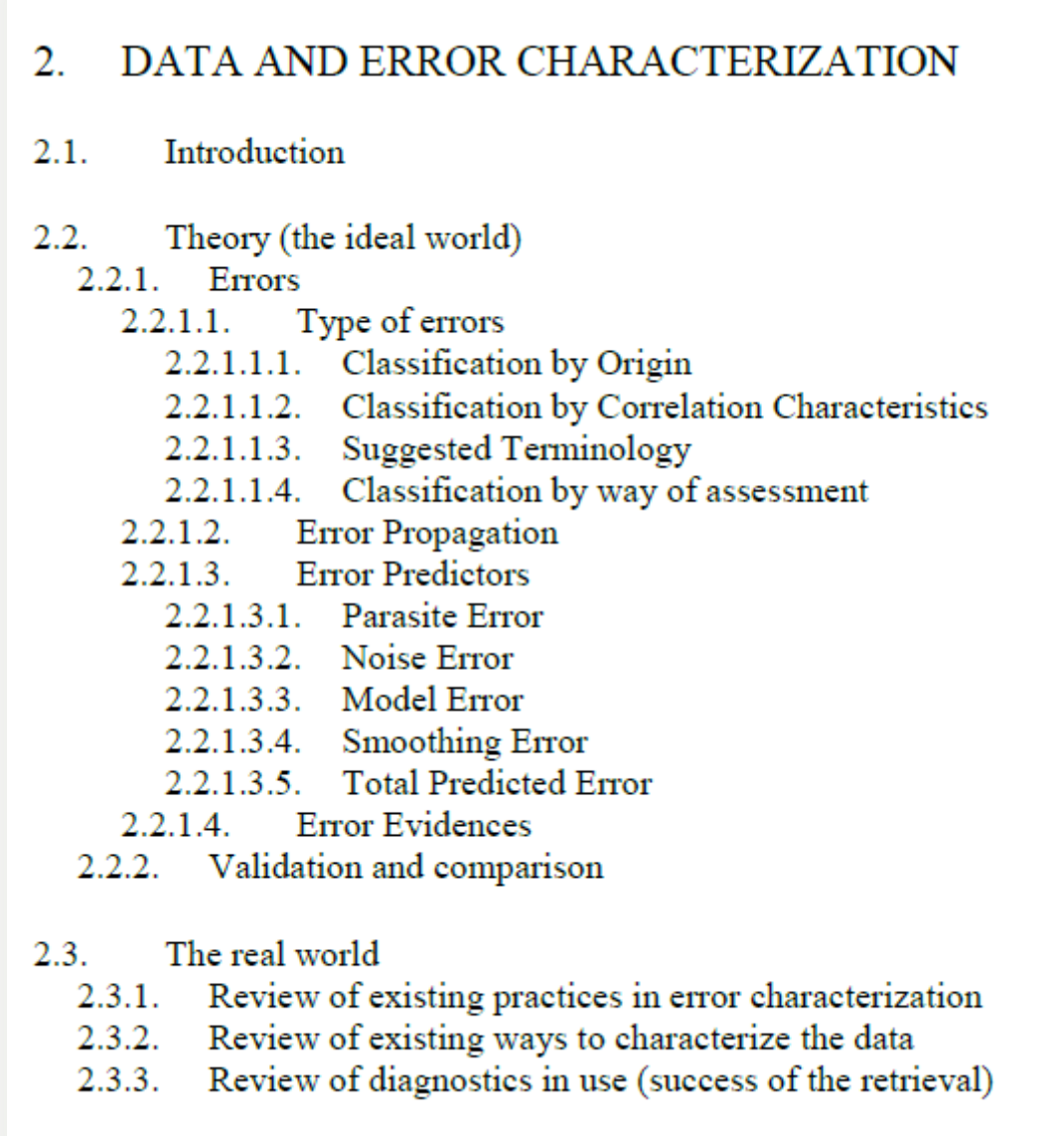
Other partners:
EOST: DLR-IMF, KNMI, IUP, LATMOS, FMI, University of Saskatchewan
VALT: AUTH, NKUA, BIRA-IASB
CRG: DLR-PA, KNMI

This work is supported by the European Space Agency

Generated by KIT IMK-ASF

Page 1-83

- **Classification of data uncertainties**
- **a common terminology, which seems to be clear to modelers**



2. DATA AND ERROR CHARACTERIZATION

2.1. Introduction

2.2. Theory (the ideal world)
- 2.2.1. Errors
  - 2.2.1.1. Type of errors
    - 2.2.1.1.1. Classification by Origin
    - 2.2.1.1.2. Classification by Correlation Characteristics
    - 2.2.1.1.3. Suggested Terminology
    - 2.2.1.1.4. Classification by way of assessment
  - 2.2.1.2. Error Propagation
  - 2.2.1.3. Error Predictors
    - 2.2.1.3.1. Parasite Error
    - 2.2.1.3.2. Noise Error
    - 2.2.1.3.3. Model Error
    - 2.2.1.3.4. Smoothing Error
    - 2.2.1.3.5. Total Predicted Error
  - 2.2.1.4. Error Evidences
- 2.2.2. Validation and comparison

2.3. The real world
- 2.3.1. Review of existing practices in error characterization
- 2.3.2. Review of existing ways to characterize the data
- 2.3.3. Review of diagnostics in use (success of the retrieval)

A classical approach: collocated measurements of the same instrument

$$s_{12}^2 = \cancel{\sigma_{0,nat}^2} + \sigma_1^2 + \sigma_2^2$$

Sample variance of differences

Natural variability

$$\sigma_1 = \sigma_2 = \sigma$$

$$\frac{s_{12}}{\sqrt{2}} \approx \sigma$$

"experimental" precision estimate

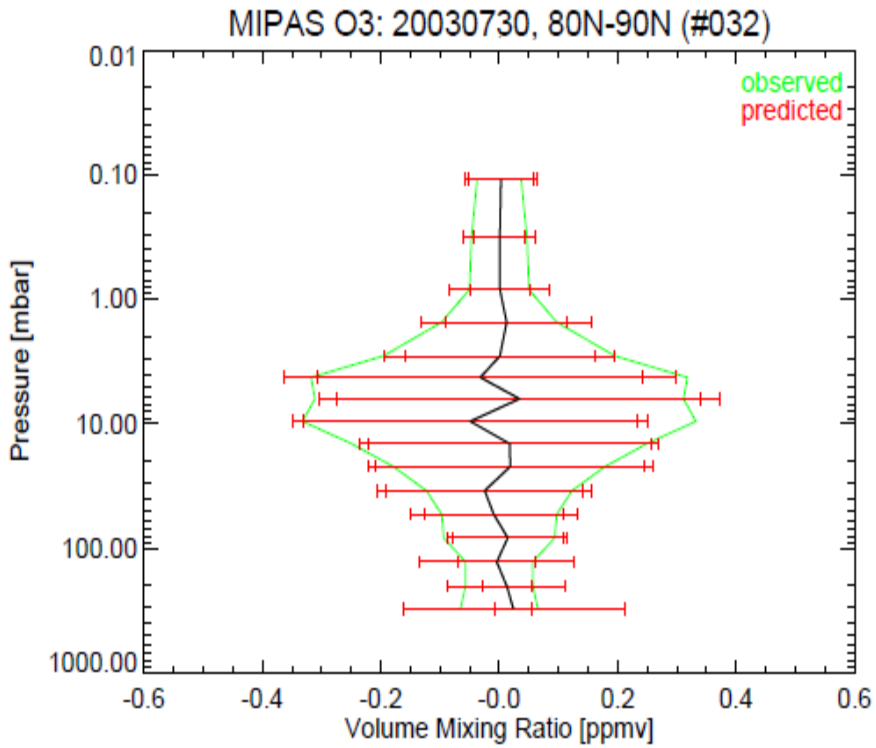
Error estimate

A classical approach: collocated measurements of the same instrument

$$\frac{s_{12}}{\sqrt{2}} \approx \sigma$$

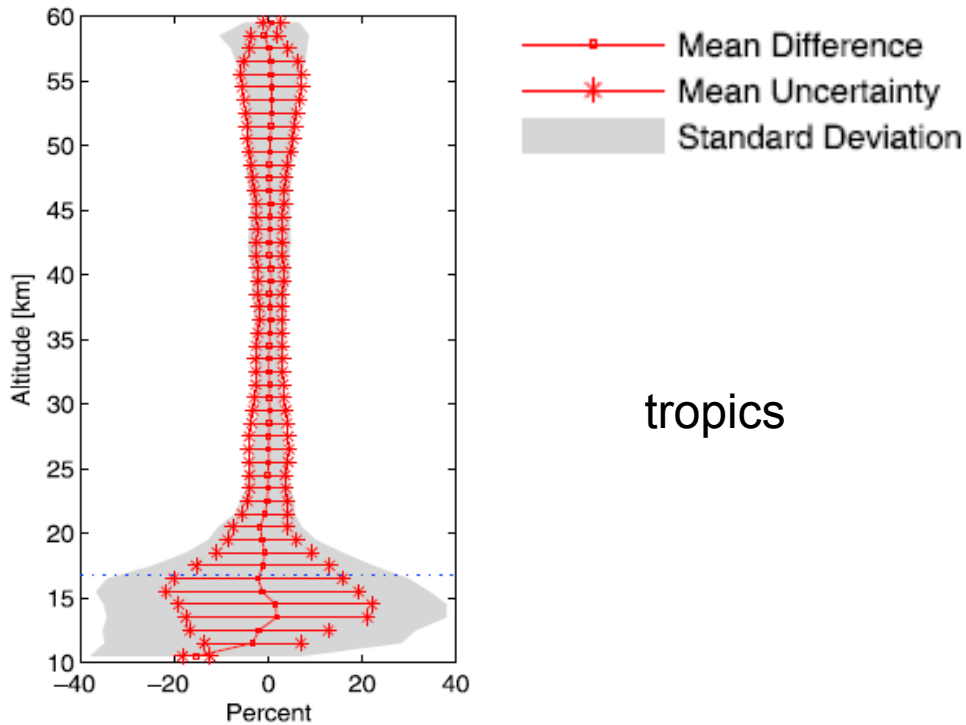
Realized for:

MIPAS



[Piccolo and Dudhia, 2007]

OSIRIS

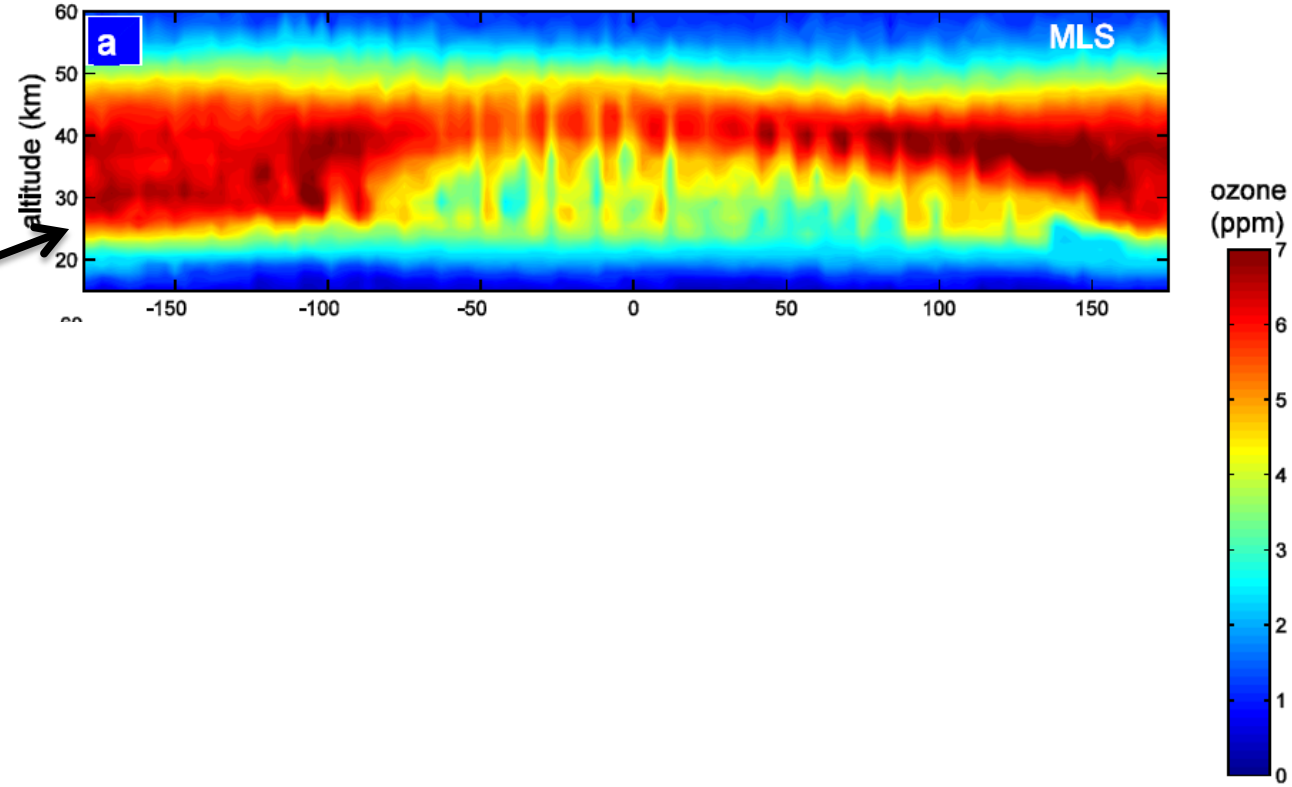


tropics

[Bourassa et al., 2011]

A spurious feature: quasi-periodic oscillations

Microwave Limb Sounder on board EOS/Aura (2004-present)



ozone mixing ratio at 60°N -70°N on 14-16 October 2007, averaged in 3° longitude bins.

Sofieva et al., AMT, 2014b

THANK YOU!

- **More details:**
  - Sofieva, V. F., Tamminen, J., Kyrölä, E., Laeng, A., von Clarmann, T., Dalaudier, F., Hauchecorne, A., Bertaux, J.-L., Barrot, G., Blanot, L., Fussen, D. and Vanhellemont, F.: Validation of GOMOS ozone precision estimates in the stratosphere, Atmos. Meas. Tech., 7(7), 2147–2158, doi:10.5194/amt-7-2147-2014, 2014., http://www.atmos-meas-tech.net/7/2147/2014/
  - Sofieva, V. F., Kalakoski, N., Päivärinta, S.-M., Tamminen, J., Laine, M., and Froidevaux, L.: On sampling uncertainty of satellite ozone profile measurements, Atmos. Meas. Tech., 7, 1891-1900, doi: 10.5194/amt-7-1891-2014, 2014, http://www.atmos-meas-tech.net/7/1891/2014/

Uncertainty estimates in Level 3 monthly zonal mean ozone profiles

- **From individual instruments.**
  - Included parameters
    - Uncertainty of the mean $\sigma^2_{mean} = \frac{s^2}{N}$
    - The mean of individual error estimates
    - inhomogeneity measures in latitude and in time)
- **Merged monthly zonal mean profiles**
  - Merging: weighted mean according to total uncertainty

$$\sigma^2 = \sigma^2_{mean} + \sigma^2_{sampling}$$