EUROPEAN COMMISSION

A new synergistic approach for tropospheric ozone profiling

Cecilia Tirelli[1], Ugo Cortesi[1], Michel Van Roozendael [2] , Arno Keppens[2], Rossana Dragani[3], Olaf Tuinder[4], Antti Arola[5], and the AURORA Consortium

[1] Institute for Applied Physics «Nello Carrara», IFAC-CNR, IT
[2] Belgian Institute for Space Aeronomy, BIRA-IASB, BE
[3] European Centre for Medium Range Weather Forecasts, ECMWF
[4] Royal Netherlands Meteorological Institute, KNMI, NL
[5] Finnish Meteorological Institute, FMI, FN

AURORA
Advanced Ultraviolet Radiation and Ozone Retrieval for Applications

THE FRAMEWORK PROGRAMME FOR RESEARCH AND INNOVATION
HORIZON 2020

Copernicus
The European Earth Observation Programme

OUTLINE

- The AURORA project
- AURORA data processing chain
- Complete Data Fusion
- Assimilation of Level-2 and Fused data products
- Conclusions

Overview of the AURORA project

Project Title: AURORA (Advanced Ultraviolet Radiation and Ozone Retrieval for Applications)

Project Duration: **36 months + 6 months extension (February 1°, 2016 – July 31°, 2019)**

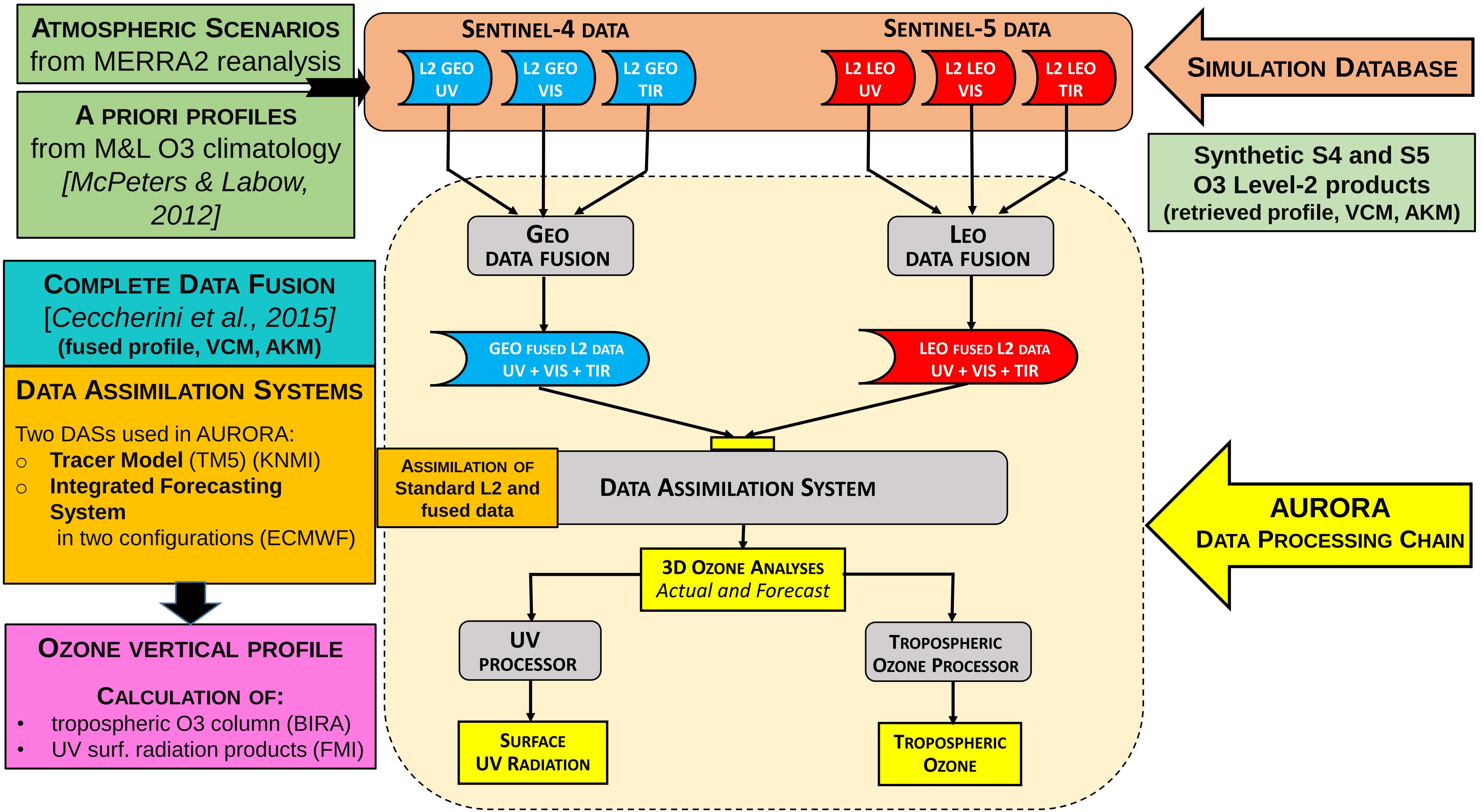
EU Framework Program: HORIZON 2020
H2020 Work Program: 2014-2015
Call: H2020-Earth Observation-2015
Topic: EO-2-2015 Stimulating wider research use of Copernicus Sentinel Data

Objectives and Scope

- to investigate the **potential of data fusion and assimilation** to convey complementary information of the atmospheric Sentinels measurements into unique geophysical products.
- to focus the **exploitation of the synergy** between simultaneous and independent measurements of the same target on **tropospheric O3** and **UV surface radiation**.
- to **reduce the complexity** of managing the high volume of Copernicus S-4 and S-5 data and **increase its quality**
- to develop **a prototype data processing system** and **demonstrate its capability to work with simulated data** as close as possible to the operational environment
- to develop **two operational downstream services** (innovative mobile App for UV dosimetry and tropospheric ozone monitoring application for prediction of air quality).

OUTLINE

- The AURORA project
- ***AURORA data processing chain***
- Complete Data Fusion
- Assimilation of Level-2 and Fused data products
- Conclusions


ATMOSPHERIC SCENARIOS
from MERRA2 reanalysis
A PRIORI PROFILES
from M&L O3 climatology
[McPeters & Labow, 2012]
SENTINEL-4 DATA
L2 GEO UV
L2 GEO VIS
L2 GEO TIR
SENTINEL-5 DATA
L2 LEO UV
L2 LEO VIS
L2 LEO TIR
SIMULATION DATABASE
Synthetic S4 and S5
O3 Level-2 products
(retrieved profile, VCM, AKM)
GEO DATA FUSION
LEO DATA FUSION
COMPLETE DATA FUSION
[Ceccherini et al., 2015]
(fused profile, VCM, AKM)
GEO FUSED L2 DATA
UV + VIS + TIR
LEO FUSED L2 DATA
UV + VIS + TIR
DATA ASSIMILATION SYSTEMS
Two DASs used in AURORA:
Tracer Model (TM5) (KNMI)
Integrated Forecasting System
in two configurations (ECMWF)
ASSIMILATION OF Standard L2 and fused data
DATA ASSIMILATION SYSTEM
AURORA
DATA PROCESSING CHAIN
3D OZONE ANALYSES
Actual and Forecast
UV PROCESSOR
TROPOSPHERIC OZONE PROCESSOR
OZONE VERTICAL PROFILE
CALCULATION OF:
tropospheric O3 column (BIRA)
UV surf. radiation products (FMI)
SURFACE UV RADIATION
TROPOSPHERIC OZONE

OUTLINE

- The AURORA project
- AURORA data processing chain
- ***Complete Data Fusion***
- Assimilation of Level-2 and Fused data products
- Conclusions

Complete Data Fusion (CDF)

Having N independent, simultaneous retrievals $\hat{\mathbf{x}}_i$ (i=1, 2, ..., N) that provide independent estimates of the atmospheric species profiles (on a common vertical grid) and characterized by the CMs $\mathbf{S}_i$ and the AKMs $\mathbf{A}_i$:

$$\mathbf{S}_i \equiv \left\langle \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \right\rangle = \left(\mathbf{K}_i^T \mathbf{S}_{yi}^{-1} \mathbf{K}_i + \mathbf{S}_{ai}^{-1} \right)^{-1} \mathbf{K}_i^T \mathbf{S}_{yi}^{-1} \mathbf{K}_i \left(\mathbf{K}_i^T \mathbf{S}_{yi}^{-1} \mathbf{K}_i + \mathbf{S}_{ai}^{-1} \right)^{-1}$$

$$\mathbf{A}_i \equiv \frac{\partial \hat{\mathbf{x}}_i}{\partial \mathbf{x}_{true}} = \left(\mathbf{K}_i^T \mathbf{S}_{yi}^{-1} \mathbf{K}_i + \mathbf{S}_{ai}^{-1} \right)^{-1} \mathbf{K}_i^T \mathbf{S}_{yi}^{-1} \mathbf{K}_i$$

σ_i : errors on $\hat{\mathbf{x}}_i$

S_{yi} : CMs of the observations

K_i : Jacobians of the forward models

S_{ai} : CMs of the a priori profiles

x_{true} : true profile

The fused product is given by:

$$\mathbf{x}_f = \left(\sum_{i=1}^{N} \mathbf{A}_i^T \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1} \right)^{-1} \left(\sum_{i=1}^{N} \mathbf{A}_i^T \mathbf{S}_i^{-1} \boldsymbol{\alpha}_i + \mathbf{S}_a^{-1} \mathbf{x}_a \right)$$

x_a : a priori

S_a : a priori covariance

$$\boldsymbol{\alpha}_i \equiv \hat{\mathbf{x}}_i - \left(\mathbf{I} - \mathbf{A}_i \right) \mathbf{x}_{ai} = \mathbf{A}_i \mathbf{x}_{true} + \boldsymbol{\sigma}_i$$

The corresponding error covariance matrix and averaging kernel matrix are given by:

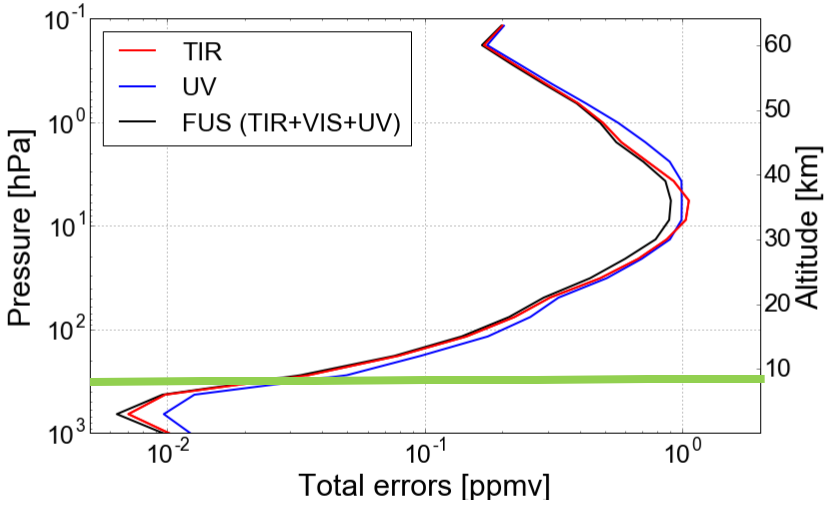
$$\mathbf{S}_f = \left(\sum_{i=1}^{N} \mathbf{A}_i^T \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1} \right)^{-1} \sum_{i=1}^{N} \mathbf{A}_i^T \mathbf{S}_i^{-1} \mathbf{A}_i \left(\sum_{i=1}^{N} \mathbf{A}_i^T \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1} \right)^{-1}$$

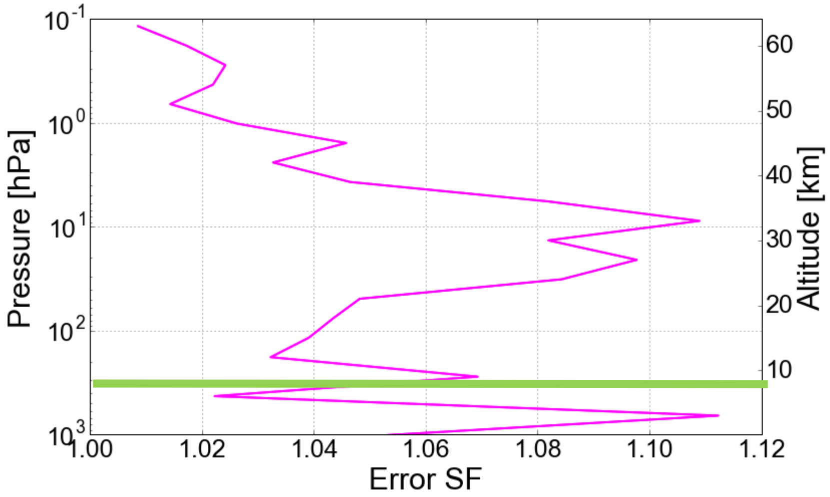
$$\mathbf{A}_f = \left(\sum_{i=1}^{N} \mathbf{A}_i^T \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1} \right)^{-1} \sum_{i=1}^{N} \mathbf{A}_i^T \mathbf{S}_i^{-1} \mathbf{A}_i$$

➔ **Reference** - Ceccherini et al. (2015) (also available at http://www.aurora-copernicus.eu/data-fusion/)

Ozone profile from CDF - Results

Here the CDF method has been used to fuse ozone profiles retrieved from simulated measurements in the TIR, VIS and UV spectral regions of Sentinel 4 (same space-time locations). Fused data shown here are simulated measurements of the first week of April 2012 (23881 analyzed pixels).





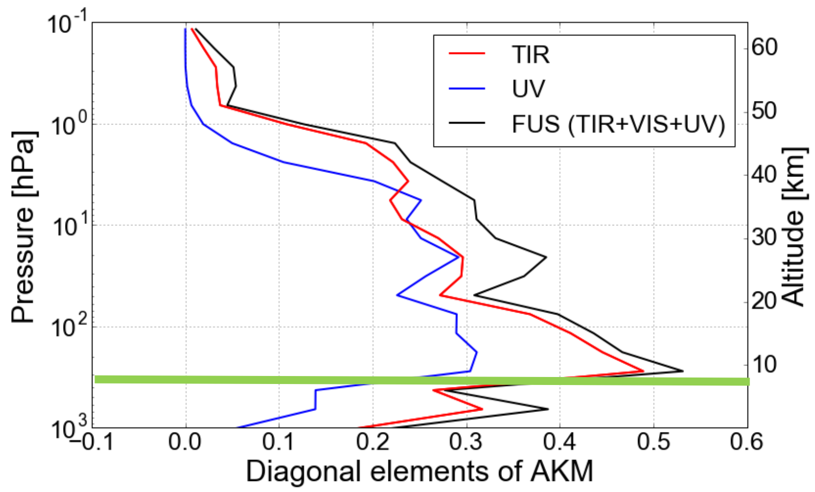
For each pressure level the error synergy factor (SF) is defined as:

$$SF_{err} = \frac{Min_{i=1,\dots,N}\,\sigma_{i,tot}}{\sigma_{f,tot}}$$

SF >1 synergy among sources of information

$\sigma_{i,tot}$ = total error of the *i-th* profile to be fused

$\sigma_{f,tot}$ = total error of the fused profile



Spectral range	TIR	UV	FUS (TIR+VIS+UV)
Number of DOF below TP	0.8	0.3	0.9

Spectral range	TIR	UV	FUS (TIR+VIS+UV)
Number of DOF	4.9	3.4	5.8

The fused product shows:

- smaller error for each pressure level
- greater diagonal element for each pressure level
- higher DOFs

The quality assessment showed that fused products have better quality than individual products for all the considered quantifiers

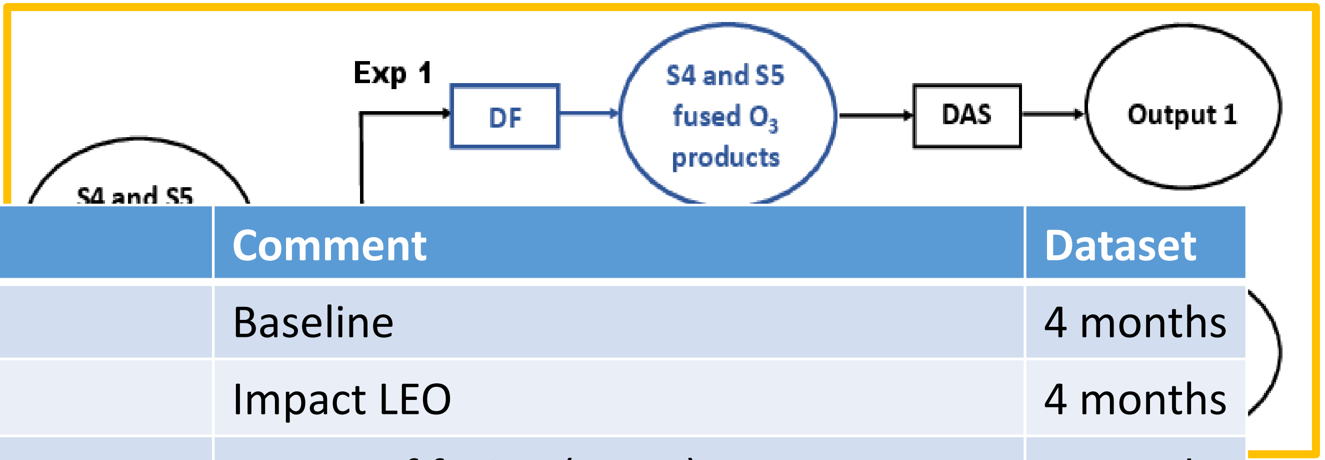
OUTLINE

- The AURORA project
- AURORA data processing chain
- Complete Data Fusion
- ***Assimilation of Level-2 and Fused data products***
- Conclusions

Data Assimilation

Two state-of-the-art Data Assimilation Systems (DASs) have been used in the AURORA project:

- Two configurations of the **Integrated Forecasting System** (IFS and C-IFS)- ECMWF
- **Tracer Model version 5** (TM5) - KNMI



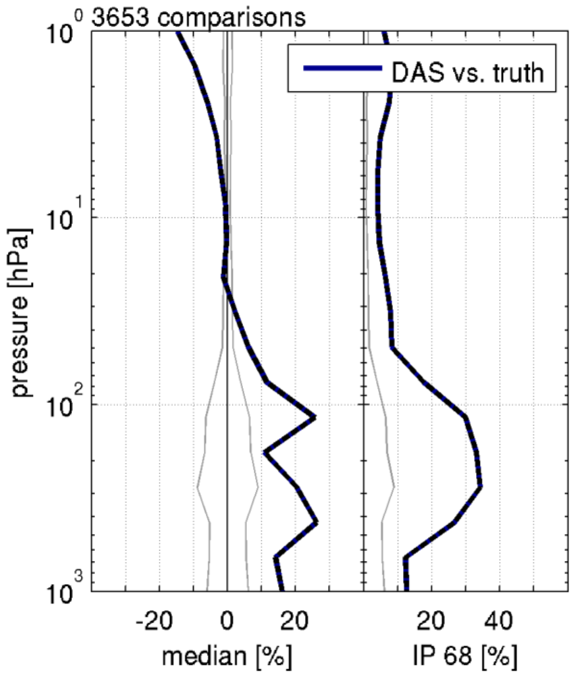
Both IFS and TM5 have been used to combine:

ID	Assimilated datasets	Setup	Comment	Dataset
1	None	BASE	Baseline	4 months
2	L2-LEO	BASE+(L2-LEO)	Impact LEO	4 months
3	Fused LEO-LEO	BASE+(LEO-LEO)	Impact of fusion (3 vs 2)	4 months
4	Fused LEO-LEO & GEO-GEO	BASE+(LEO-LEO)+(GEO-GEO)$^{(S4)}$	Impact of GEO (4 vs 3)	4 months
5	Fused LEO-LEO & GEO-GEO (including TEMPO & GEMS simulated data)	BASE+(LEO-LEO)+ (GEO-GEO)$^{(S4;\ TEMPO;\ GEMS)}$	Impact of 3 GEO instead of 1 (5 vs 4)	1 month
6	Fused LEO-GEO	BASE+(LEO-GEO)$^{(S4)}$	Impact of cross-platform fusion (6 vs 4)	4 months
7	L2-LEO+L2-GEO	BASE+(L2-LEO-L2-GEO)	Impact of GEO (7 vs 2) Impact of fusion (7 vs 4)	4 months

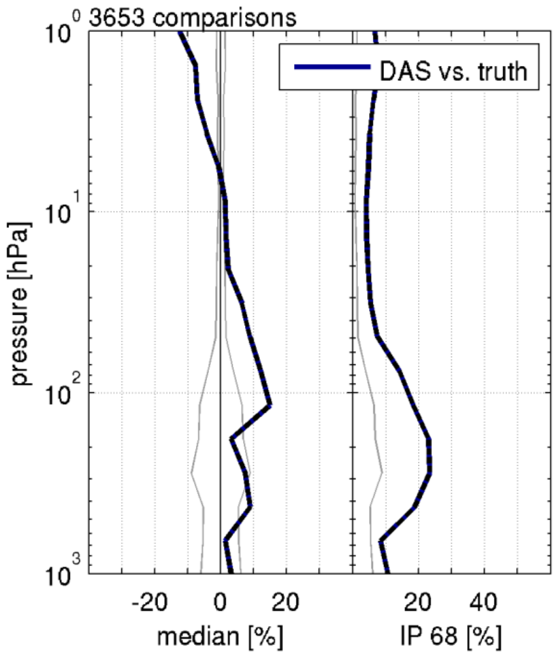
TM5

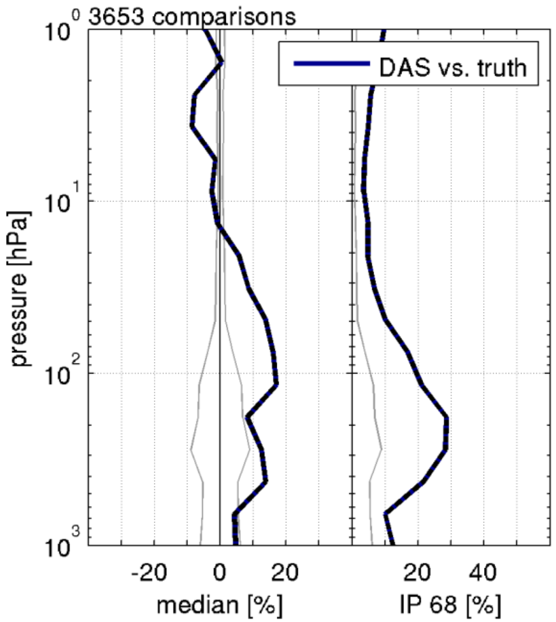
global **DAS versus truth** bias and spread

BASE

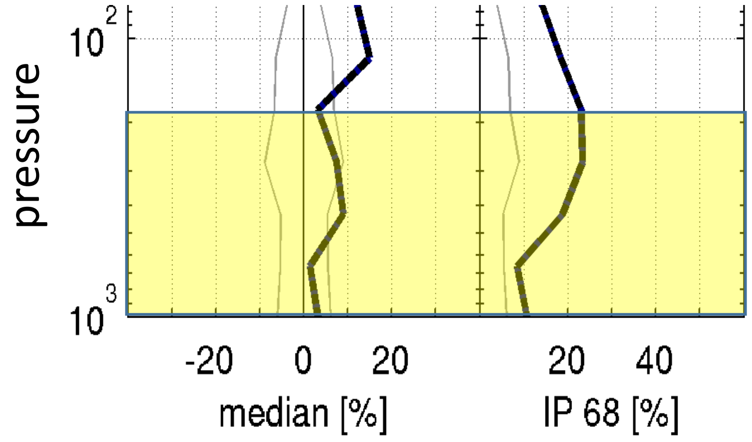


Vertical sampling	45 levels
Horizontal sampling	3° x 2°
Temporal sampling	3 h

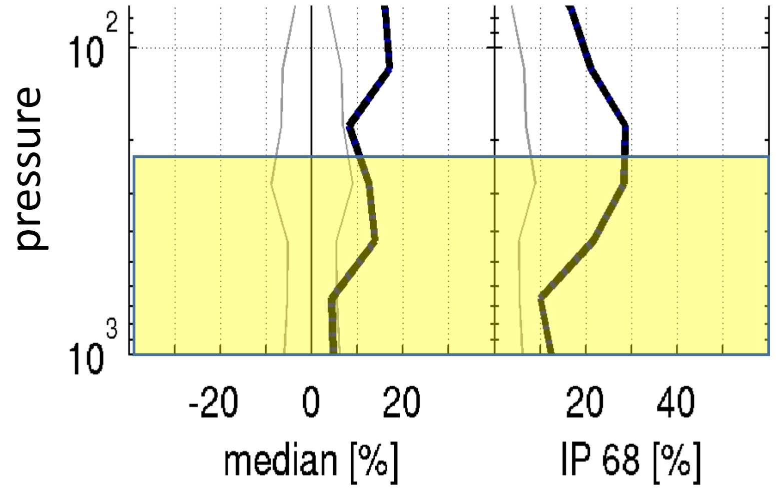




L2 LEO & L2 GEO



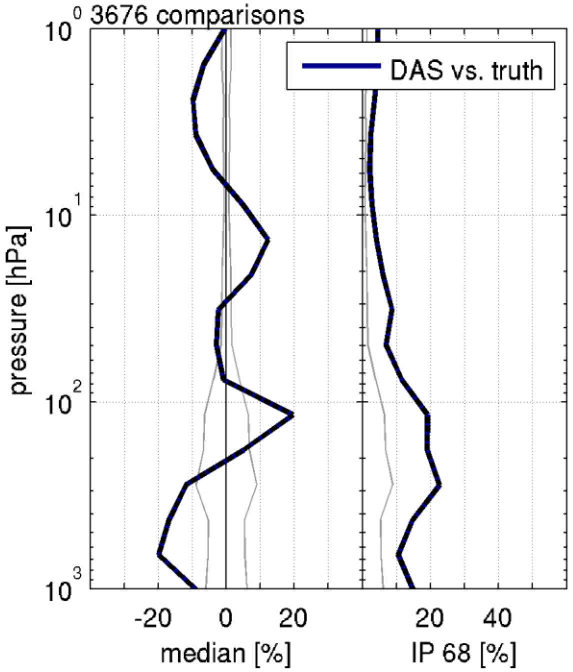
FUSED LEO-GEO



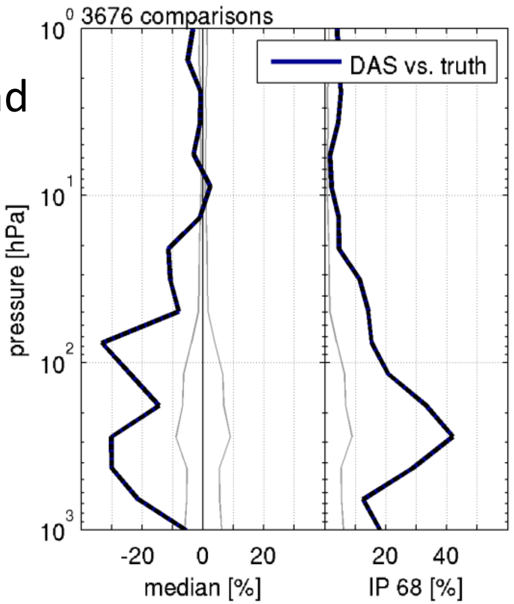
IFS

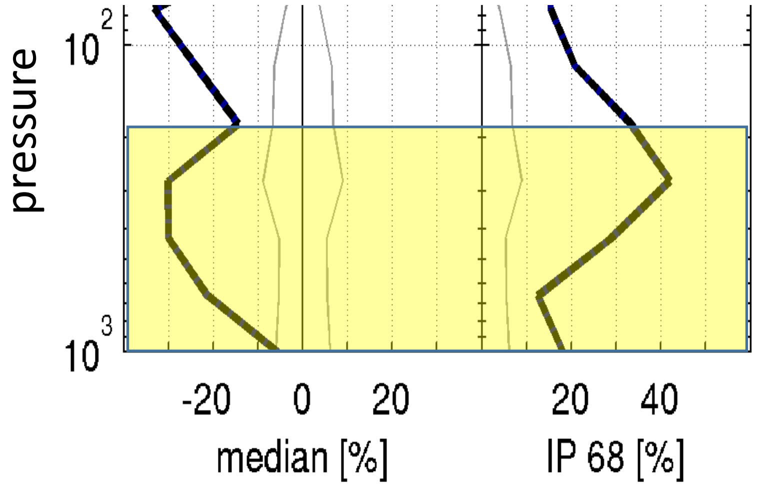
global **DAS versus truth** bias and spread

Base

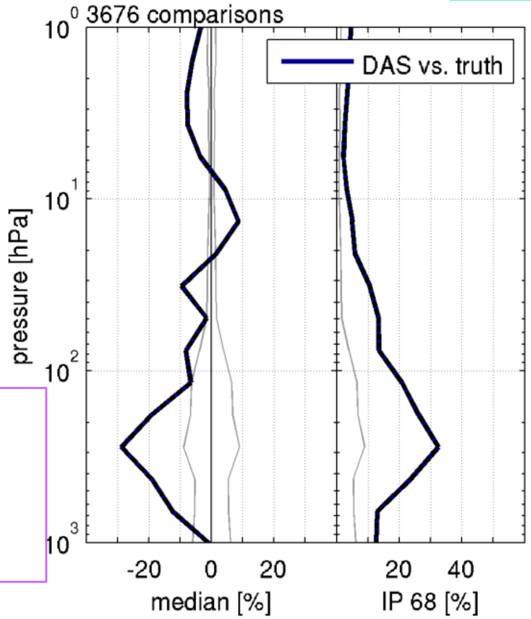


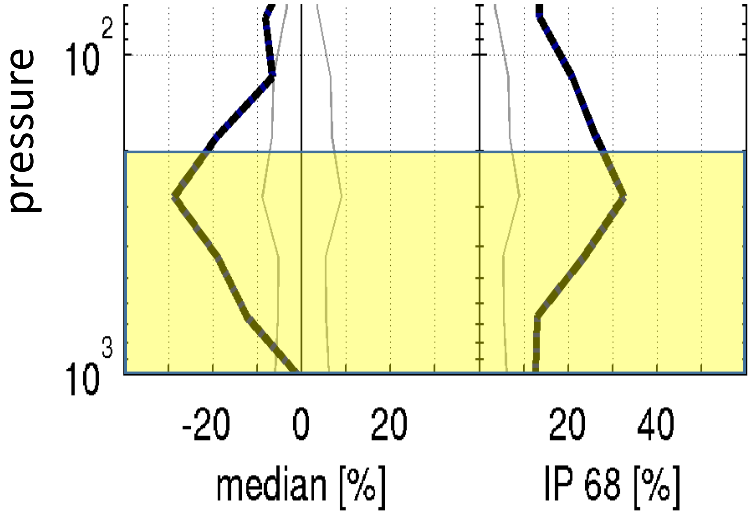
L2 LEO & L2 GEO





FUSED LEO-GEO



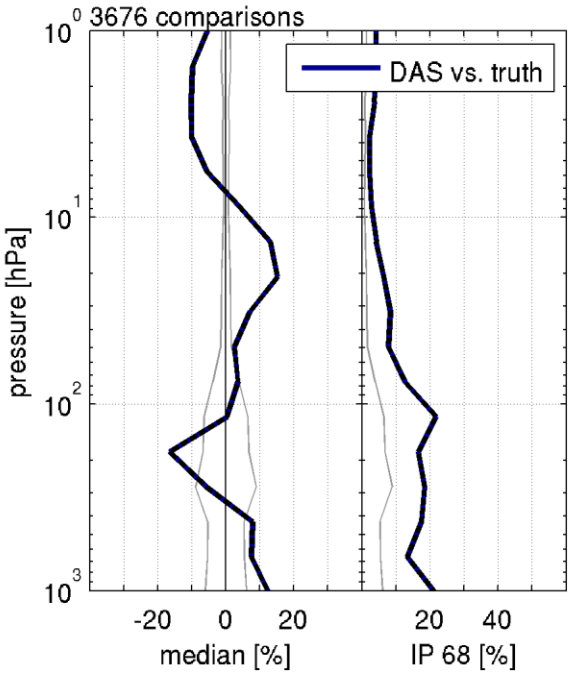


Vertical sampling	138 levels
Horizontal sampling	0.5° x 0.5
Temporal sampling	3 h

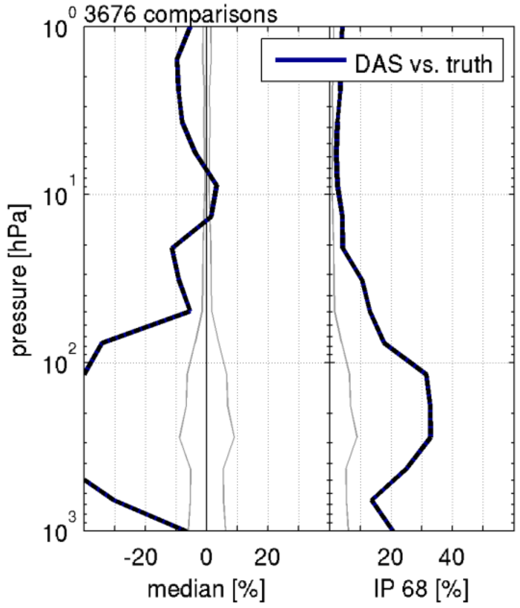
C-IFS

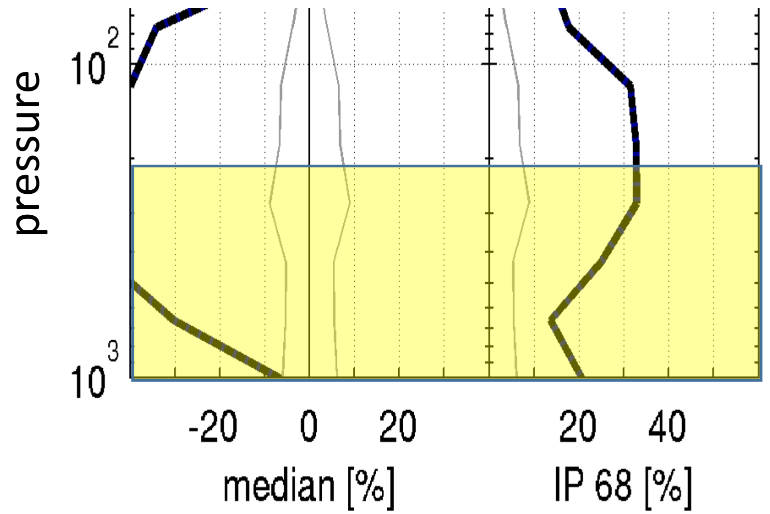
global **DAS versus truth** bias and spread

BASE

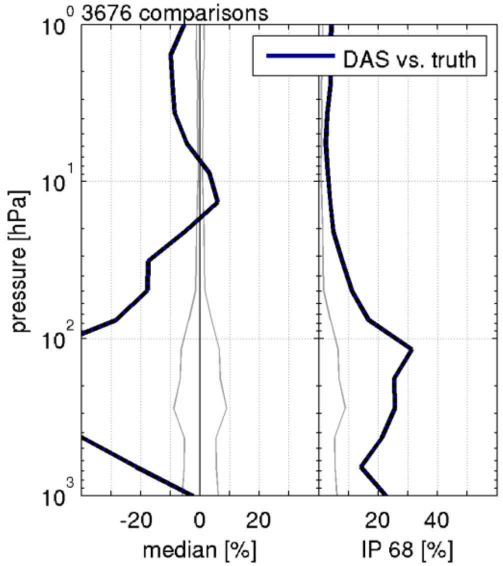


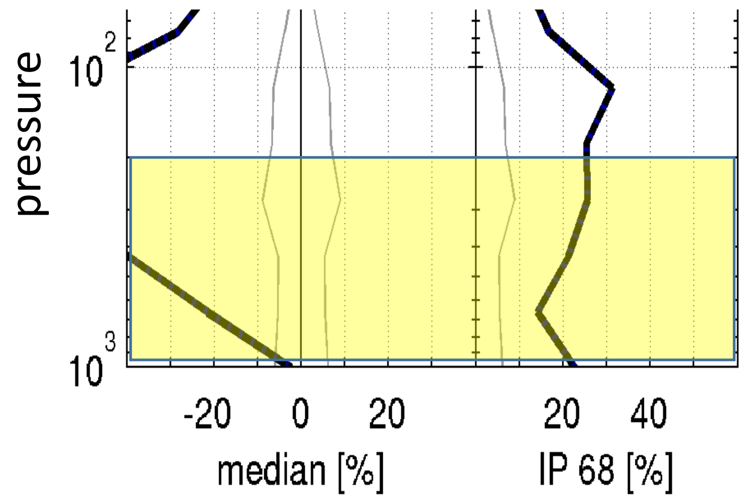
L2 LEO & L2 GEO





FUSED LEO-GEO





Vertical sampling	61 levels
Horizontal sampling	0.75° x 0.75
Temporal sampling	3 h

Ozone profile from Data Assimilation - First results

These are the outcomes of the preliminary analysis, further evaluations are ongoing.
The obtained results seem to be highly dependent from:

- the DAS
- the altitude in the atmosphere

Main comments:

- TM5 and (C-) IFS systems are not simply comparable because they are based on different methods (Kalman Filter and 4D-Var) (Note that TM5 Does not assimilate total column L2 VIS data)
- Data coverage is patchy (reduction of Sentinels pixels was necessary)

Conclusions from first analysis:

- TM5: agreement with the reference atmosphere improved over the full domain, (C-) IFS : agreement is improved in the stratosphere, for UTLS and troposphere only **fused assimilation is improved**
- CDF preceding assimilation does not result in information loss while the assimilation **input data size is significantly reduced**
- **O3 surface concentration** is increased for all experiments and all DASs but the strong ozone underestimation of the a priori profiles might be reduced by CDF.
- **CDF reduces vertical oscillations** (and spread) in the troposphere acting as an intelligent averaging process including vertical smoothing

Conclusions

- Data fusion: the quality assessment showed that **fused products have better quality than individual products** for all the considered quantifiers and improve the quality of tropospheric ozone product

- Data assimilation of L2 simulated data and fused products:

  - CDF preceding assimilation guarantees a significantly **reduced input data size** to the assimilation system
  - **CDF reduces vertical oscillations** (and spread) in the troposphere acting as an intelligent averaging process including vertical smoothing
  - the strong ozone underestimation of the retrievals' prior profiles at the surface might be reduced by data fusion, giving **more weight to the retrieval** (although this is farther off from the virtual truth for TM5)

The main goal of the data assimilation process is to provide global ozone fields and forecast (taking into account there are losses of accuracy along the processing chain)

There's an impact!
This makes us think that in a real case study we could potentially get interesting results.

AURORA web-site: http://www.aurora-copernicus.eu/

AURORA contact: aurora.coord@ifac.cnr.it

Assimilation experiments comparisons