Ozone satellite data synergy and combination with non-satellite data in the AURORA project

Ugo Cortesi(1), Cecilia Tirelli(1), Antti Arola(2), Rossana Dragani(3), Arno Keppens(4), Edo Loenen(5), Andrea Masini(6), Marc Bonazountas(7), Chrisovalantis Tsiakos(7), Ronald van der A(8), Koen Verberne(9) and the AURORA Consortium

(1) Institute for Applied Physics «Nello Carrara», IFAC-CNR, IT
(2) Finnish Meteorological Institute, FMI, FN
(3) European Centre for Medium Range Weather Forecasts, ECMWF
(4) Belgian Institute for Space Aeronomy, BIRA-IASB, BE
(5) Science and Technology BV, S[&]T, NL
(6) Flyby Srl, IT
(7) Epsilon International SA, GR
(8) Royal Netherlands Meteorological Institute, KNMI, NL
(9) Datacraft, NL
The AURORA project:

• Basic Information
• Objectives
• Overall concept and Data Processing Chain
• Atmospheric scenario and data simulation
• Data fusion and data assimilation
• AURORA Technological Infrastructure
• Validation of AURORA products
• Applications
AURORA project

EU Framework Program: HORIZON 2020

Research Area: Leadership in Enabling and Industrial Technologies (LEIT)
Sub-program: Space

Call: H2020-Earth Observation-2015
Topic: EO-2-2015 Stimulating wider research use of Copernicus Sentinel Data

Project Title: Advanced Ultraviolet Radiation and Ozone Retrieval for Applications
Project Duration: 36 months (February 1°, 2016 – January 31°, 2019)
AURORA website: http://www.aurora-copernicus.eu/
AURORA: Objectives of the project

**SCIENCE**

- 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 (i.e., \( O_3 \) vertical profile) on **tropospheric O\(_3\)** and **UV surface radiation**

**TECHNOLOGY**

- 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.

**APPLICATION**

- To develop **two operational downstream services** (innovative mobile App for UV dosimetry and tropospheric ozone monitoring for prediction of air quality)
AURORA overall concept and data processing chain

Individual Ozone products from LEO and GEO Sentinels (UV, VIS and TIR bands)

DATA COMBINATION
Fusion and Assimilation

Unique geophysical product
Ozone vertical profile
Calculation of tropospheric \( O_3 \) and UV surface radiation

SIMULATION DATABASE

L2 GEO UV
L2 GEO VIS
L2 GEO TIR
L2 LEO UV
L2 LEO VIS
L2 LEO TIR

DATA ASSIMILATION SYSTEM
Integrated Forecasting System

GEO DATA FUSION
LEO DATA FUSION

GEO FUSED L2 DATA
UV + VIS + TIR
LEO FUSED L2 DATA
UV + VIS + TIR

3D OZONE ANALYSES
Actual and Forecast

UV PROCESSOR
TROPOSPHERIC OZONE PROCESSOR

SURFACE UV RADIATION
TROPOSPHERIC OZONE
Atmospheric Scenario and Data Simulation

**Atmospheric scenario** ➔ definition of the «true» state of the atmosphere for forward calculation and generation of S4 and S5 synthetic measurements (ECMWF).

Atmospheric scenario (4 months of data: Apr-Jul 2012) are generated from the MERRA 2 re-analysis, provided by GMAO at NASA Goddard Space Flight Center.

**Ozone climatology** ➔ The McPeters and Labow climatology is used as a priori in the retrieval process (McPeters & Labow, 2012).

**Sentinel-4 and Sentinel-5 data simulation**

- LEO and GEO UV L2 products for Ozone and associated VCMs and AKMs (FMI, with contributions from KNMI).
- LEO and GEO VIS L1 and L2 data for Ozone and associated VCMs and AKMs (BIRA-IASB, with contributions from KNMI).
- LEO and GEO TIR L2 products for Ozone and associated VCMs and AKMs (IFAC-CNR).
Complete Data Fusion

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

$$S_i = \begin{pmatrix} \sigma_i \sigma_i^T \end{pmatrix} = \left( K_i^T S_{yi}^{-1} K_i + S_{ai}^{-1} \right)^{-1} K_i^T S_{yi}^{-1} K_i \left( K_i^T S_{yi}^{-1} K_i + S_{ai}^{-1} \right)^{-1}$$

$$A_i = \frac{\delta \hat{x}_i}{\delta x_{true}} = \left( K_i^T S_{yi}^{-1} K_i + S_{ai}^{-1} \right)^{-1} K_i^T S_{yi}^{-1} K_i$$

The fused product is given by:

$$x_f = \left( \sum_{i=1}^{N} A_i^T S_i^{-1} A_i + S_a^{-1} \right)^{-1} \left( \sum_{i=1}^{N} A_i^T S_i^{-1} a_i + S_a^{-1} x_a \right)$$

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

$$S_f = \left( \sum_{i=1}^{N} A_i^T S_i^{-1} A_i + S_a^{-1} \right)^{-1} \left( \sum_{i=1}^{N} A_i^T S_i^{-1} A_i \right)$$

$$A_f = \frac{1}{\sum_{i=1}^{N} A_i^T S_i^{-1} A_i} \left( \sum_{i=1}^{N} A_i^T S_i^{-1} A_i \right)$$

Example: S4-S4 (GEO-GEO)

1 coincidence grid cell

1 TIR VMR profile
1 UV VMR profile
1 VIS total column

1 FUSED VMR profile
One week analysis for S4-S4 data fusion

**Average total errors**

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

\[
SF_{err} = \frac{\min_{i=1}^{N} \sigma_{i, tot}}{\sigma_{f, tot}}
\]

SF > 1 indicates synergy among sources of information.

**Average error SF**

**Average of the diagonal elements of the AKMs**

\[
\sigma_{i, tot} = \text{total error of the } i\text{-th profile to be fused}
\]
\[
\sigma_{f, tot} = \text{total error of the fused profile}
\]

**Average number of degrees of freedom (DOF)**

<table>
<thead>
<tr>
<th>Spectral range</th>
<th>TIR</th>
<th>UV</th>
<th>VIS</th>
<th>FUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of DOF</td>
<td>4.90</td>
<td>3.41</td>
<td>0.97</td>
<td>5.72</td>
</tr>
</tbody>
</table>
Data Assimilation

State-of-the-art Data Assimilation Systems (DASs) will be used to combine the LEO and GEO fused profiles, and contrasted with the assimilation of standard retrievals.

Two DASs are available to AURORA:

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

Both IFS and TM5 will be used to run the following experiments:
# Plan of assimilation experiments

<table>
<thead>
<tr>
<th>ID</th>
<th>Assimilated datasets</th>
<th>Setup</th>
<th>Comment</th>
<th>Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>None</td>
<td>BASE</td>
<td>Baseline</td>
<td>4 months</td>
</tr>
<tr>
<td>2</td>
<td>L2-LEO</td>
<td>BASE+(L2-LEO)</td>
<td>Impact LEO</td>
<td>4 months</td>
</tr>
<tr>
<td>3</td>
<td>Fused LEO-LEO</td>
<td>BASE+(LEO-LEO)</td>
<td>Impact of fusion (3 vs 2)</td>
<td>4 months</td>
</tr>
<tr>
<td>4</td>
<td>Fused LEO-LEO &amp; GEO-GEO</td>
<td>BASE+(LEO-LEO)+(GEO-GEO)(^{S4})</td>
<td>Impact of GEO (4 vs 3)</td>
<td>4 months</td>
</tr>
<tr>
<td>5</td>
<td>Fused LEO-LEO &amp; GEO-GEO (including TEMPO &amp; GEMS simulated data)</td>
<td>BASE+(LEO-LEO)+(GEO-GEO)(^{S4; \text{TEMPO}; \text{GEMS}})</td>
<td>Impact of 3 GEO instead of 1 (5 vs 4)</td>
<td>1 month</td>
</tr>
<tr>
<td>6</td>
<td>Fused LEO-GEO</td>
<td>BASE+(LEO-GEO)(^{S4})</td>
<td>Impact of cross-platform fusion (6 vs 4)</td>
<td>4 months</td>
</tr>
<tr>
<td>7</td>
<td>L2-LEO+L2-GEO</td>
<td>BASE+(L2-LEO-L2-GEO)</td>
<td>Impact of GEO (7 vs 2)</td>
<td>4 months</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Impact of fusion (7 vs 4)</td>
<td></td>
</tr>
</tbody>
</table>
Total Ozone Column from TM5 model output on 05-05-2012

Relative Difference
AURORA technological infrastructure: Geo-Database

Key features:
- Modular and extensible architecture, ready for big-data analysis
- Reliable and scalable storage for data and metadata
- Usage of OPeNDAP protocol in the communication with the geo-database
- OGC CSW metadata services to enable searching through the huge amount of data in the geo-database
- Adoption of OGC standards for maximum interoperability with other systems and infrastructures
- Exploitation of SaaS, PaaS, CaaS cloud services

OPeNDAP:
- OPeNDAP enables remote datasets to be accessed through data analysis and visualization packages.
- OPeNDAP handles transport, translation and subsetting of data residing in most Earth science data formats.
- Currently OPeNDAP is used and supported by institutions like NOAA, NASA, NSF, ABM, etc.

AURORA Geo-DATABASE

<table>
<thead>
<tr>
<th>Data &amp; Metadata Webservices</th>
<th>Access to data using the same protocol (OPeNDAP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web Catalog Service</td>
<td>Web services</td>
</tr>
<tr>
<td>OPeNDAP Service</td>
<td>GIS clients</td>
</tr>
</tbody>
</table>

Database (Metadata)
- External data
- GEO data files
- LEO data files

Hadoop Distributed File System

Browser access
AURORA technological infrastructure: Data Processing Chain (DPC)

- **AURORA DPC**: component controlling all the processing to be performed on the stored data.
- **DPC based on a distributed approach**: each tool will run on the site of one partner, triggered when elaboration is needed and upload resulting data on the Geo Db at the end of the process.
- **Key feature of the DPC infrastructure**: to hook processors with minimum or negligible effort
- **Entire elaboration managed by a central DPC Manager**: it coordinates all the jobs performed by the various tools provided by the AURORA partners.
Applications

AURORA aims to develop two operational downstream services using innovative mobile App for **UV dosimetry** and **tropospheric ozone monitoring application** for **major cities and regional prediction of air quality** reaching a pre-market version at the end of the project.

Pre-market service on Urban Pollution Monitoring

http://www.happysun.it/

Personal UV dosimetry
supports and induces a new life style for healthy solar light exposure

The very first sensor-less / satellite-based UV personal dosimeter:

- No wearables needed
- No smartphone exposure to the sun needed
- Never-ending calibration
- Clinically tested

- Phototype and skin UV sensitivity assessment
- Sunbath & sunscreen application planner
- UVB & UVA erythema effective dose measurement in real-time
- Warning & recommendation for sunscreen re-application
- Alerts for safe sunbath termination

www.happysun.co.uk (soon available for USA & Canada)
What is AIR-Portal?

**Information:**
- Satellite
- Local KNMI, DCMR (local air quality network)
- Crowdsourcing (e.g., Twitter / iSPEX)

**Pollution:**
- External / Foreign
- Human induced (e.g., transportation, industry)
- Background / natural phenomena (e.g., volcanic dust)

EU + large EU cities have indicated air quality as the largest challenge for urban areas.

AIR-Portal combines various levels of data sources:
- Earth Observation
- Ground-based measurements
- Meteorological forecasts
- Land use

AIR-Portal uses modelling to produce historical, current and forecasted air quality information at a resolution of 100 meter.

- Monitoring- and scenario-based dashboard on air quality for policy developers, citizens and industry
- Allows cities to implement policies
- Citizens can become aware and act to get a better quality of life.
What is AIR-Portal?

EU + large EU cities have indicated air quality as the largest challenge for urban areas.

AIR-Portal combines various levels of data sources:
- Earth Observation
- Ground-based measurements
- Meteorological forecasts
- Land use

AIR-Portal uses modelling to produce historical, current and forecasted air quality information at a resolution of 100 meters.

- Monitoring-and scenario-based dashboard on air quality for policy developers, citizens, and industry
- Allows cities to implement policies
- Citizens can become aware and act to get a better quality of life.
AURORA – Advanced Ultraviolet Radiation and Ozone Retrieval for Applications.

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

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