PHIDIAS

Prototype of HPC/Data Infrastructure for On-demand Services

PHIDIAS ATMOSPHERE USE CASE :
Steps forward in detection and identification of anomalous atmospheric events from space

SPASCIA Space Science Algorithmics

AERIS

Icare

IPSL

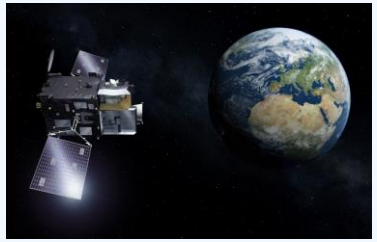
HYGEOS

SRON Netherlands Institute for Space Research

The PHIDIAS project has received funding from the European Union's Connecting Europe Facility under grant agreement n° INEA/CEF/ICT/A2018/1810854.

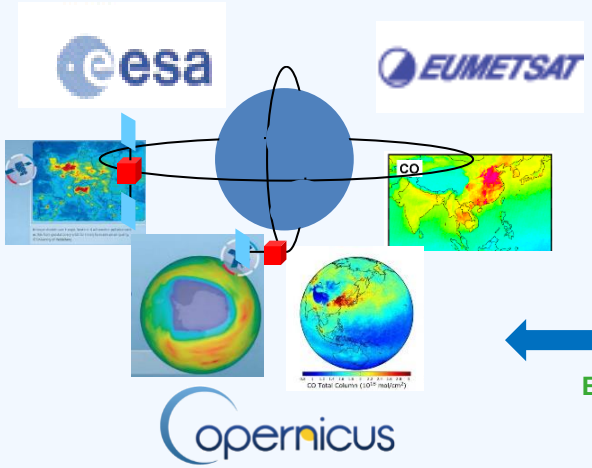
ATMOSPHERE USE CASE : CONTEXT

European atmospheric sounding missions will deliver each day several TB (terabytes) of raw datacubes at high spatial/temporal/spectral resolutions. This represents an unprecedented amount of atmospheric data, with improved quality and coverage.



Onboard **geostationary MTG satellite**, the Sentinel-4 mission comprises an **UVN** instrument and data InfraRed Sounder (**IRS**), completed by **FCI**

Each 30 minutes over Europe



Each day over the globe

Onboard **polar-orbiting MetOp SG satellite**, the Sentinel-5 mission comprises an **UVNS** instrument and data from **IASI-NG**, **MetImage** and **3MI**

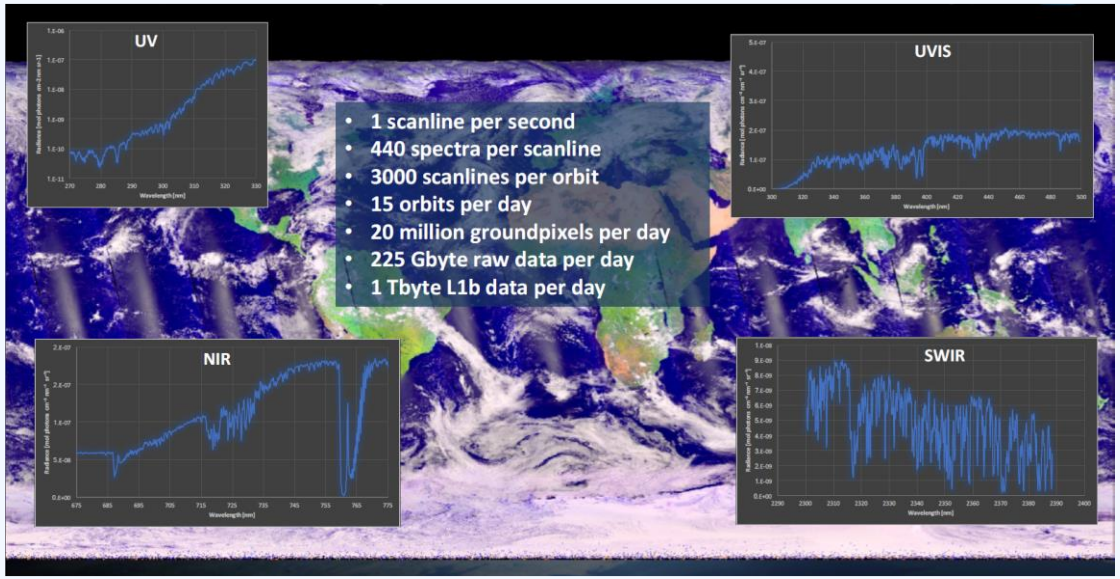
How to comprehensively deal with all available information ?

- Key challenge is to provide the capacity of “intelligent” screening of large amounts of satellite data for targeting scenes or events of interest in view of their dedicated processing or exploitation.

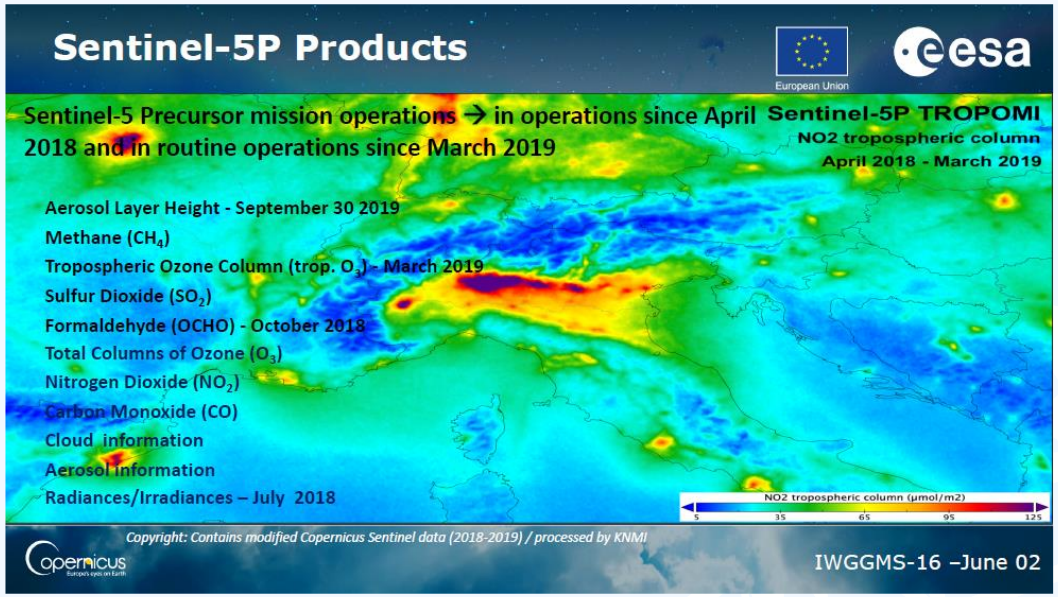
PHIDIAS addresses those needs by using HPC and HPDA capacities of **screening approaches in an operational context,** for detection and identification of atmospheric composition events.

S5P DATA FOR PHIDIAS USE CASE AND PROOF OF CONCEPT

- This use case proposes to develop, test and prototype the approach with Sentinel 5 Precursor data/products : Level 1b data (calibrated radiances) and Level 2 products (atmospheric contents of NO2, CO, …)



20 millions of ground pixels, 500 Gbytes of L1b data, per day

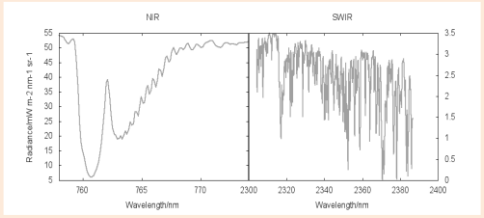


~ 10 L2 products (gas pollutant content, clouds, aerosols, UV index) distributed operationnaly from April 2018

freely available to users from Copernicus / ESA :https://s5phub.copernicus.eu.
Available from AERIS/ICARE French data center : https://www.icare.univ-lille.fr/asd-content/archive/?dir=S5P/

2 PROCESSING CHAINS DEVELOPED BY PHIDIAS :

SYSTEMATIC, OPERABLE DETECTION OF ATMOSPHERIC INFORMATION OF INTEREST, TARGETTING OVER LARGE AMOUNT OF DATA



L1B data

1. PCA-based screening of L1B data

On the flow, realtime detection & analysis of anomalous events

Catalogue/Metadata of selected and characterised L1B data

Earth observation from space

Sentinel 5 Precursor (S5P) L1B, L2

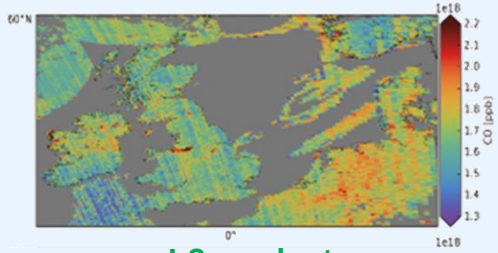
PHIDIAS Pilot users :

SRON Netherlands Institute for Space Research

AERIS

Complementary dedicated services

Early warning, monitoring, decision support



L2 product

2. AI-based Detection of gas plume and sources from L2

On demand processing of the data for detection of plumes and sources over a region/period specified by the user

Processed L2 products with additional characteristics

S5P TROPOMI PCA-based screening of L1 data

A protype chain has been implemented, and is delivered to ICARE for implementation on HPC

Build a reference database of spectra statistically representative of the atmospheric variability for generating the basis of principal components

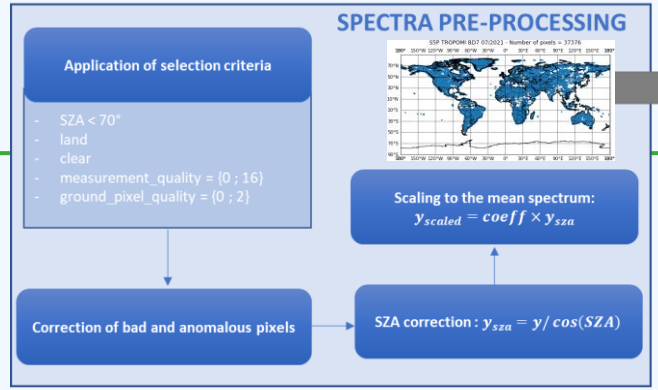
- More than 300 000 spectra (after filtering).
- Huge work done for pre-processing and selection of the spectra : multiple filtering, corrections, scaling …
- Challenge to practically define “representative variability” in the SWIR (impact of solar spectrum, SZA variability, clouds and aerosols, glint, SAA, …)

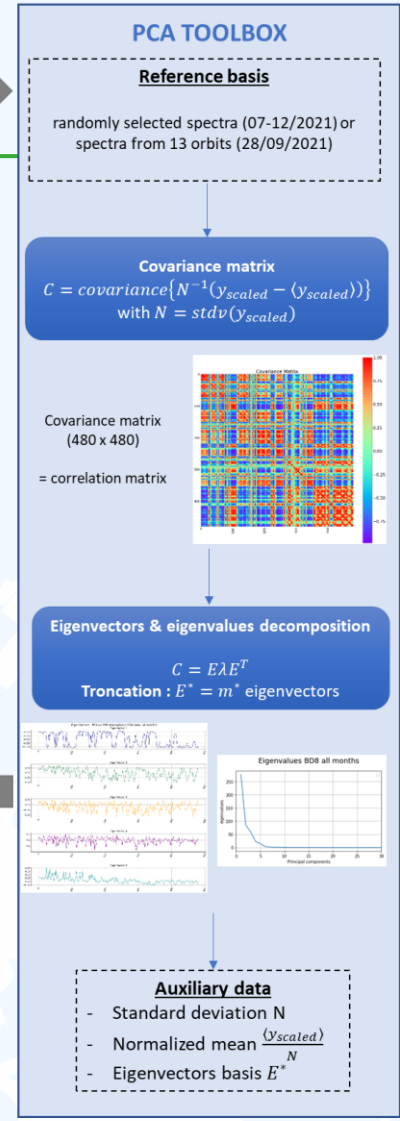
The pre-processing is a critical part, applied both to reference and processed spectra.

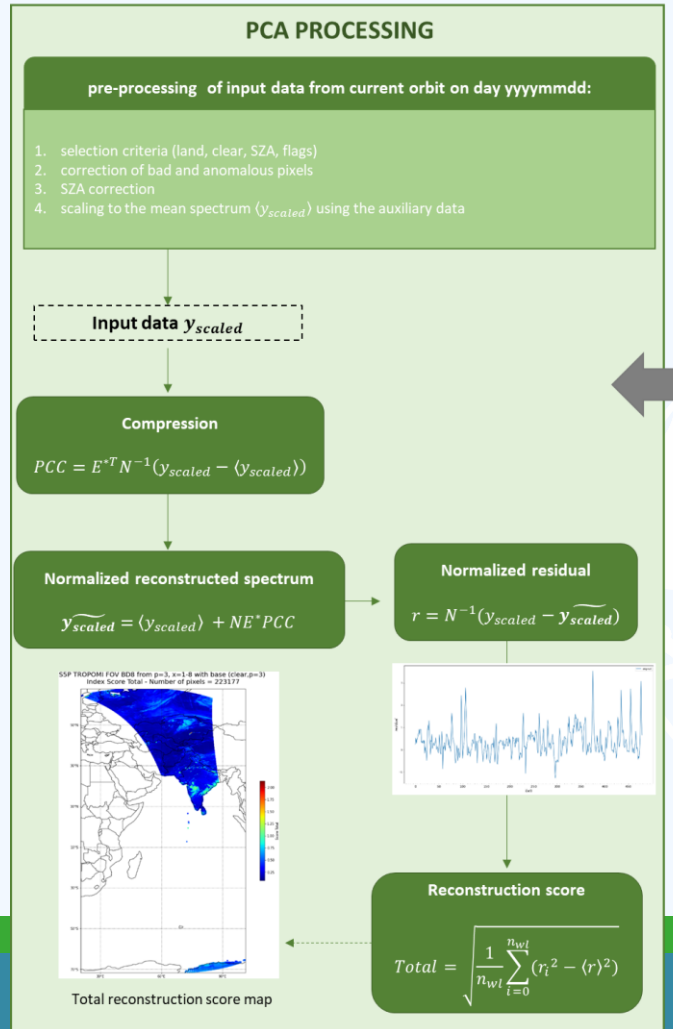
Processing : projection of the current spectrum on the PC basis, truncated reconstruction, computation of the spectral residual (original minus reconstructed spectrum) and mapping of reconstruction scores.

Scores quantify the reconstruction performances, focused on spectral domains of interest : CO, CH4, surface, (others ?)

Mapping of anomalies, that basically correspond to “rare”, “unexpected”, “extreme” events







Prototype performance (to be optimized) :

- ~ 4 minutes per orbit
- ~ 1 hour per day

In the present version, 11 score indicators has been defined in BD7 and BD8

- CO indicators are validated for fires, and are testing for pollution events (over cities, plants)
- CH4 indicators to be tested
- Combination of several indicators allows to avoid ambiguities in detection, and should allow the identification of detected events.

The analysis of full residuals are also promising for improvements of detection/characterization, and for identification of other features.

On going work : detection thresholds and dedicated post-processing for automatic detection and alerts (this will benefit from experience gained on IASI PCA studies supported by CNES)

Input data y_{scaled}

Compression

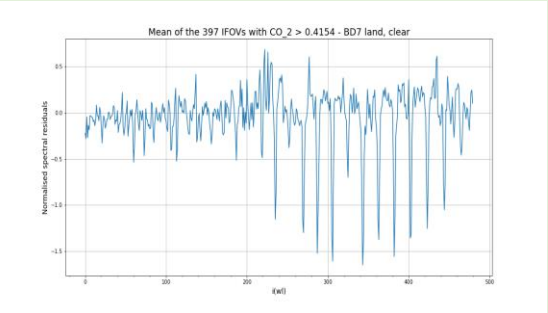
$$PCC = E^{*T} N^{-1} (y_{scaled} - \langle y_{scaled} \rangle)$$

Normalized reconstructed spectrum

$$\widetilde{y_{scaled}} = \langle y_{scaled} \rangle + NE^{*}PCC$$

Normalized residual

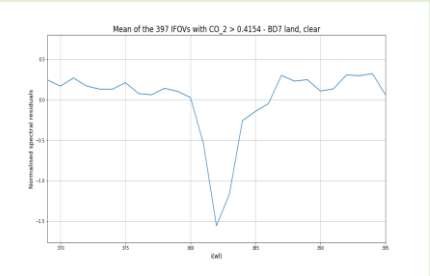
$$r = N^{-1}(y_{scaled} - \widetilde{y_{scaled}})$$



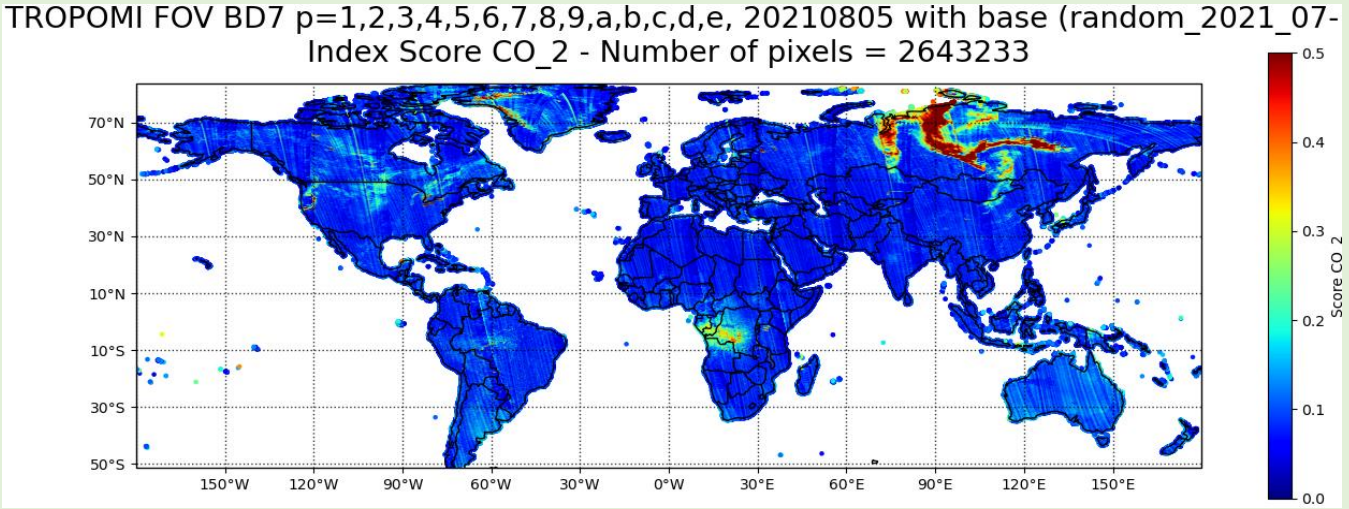
Dedicated spectral windows for indicators

Reconstruction score

$$CO = \sqrt{\frac{1}{n_{CO}} \sum_{i=0}^{n_{CO}} (r_i{}^2 - \langle r \rangle^2)}$$



CO reconstruction score map



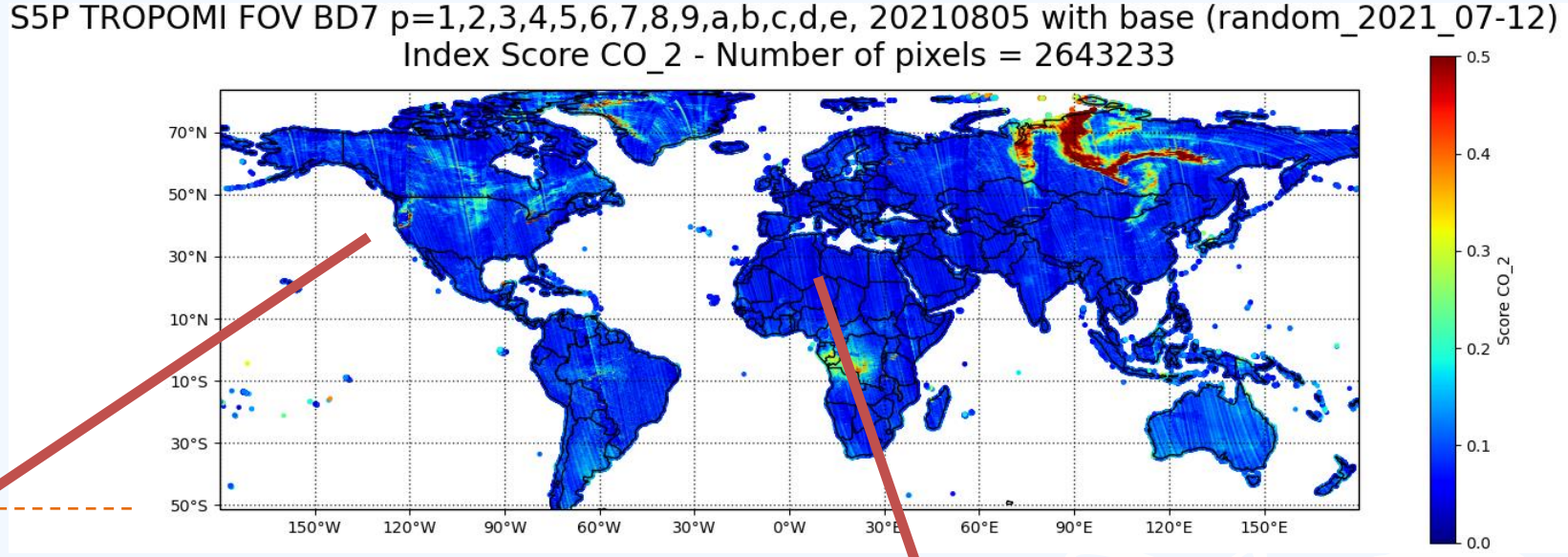
Detection of fires: 05/08/2021

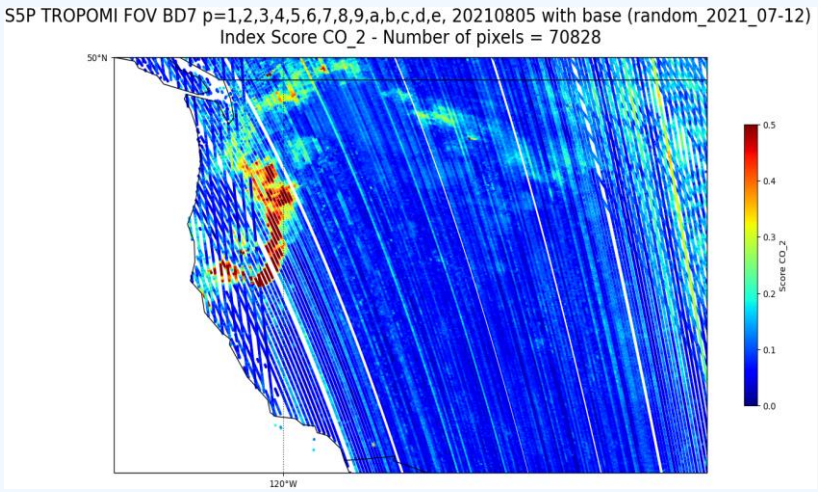
S5P PCA processing is well designed for "operation screening" of atmospheric information (oriented on extreme events) :

- Automatic, real time, all weather(clouds) processing of large amount of data, for detection and characterization of scenes of interest

A prototype processing will be available for scientific evaluation and exploration.

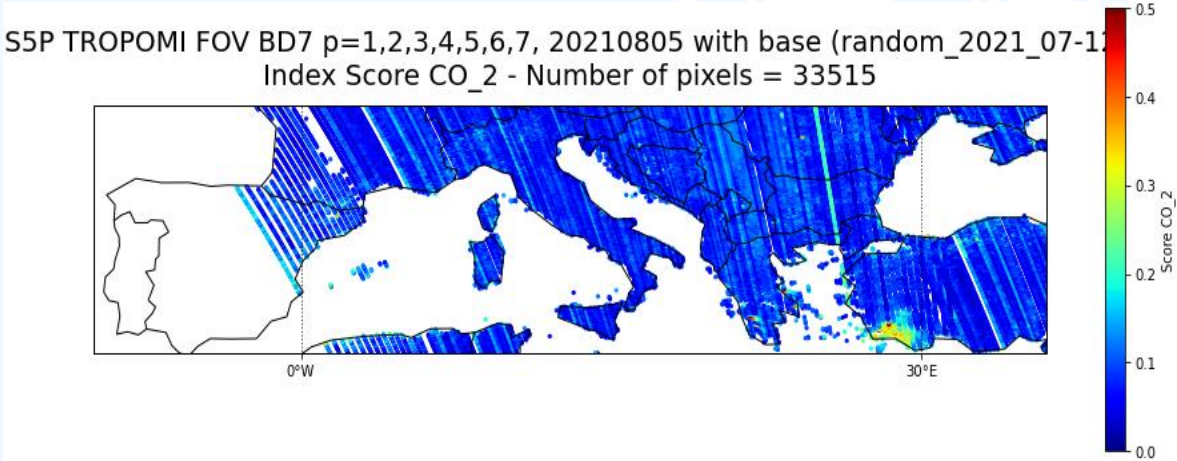
- Evaluation of added values, definition of objectives for systematic /on the flow processing, historical reprocessing ?





Detection of CO anomalies :

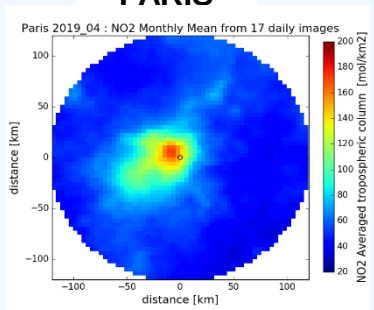
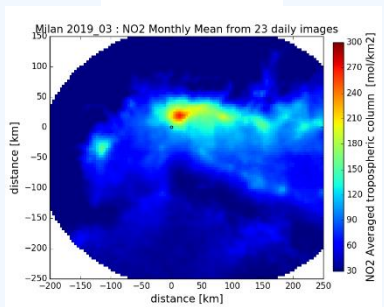
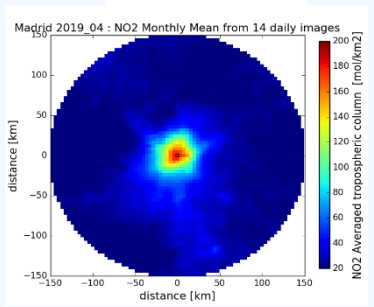
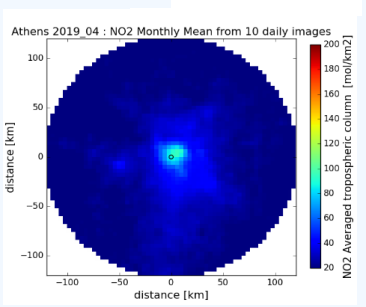
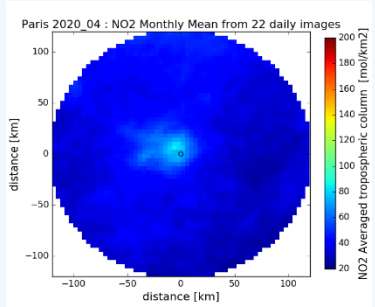
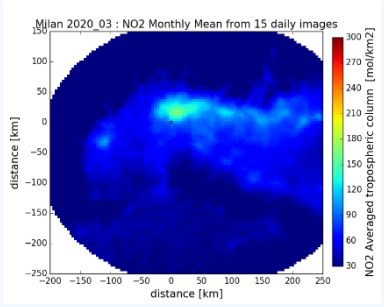
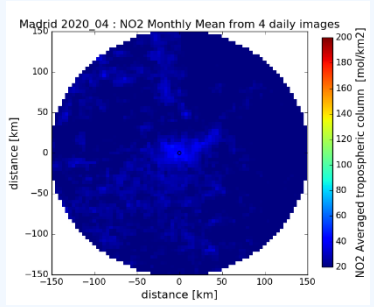
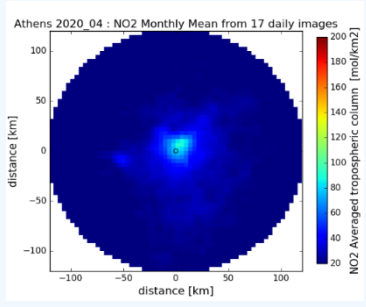
fires in US, in turkey , Greece, and Siberia



S5P TROPOMI L2 product processing for source detection

P. PRUNET, O. LEZEAUX, C CAMY-PEYRET, H. THEVENON

- **A first study (Prunet et al., supported by PHIDIAS) of the Covid impact on city pollution levels has shown promising potential for monitoring :**
  - From ESA S5P measurements of the **air pollutant NO_2 from space**, we have assessed the impact of the human activity reduction on air pollution by comparing the first 4-months periods of 2019 and 2020 on a daily, weekly and monthly basis for 4 major cities in Europe

	PARIS	MILAN	MADRID	ATHENS
2019 monthly averages				
2020 Monthly averages				

Reductions in the pollution level (using NO_2 tropospheric column as a proxy) have been observed from Mid March and for April 2020 (52% +/-9% for Paris; 28% +/-8% for Milan region; 54% +/-16% for Madrid; not significantly observed for Athens), as compared to the same periods in 2019

Objectives : Consolidate and systematize for operable use on HPC; Improve (emission quantification ?); Generalize (several species ?)

- **Implement an automatic processing of pollutant source detection and characterization for on demand use over regions/periods of interest**

Consolidated processing chain on NO2 source detection

Tests to three sites (image pixel size 80x80 km)

- Belchatow, Geismar, Matimba

Surface flux computation from temporally averaged images :

Emission maps using equations of continuity (based on Beirle et al., 2019)

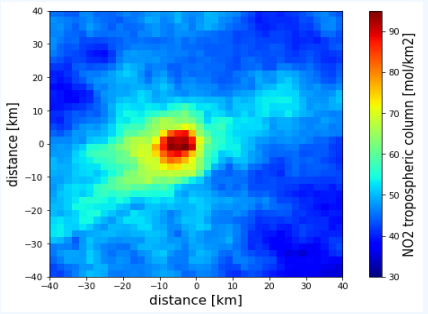
Flux = L C **w**

$\nabla(LC\mathbf{w}) = \text{Emission} - \text{Sink}$

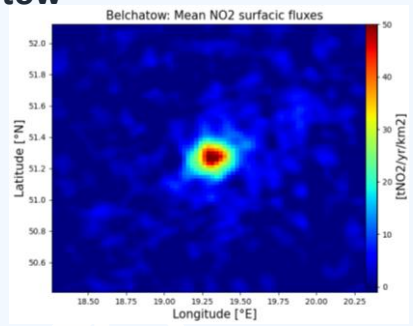
$\text{Emission} = \nabla(LC\mathbf{w}) + LV/\tau$

w → wind field,
L → is ratio Nox/NO2
C → tropospheric columns from S5P
τ → time scale for concentration loss

Belchatow



L2 NO2 colum, temporally averaged over 30 days and spatially integrated on 2km grid



Surface flux mapping at 2km resolution

Source detection and characterisation:

Step 1: First approximation of the number of sources by identifying the local peaks in the region.

Step 2: For each peak, applying a Gaussian function to fit the emission data and estimating corresponding Gaussian parameters

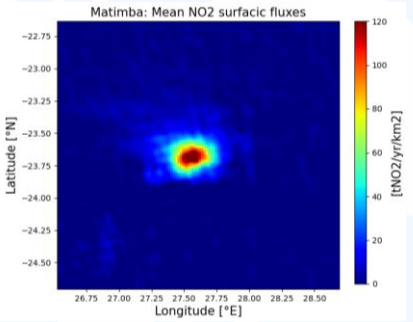
Step 3: Comparing the retrieved emissions from the Gaussian function to the true emissions.

Step 4: Repeating steps 1-3 on the residual (true - retrieved) emissions until a predefined magnitude (for instance 85% in following sites)) of the emission is retrieved.

Matimba :

strong, isolated site

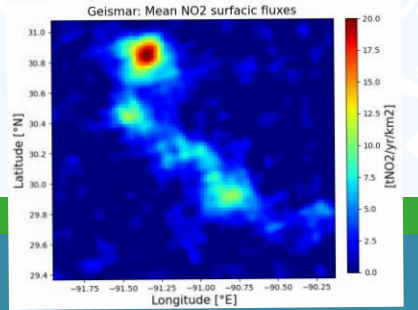
Max: ~200 $tNO2.yr^{-1}.km^{-2}$



Geismar :

Complex, weaker emissions

Max: ~20 $tNO2.yr^{-1}.km^{-2}$

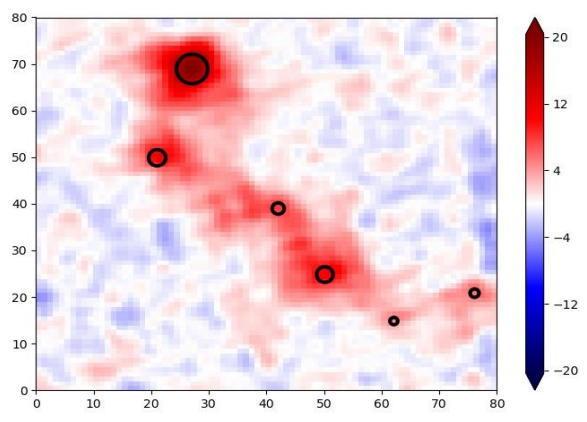


The NO2 processing has been evaluated, then implemented by ICARE for on-demand running on HPC

PHIDIAS
Prototype of HPC/Data Infrastructure for On-demand Services

Product : map and results output file

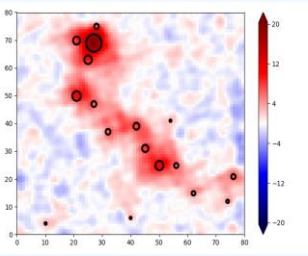
Geismar : 6 significative sources detected, located and characterized



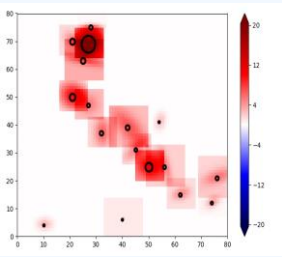
Source index	latitude index	longitude index	Latitude (degree)	Longitude (degree)	Strength (ton/year)	σ_x (km)	σ_y (km)	σ_θ (degree)	background value (ton/year)	Fit value
1	69	27	30.488	-91.174	1090.09	3.06	3.51	207.54	4.48	1.34
2	50	21	30.317	-91.237	965.92	4.23	3.61	208.55	0	1.47
3	25	50	30.092	-90.935	436.03	2.87	3.73	271.12	3.98	0.75
4	39	42	30.218	-91.018	388.9	2.44	5.88	135.58	2.58	1.44
5	21	76	30.056	-90.663	174.43	1.74	4.19	70.95	1.73	0.65
6	15	62	30.002	-90.81	78.4	1.48	2.33	71.72	1.27	0.85

Evaluation

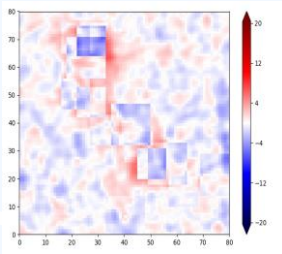
True flux



Retrieved flux



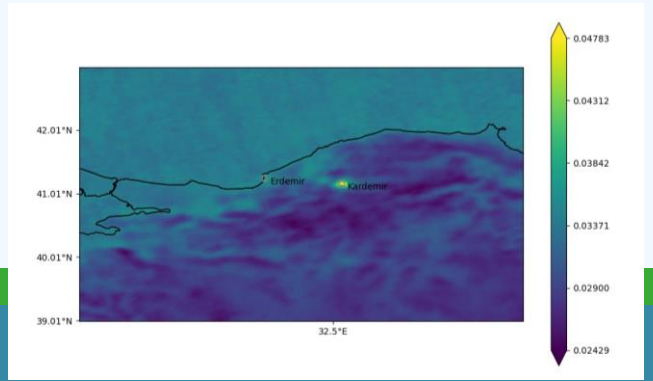
Residual flux (85% of the total emission is retrieved)



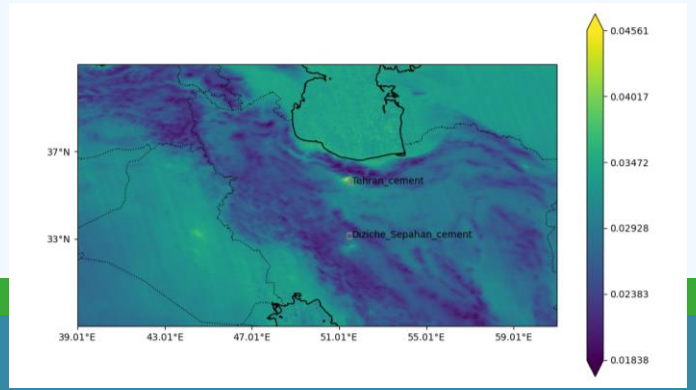
On going work to apply this processing on S5P CO L2 product

Examples of test sites : averaged (3 months) and regridded (1km) CO column maps (Moles/m2)

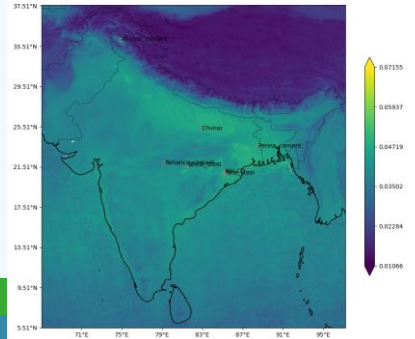
Turkey



Iran



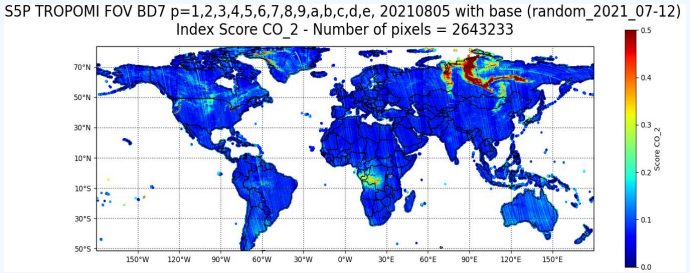
India

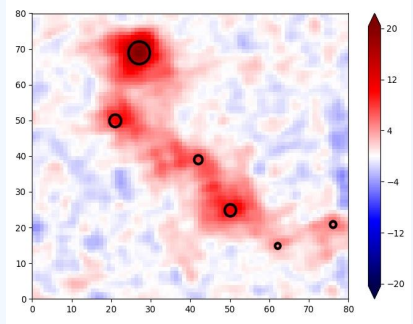


STATUS AND PERSPECTIVES

L1 PCA for extreme events and L2 NO2 (CO ?) source detection prototype will be ready for accessing/processing S5P /TROPOMI products on HPC at ICARE and CINES

- Scientific evaluation of interest and usages of exploring and screening huge volume of data
- Consolidation for on the flow and/or on demand services





Next steps : evaluating the added values for air quality and climate challenges

- Historical series and statistics on extreme events ? Real time monitoring and alert ? Comparison/complementarities with TROPOMI L2 products ?
- Adressing CH4 analysis and monitoring with these 2 complementary processings ?
- …

Perspectives : exploiting the existing and comming space-based atmospheric sounders

- Combined exploitation of several atmospheric sounders (e.g., IASI + S5P)
- Preparing the exploitation of next generation instruments : S5 and S4, IASI-NG, MTG-IRS

PHIDIAS

Prototype of HPC/Data Infrastructure for On-demand Services

JOIN OUR COMMUNITY

www.phidias-hpc.eu

@PhidiasHpc

Company/phidias-hpc