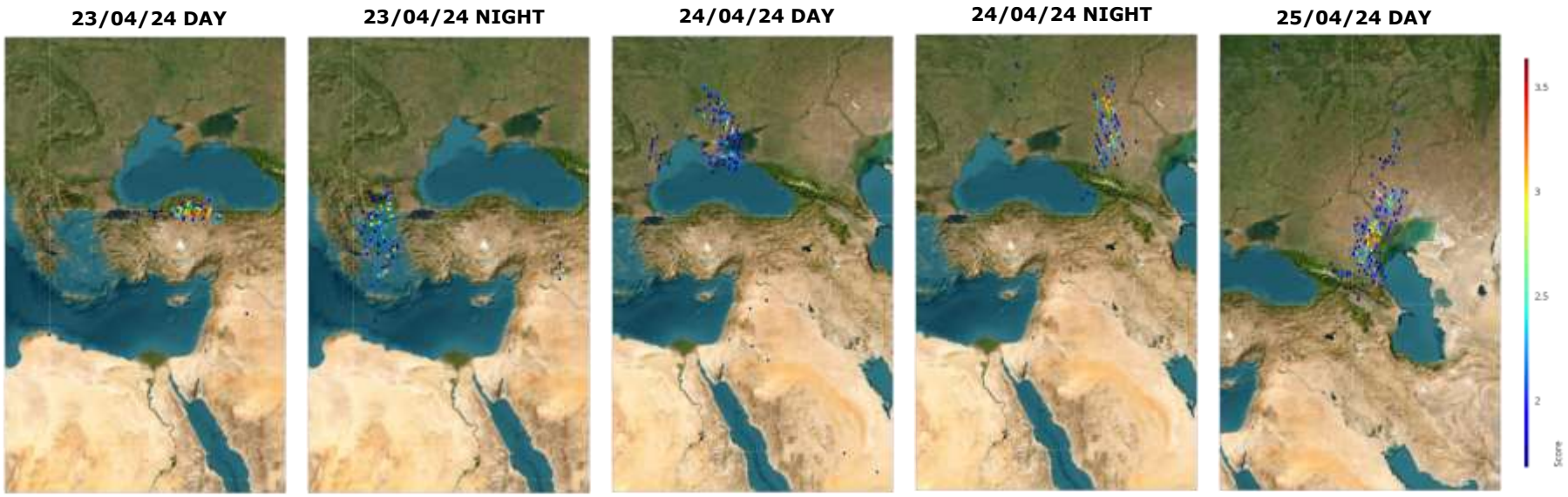
Principal Component Analysis of IASI Level 1 measurements for the detection of atmospheric events



Detection of a desert dust plume around Black Sea, April 2024

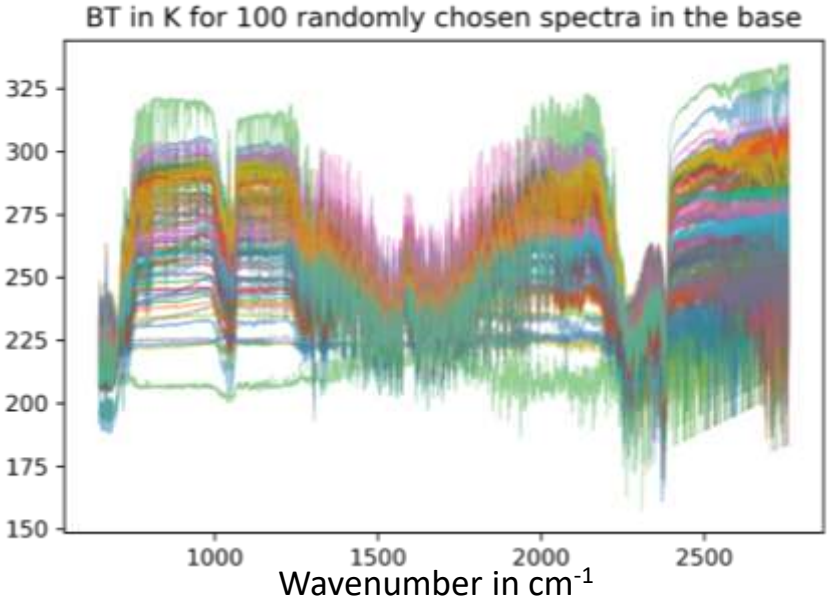
Sarah Pipien[1], Pascal Prunet[1], Claude Camy-Peyret[2,1], Dominique Jolivet[3], Bruno Monsterleet[3], Anne Boynard[1,4], Cathy Clerbaux[4], Jean-Baptiste Joan[5], Nicolas Pascal[5], Patrice Henry[6], Laura Le Barbier[6], Thierry Trema[6]

[1]SPASCIA, [2]IPSL, [3]HYGEOS, [4]LATMOS, [5]AERIS/ICARE , [6]CNES

IASI conference 2024, Nancy 2th December 2024

PCA of the variability of IASI L1 spectral measurements:

Analysing spectra variability to focus on rare /extreme events



Build an ensemble of IASI measured spectra, representative of the entire variability of the measurements :

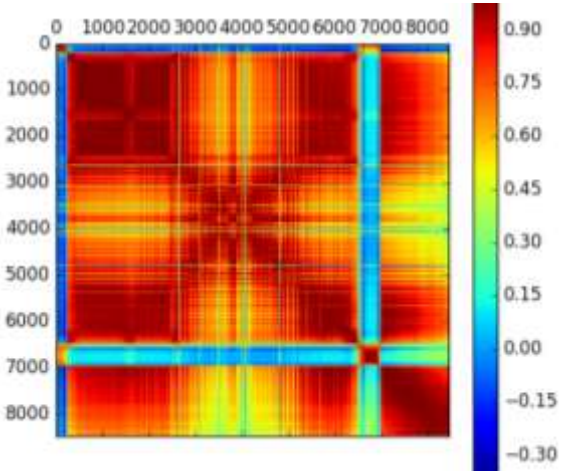
- Global, 1 year (120 000 spectra randomly selected), all conditions (air masses, surfaces, clouds and aerosols)

From the reference set (normalised by the instrument noise)? Compute the covariances matrix of the centred normalised spectra (dimension: 8461 x 8461)

- This gives the atmospheric variability as seen by the instrument

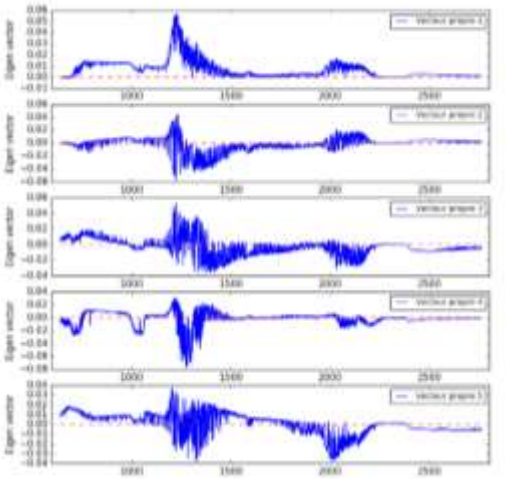
Apply PCA : decomposition of 8461 "orthogonal" components of the measurements from the most variable one to the least variable one.

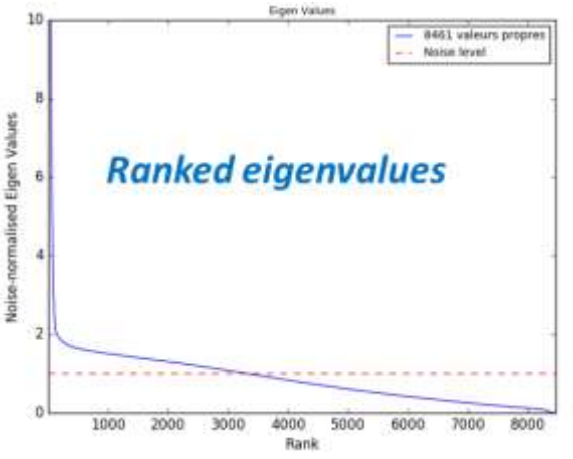
- Eigenvectors : orthogonal principal direction of variability
- Eigenvalues : measure of the variability represented by a given component



Off-diagonal : spectral corelations

(Diagonal : spectral variance)

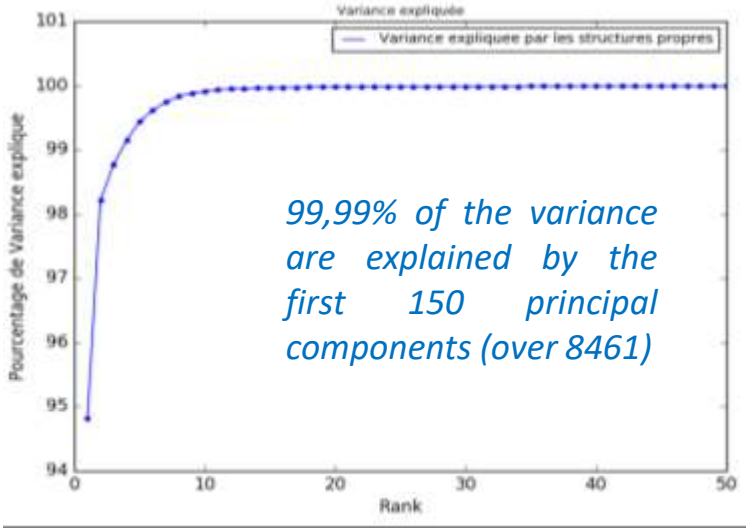




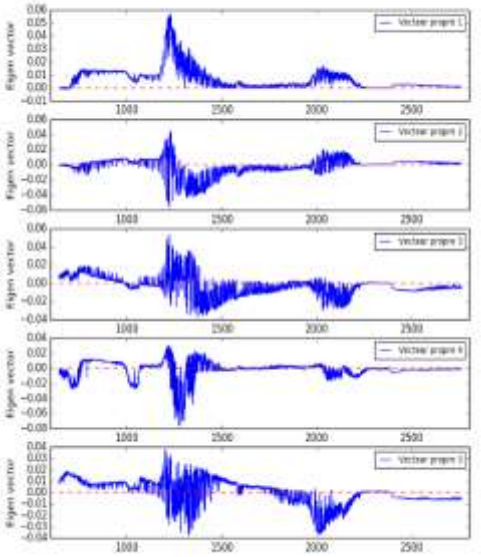
PCA of the variability of IASI L1 spectral measurements

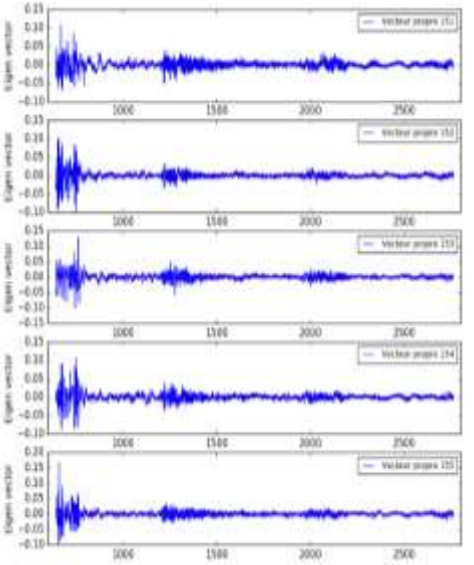
Because the atmosphere/surface variability is spectrally correlated in the spectra, the number of significant eigenvectors, that contains the signal, is drastically smaller than spectrum dimension.

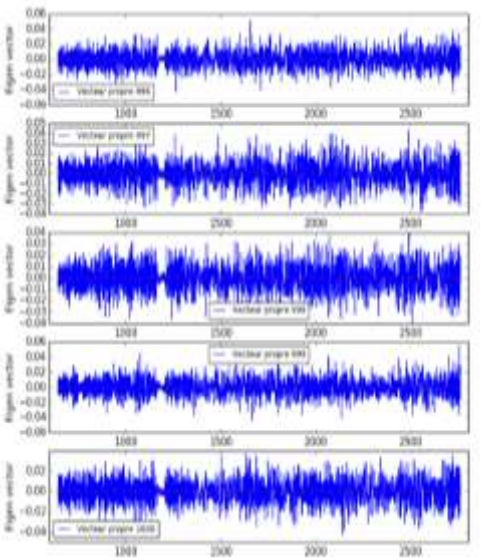
Because the measurement is dominated by atmospheric signal, and because the noise is spectrally uncorrelated, the noise is concentrated in the less significant eigenvectors.



99,99% of the variance are explained by the first 150 principal components (over 8461)







- **Possibility to separate signal and noise in the measurement**
- **Significant eigenvectors, (i.e. principal components (or PC) from rank 1 to ~ 150-200) are ranked from strongly variable signal (surface and clouds) to the smallest variability captured by the statistics (trace gases or rare events).**

Significative experience with IASI PCA at EUMETSAT for compression and denoising …

IASI-PCA approach for detection

Nominal case: compression and noise filtering
Strong experience from EUMETSAT with the IASI operational L1D (PC compressed)

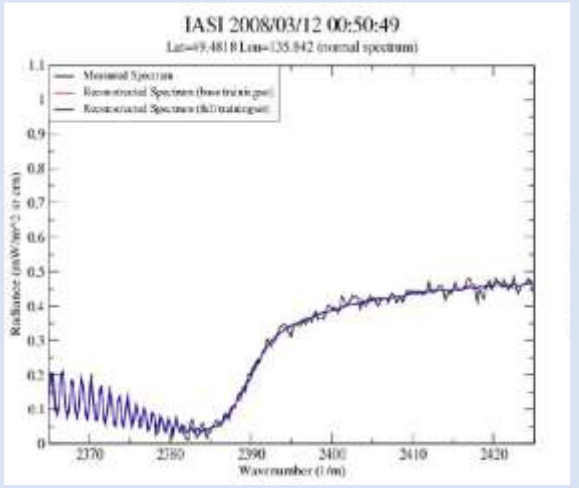
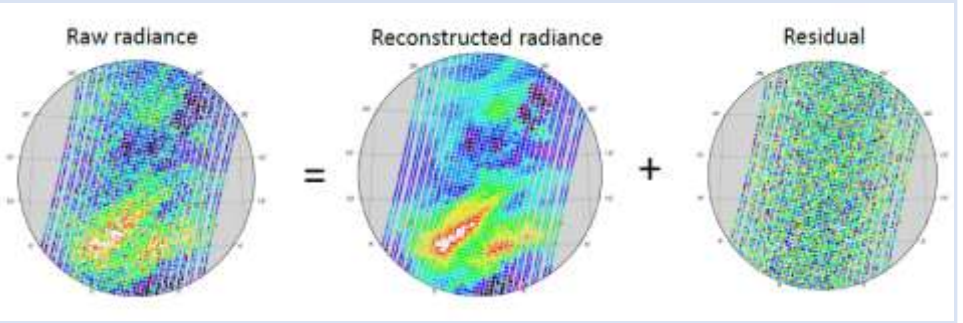
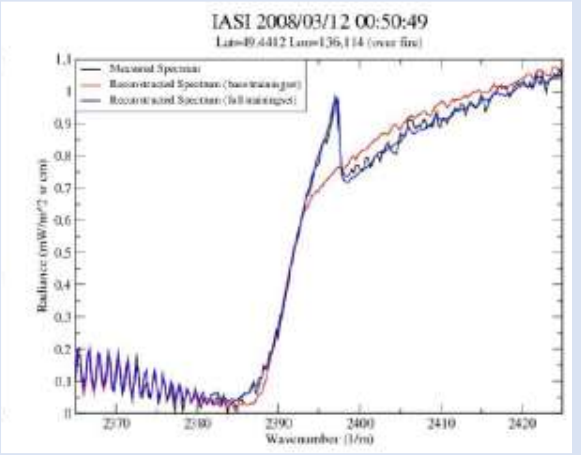


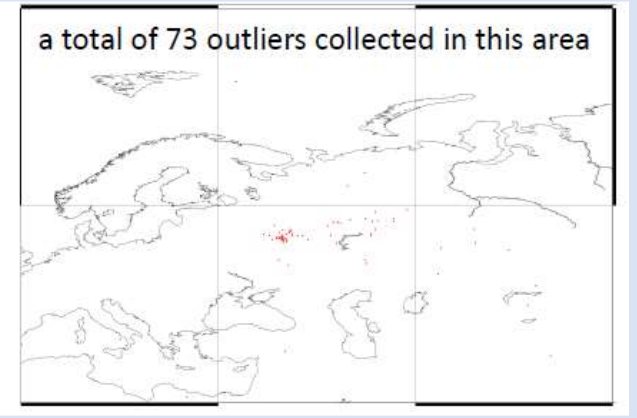
Illustration of IASI PCA decomposition : credit EUMETSAT



Exceptional cases : outliers detection and residual signal filtering
Potential for rare event detection and identification

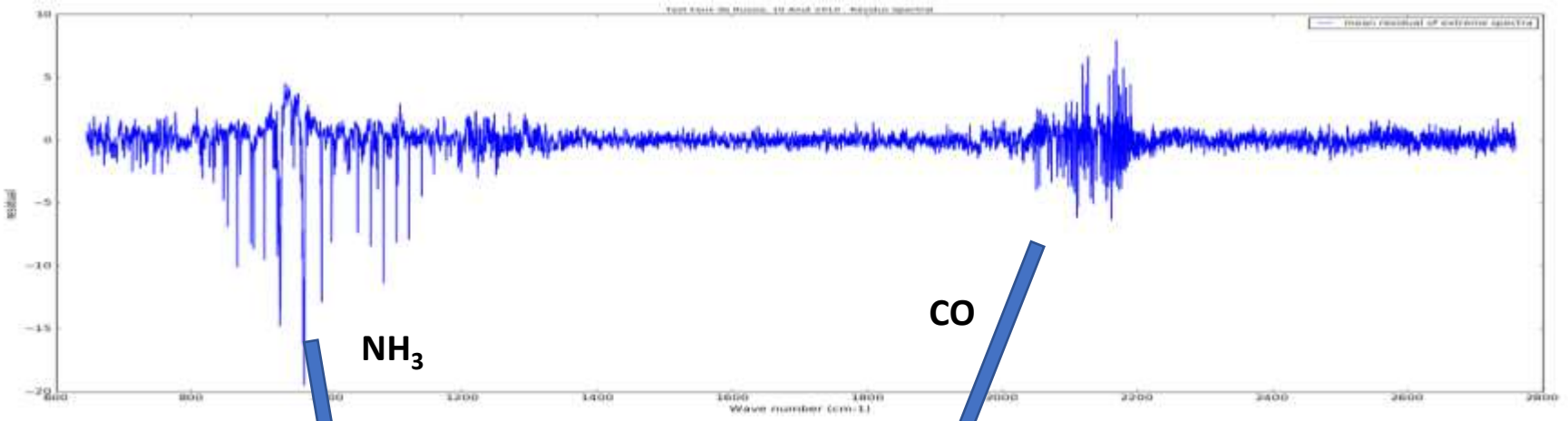


Reconstruction scores : **Fast detection of exceptional observations**



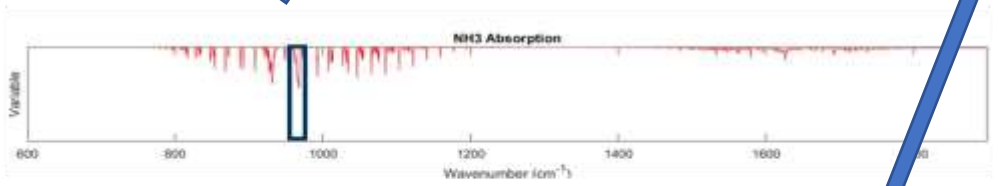
Analysis of the residual signal: **event identification and characterisation**

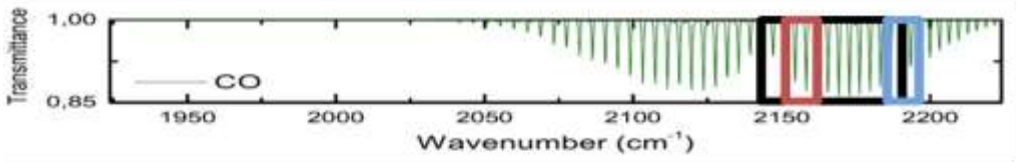
How to deal with the residuals : reconstruction score and indicators

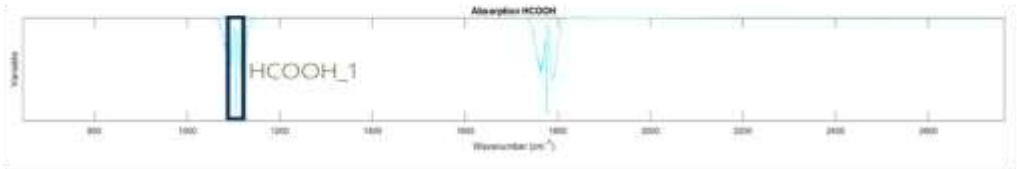


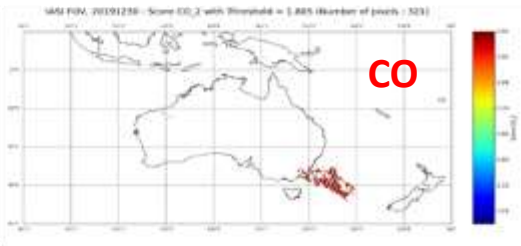
NH_3

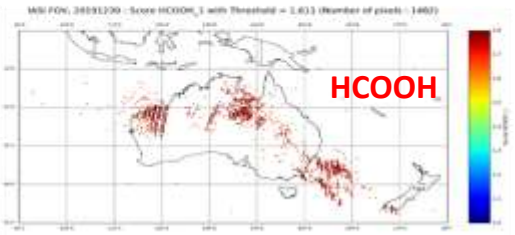
CO











Reconstruction Score (RS) :

$$RS = \sqrt{\frac{1}{m}\sum_{i=0}^{m} r_i^2}$$

It allows the detection of outiers : a high score (larger than 1) indicate that the residual spectrum contains signal (and not only noise as in nominal situations).

Indicators : dedicated reconstruction scores targeted on a limited spectral interval (associated to a species) :

$$\text{Indicator} = \sqrt{\frac{1}{m_{\mu_1\mu_2}}\sum_{\mu_1}^{\mu_2} r_i^2}$$

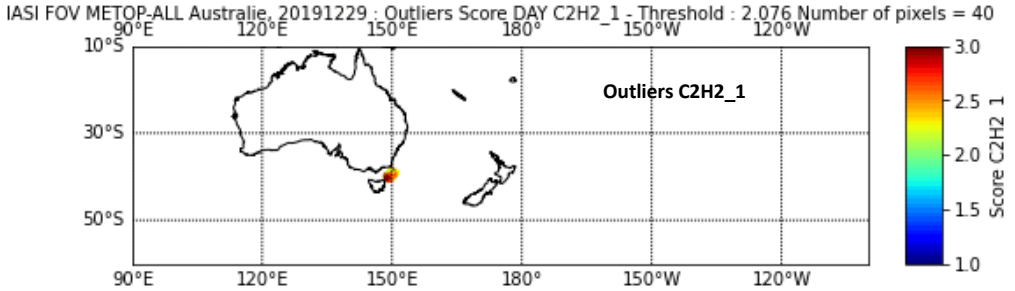
Focus on the $[\mu_1\ \mu_2]$ spectral domain

ILLUSTRATION : Detection and identification of NH_3, CO and HCOOH in Australian fire plume

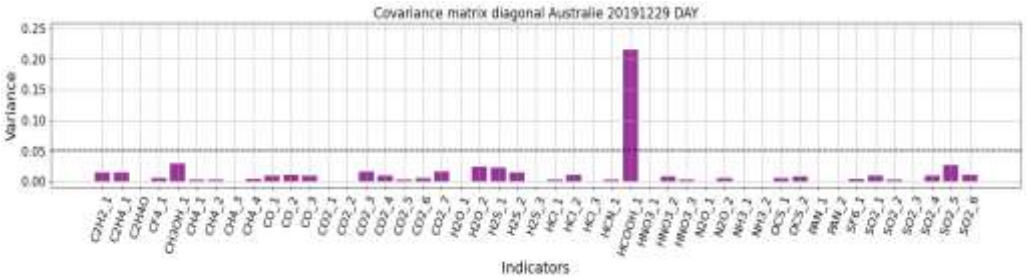
First tests and validations : detection and identification of fires from IASI

A larger number of fire events has been processed and analysed from several years of global IASI data

Australia Dec. 2019 / Jan. 2020



Testing and consolidating automatic detection and characterisation of fires on a daily basis, from indicators maping and code-detection information.



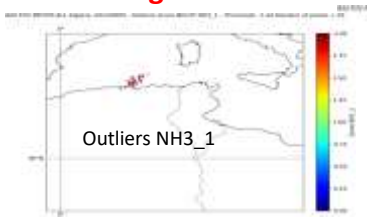
Example of Autralian fires : CO, NH_3, HCOOH, C_2H_2, C_2H_4, CH_3OH, …

➢ Detection, location, Species identification from the analysis of code-detection

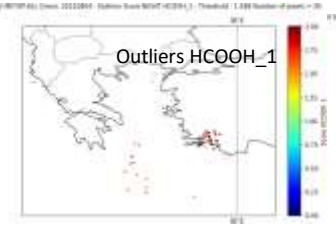
2021 summer fires

Europe

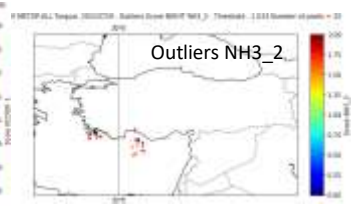
Algeria



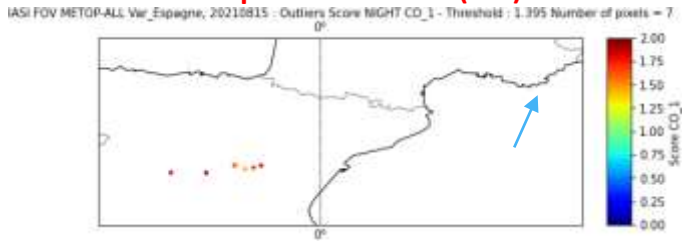
Greece



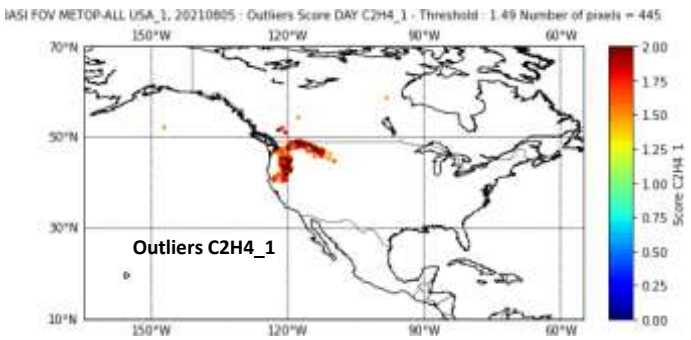
Turkey



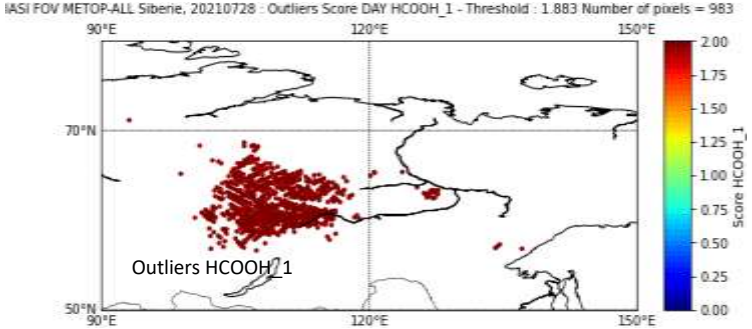
Spain and France (Var)



United-States



Siberia



Detection capabilities in “all weather” conditions

This processing has been evaluated and tested on different datasets for case studies and intercomparisons

➢ ability of the processing to work on large amounts of data and to process clear as well as cloudy conditions.

Indonesia, October 2015

Fire plume over clouds

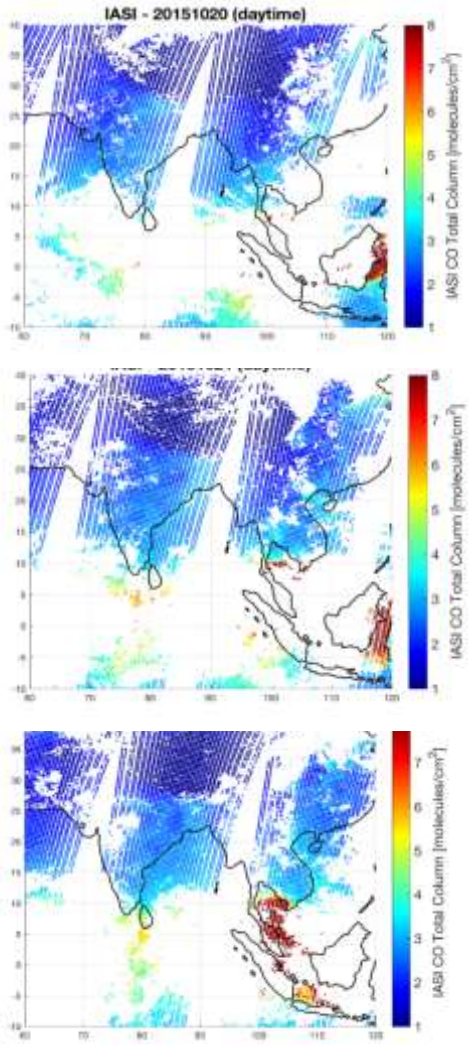
Detected by :

IASI CO L2 product (left)

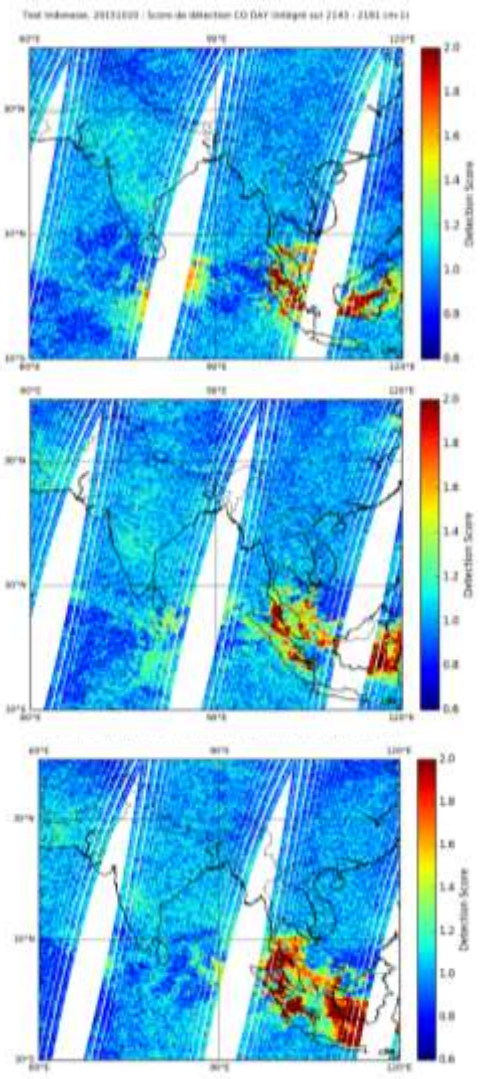
IASI-PCA product (middle)

Cloud fraction (right) show the detection over clouds by IASI-PCA

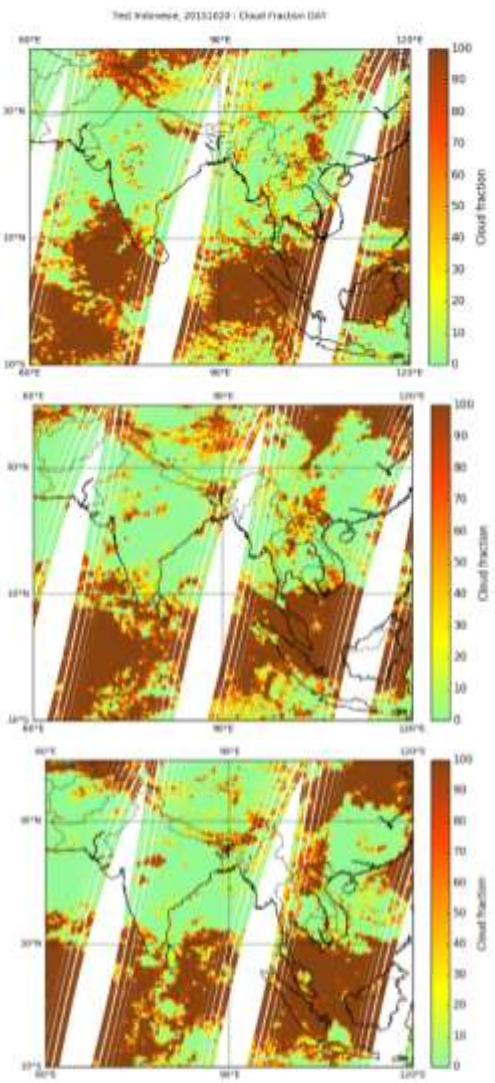
CO IASI L2 product



IASI-PCA score CO



Cloud fraction



Detection capabilities in “all weather” conditions : Argentine, June 2024

This processing has been evaluated and tested on different datasets for case studies and intercomparisons

- ability of the processing to work on large amounts of data and to process clear as well as cloudy conditions.

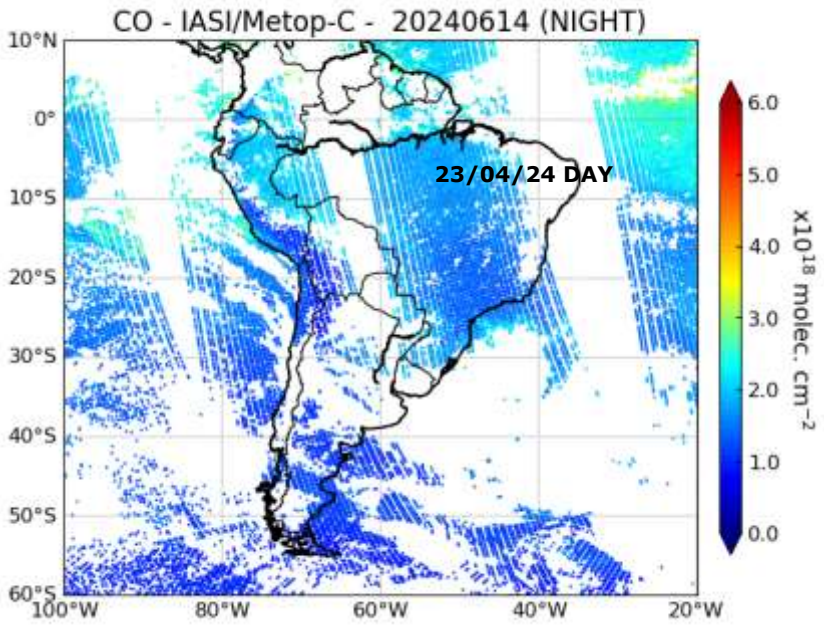
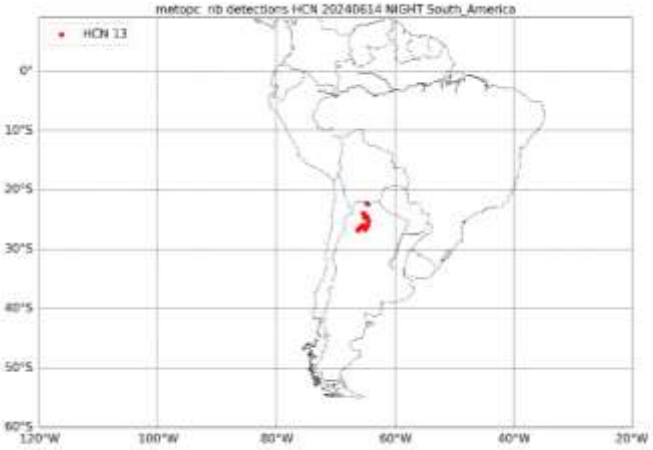


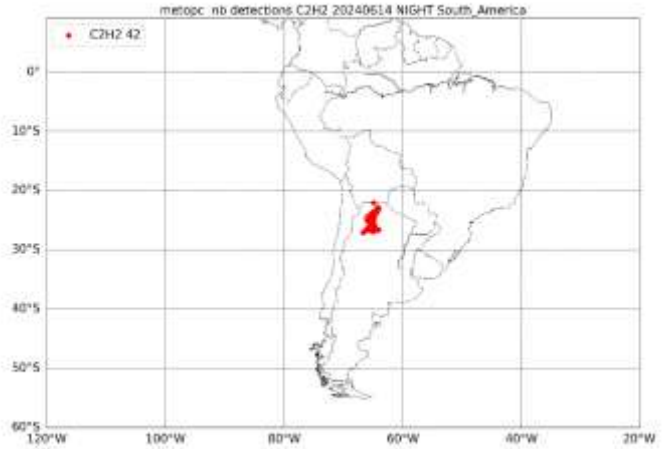
Illustration of the detection and identification of a fire plume over clouds

Not detected by IASI L2 CO or other products

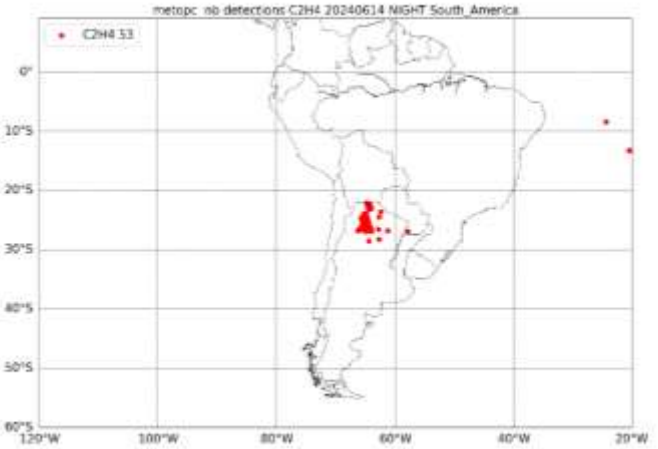
HCN indicator – IASI-C – 20240614 NIGHT



C_2H_2 indicator – IASI-C – 20240614 NIGHT



C_2H_4 indicator – IASI-C – 20240614 NIGHT

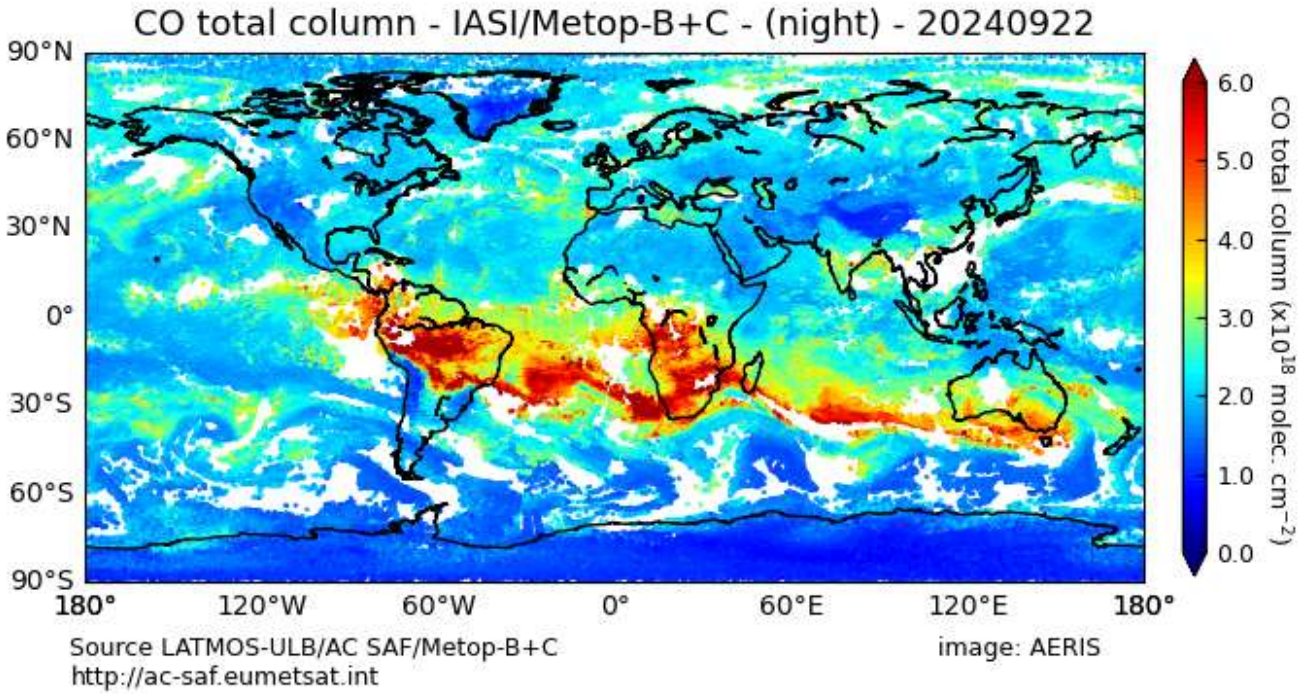


Amazonia fires, September 2024

This processing has been evaluated and tested on different datasets for case studies and intercomparisons

- ability of the processing to detect and identify plume of different species (CO, CO_2, VOCs).

IASI L2 CO product



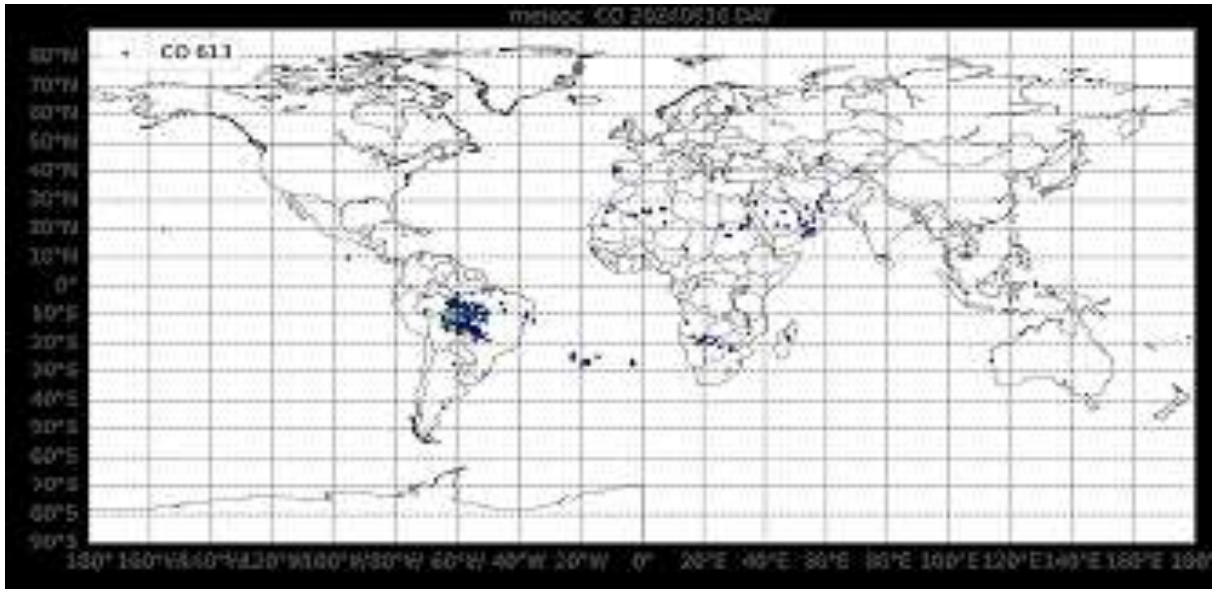
Source LATMOS-ULB/AC SAF/Metop-B+C
http://ac-saf.eumetsat.int

image: AERIS

Amazonia fires well observed and quantified from IASI L2 CO plume

IASI-PCA

CO indicator – IASI-C – from 16 to 25/09/2024 (DAY and NIGHT)

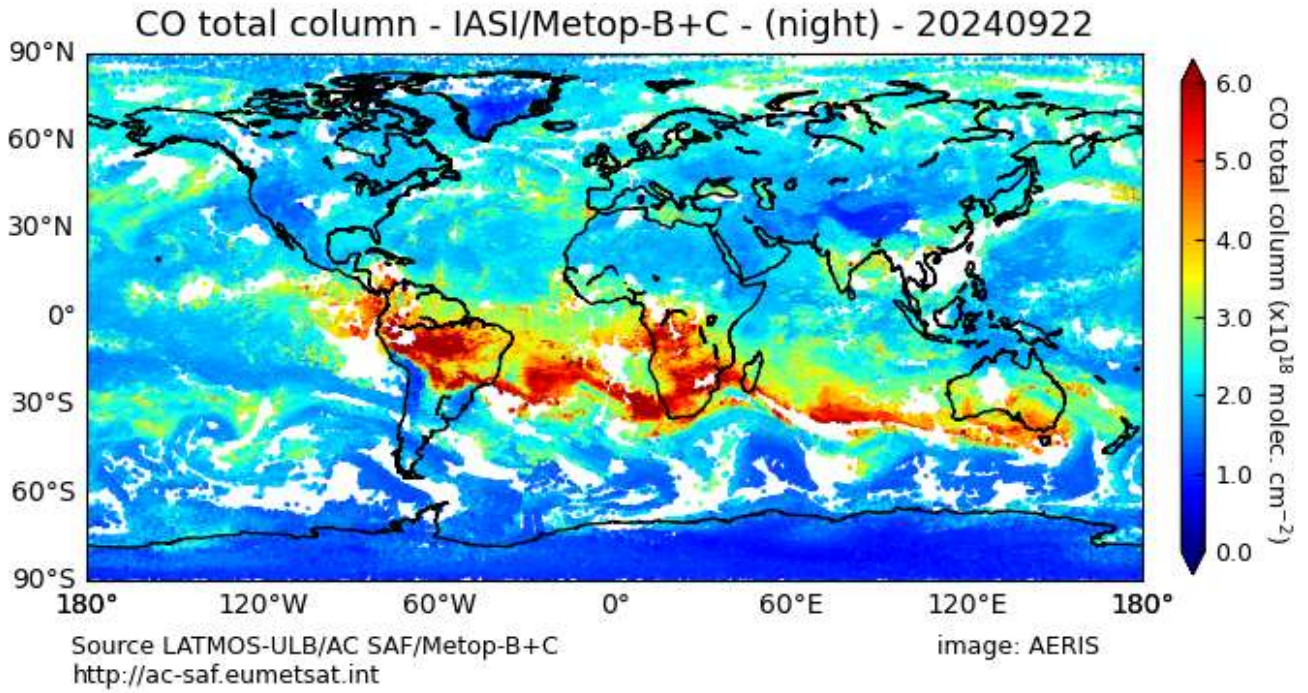


Amazonia fires also detected and identified by PCA for CO species

Multi-species detections : Amazonie fires, September 2024

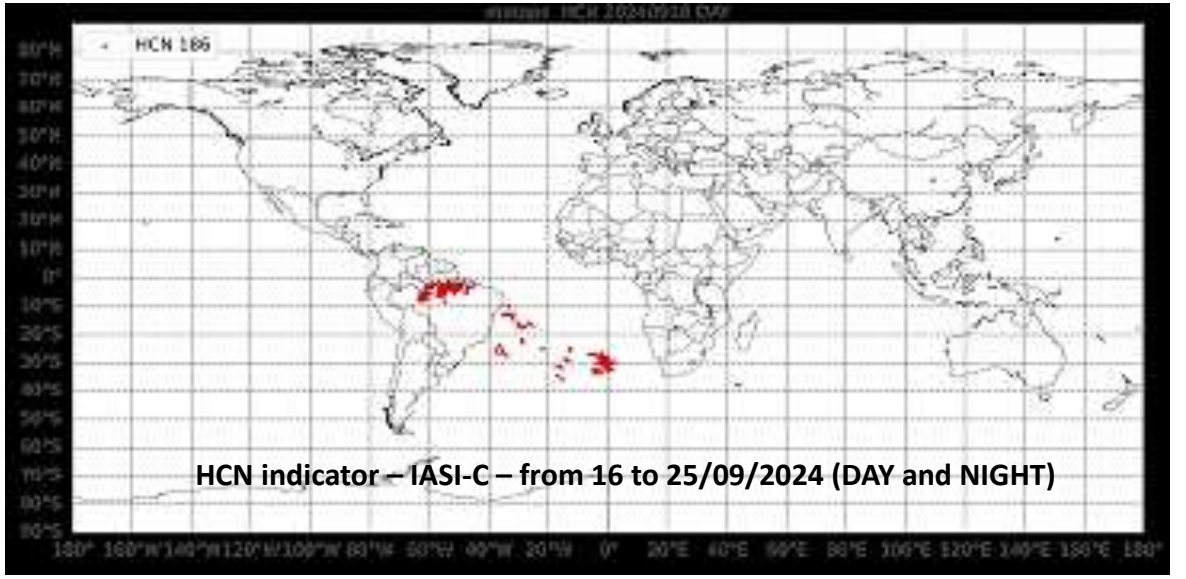
This processing has been evaluated and tested on different datasets for case studies and intercomparisons

- ability of the processing to detect and identify plume of different species (CO, CO_2, VOCs).

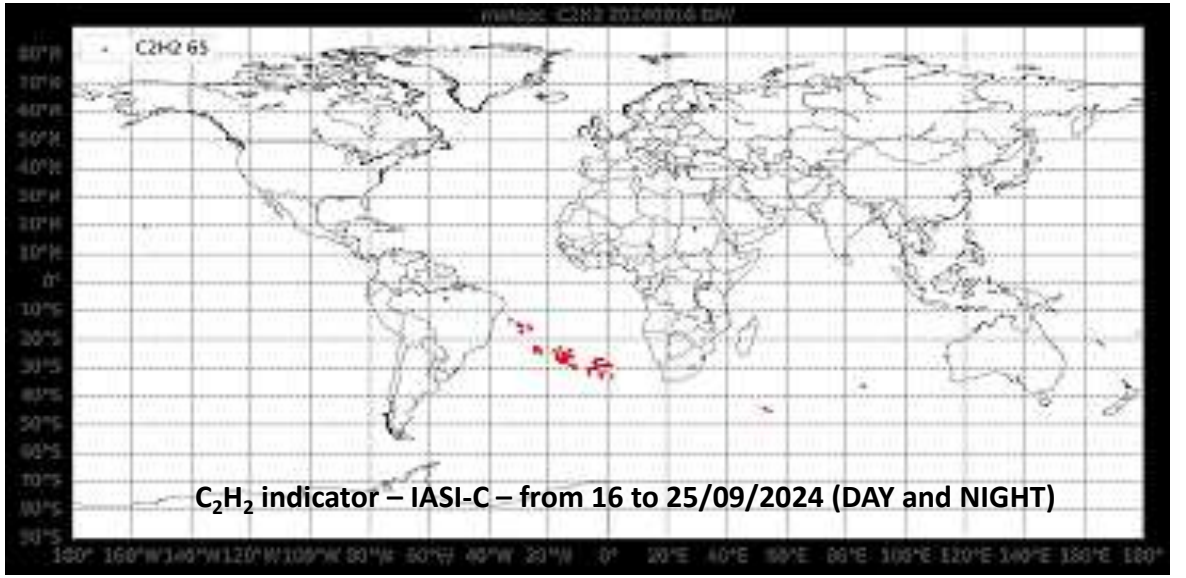


Source LATMOS-ULB/AC SAF/Metop-B+C
http://ac-saf.eumetsat.int

image: AERIS



HCN indicator – IASI-C – from 16 to 25/09/2024 (DAY and NIGHT)



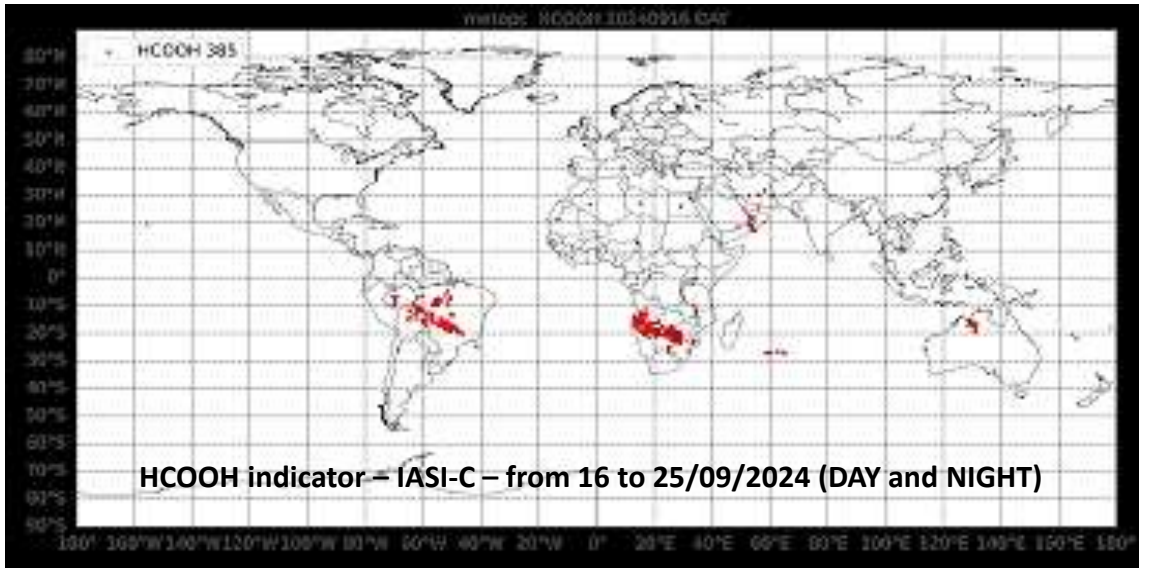
C_2H_2 indicator – IASI-C – from 16 to 25/09/2024 (DAY and NIGHT)

Amazonie fires also detected and identified by PCA for HCN, C_2H_2, HCOOH, CH_3OH

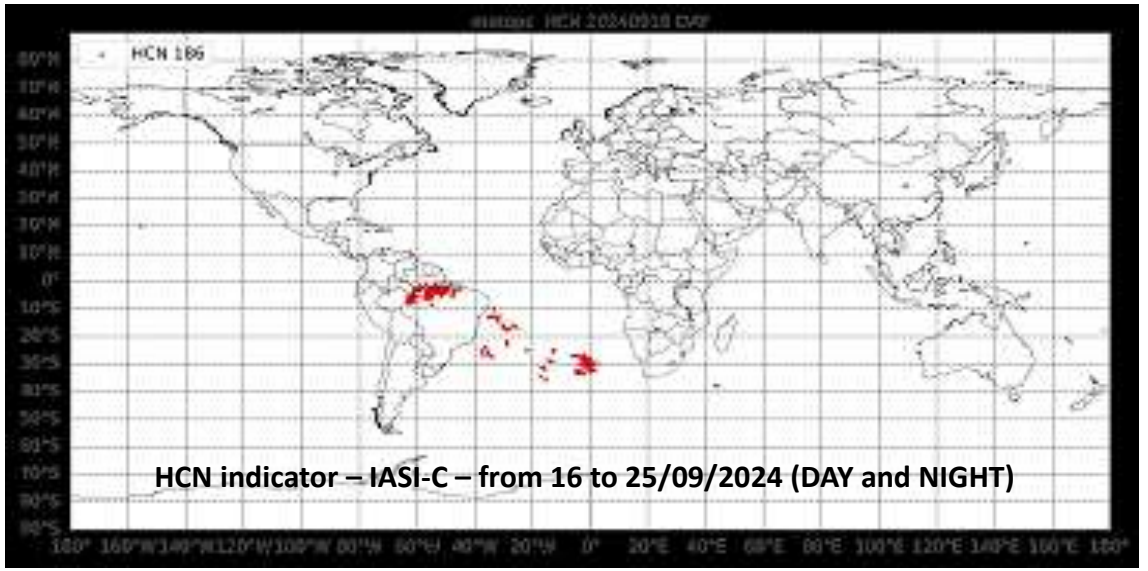
Multi-species detections : Amazonie fires, September 2024

This processing has been evaluated and tested on different datasets for case studies and intercomparisons

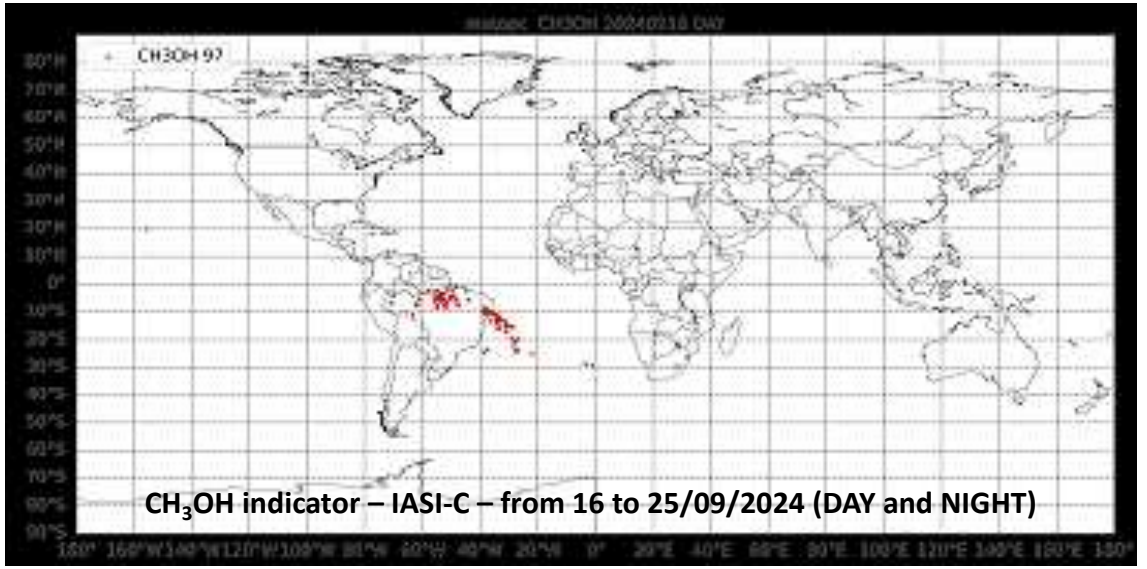
- ability of the processing to detect and identify plume of different species (CO, CO_2, VOCs).



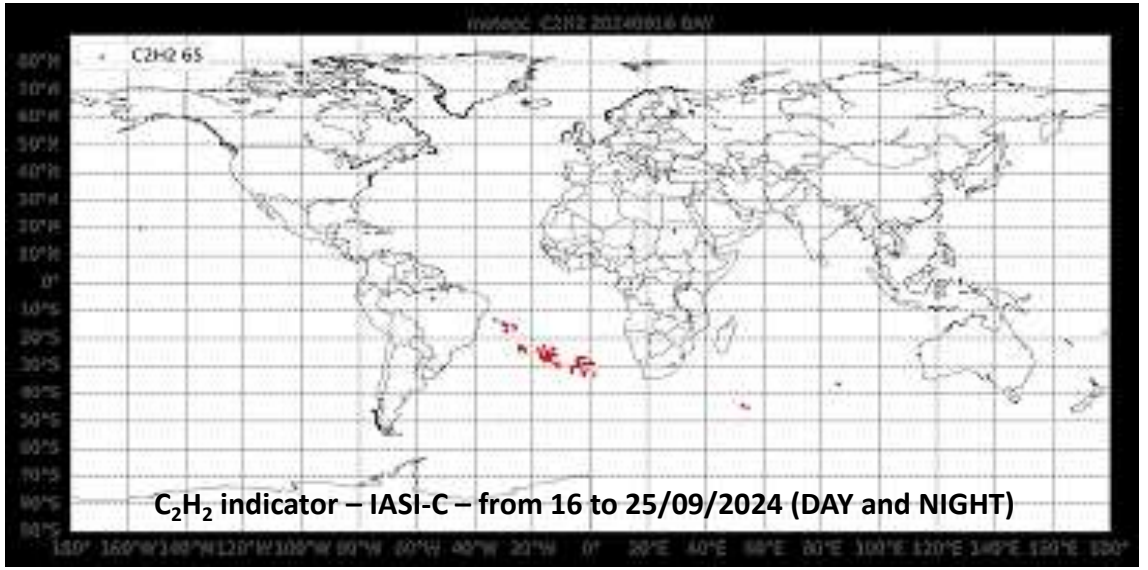
HCOOH indicator – IASI-C – from 16 to 25/09/2024 (DAY and NIGHT)



HCN indicator – IASI-C – from 16 to 25/09/2024 (DAY and NIGHT)



CH_3OH indicator – IASI-C – from 16 to 25/09/2024 (DAY and NIGHT)



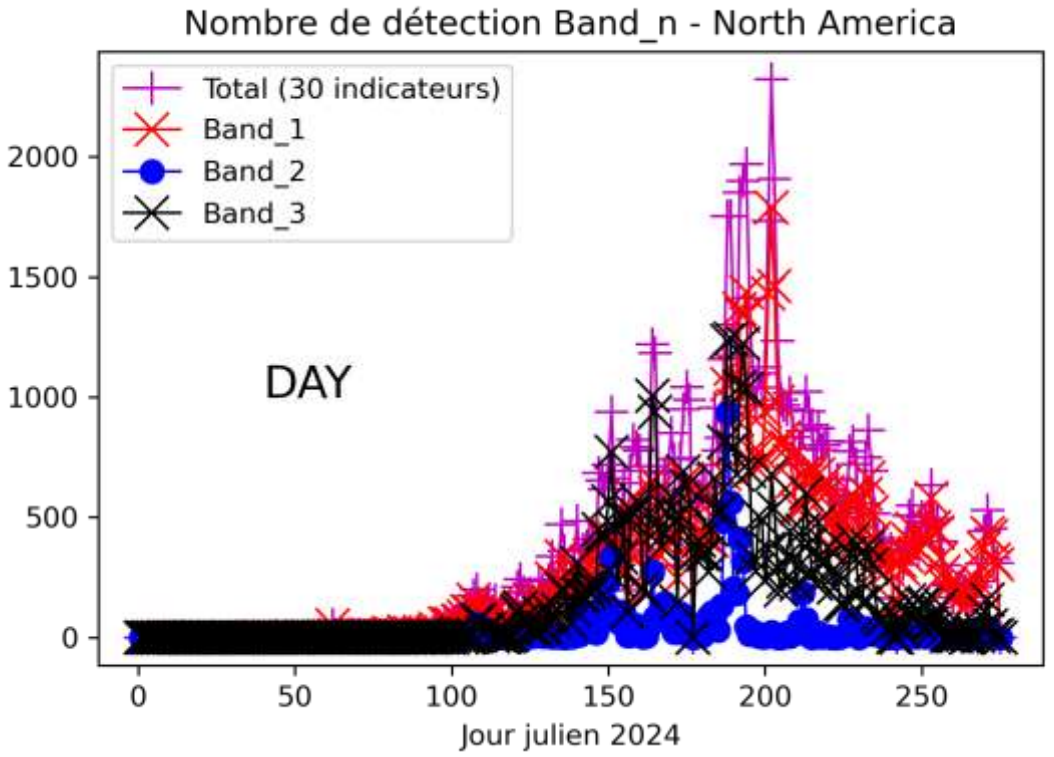
C_2H_2 indicator – IASI-C – from 16 to 25/09/2024 (DAY and NIGHT)

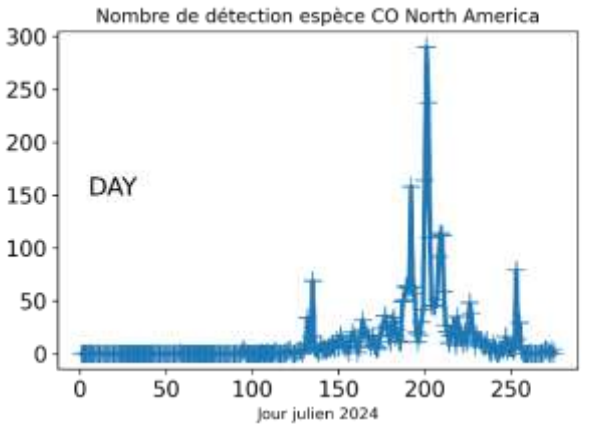
Statistics and monitoring of detections : North America (DAY)

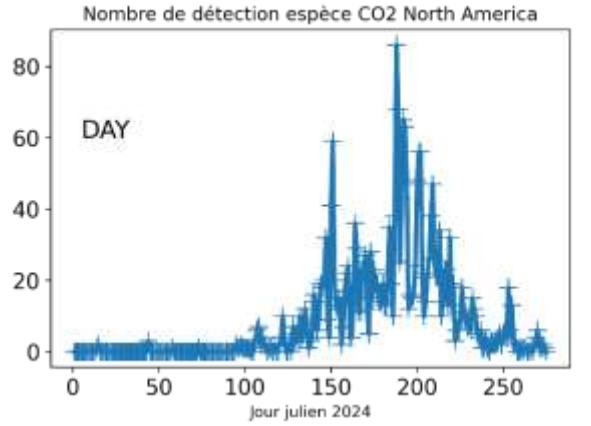
This processing has been evaluated and tested on different datasets for case studies and intercomparisons

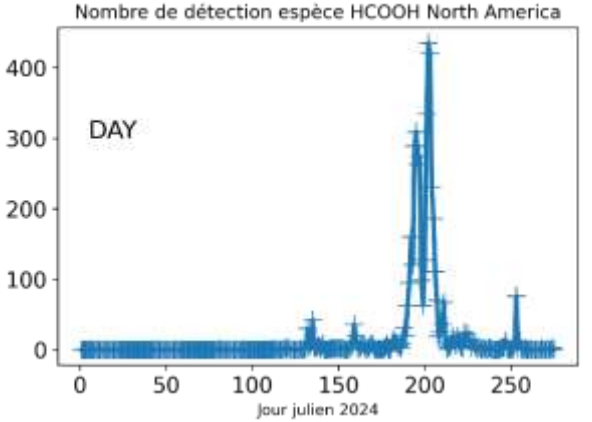
- On going analysis of 9 months of global monitoring of events

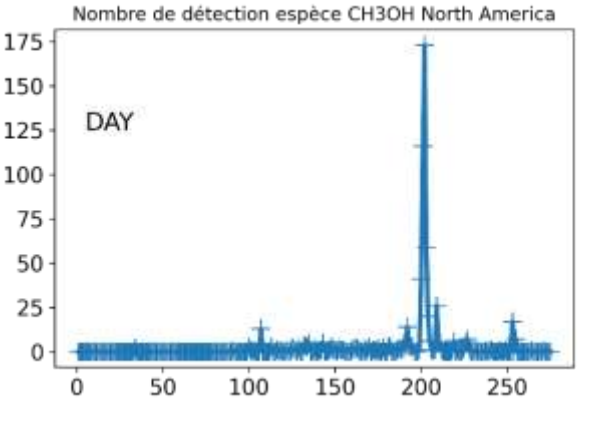
Number of daily detections of events

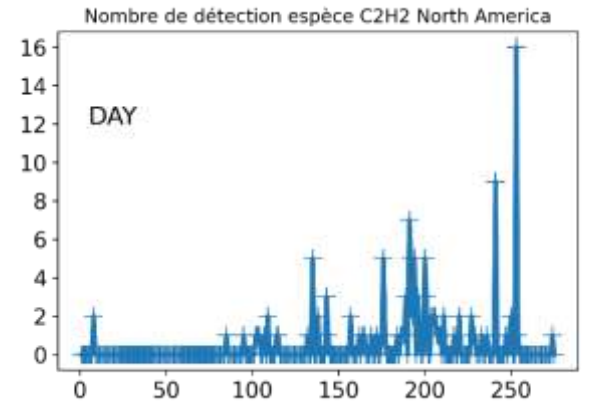


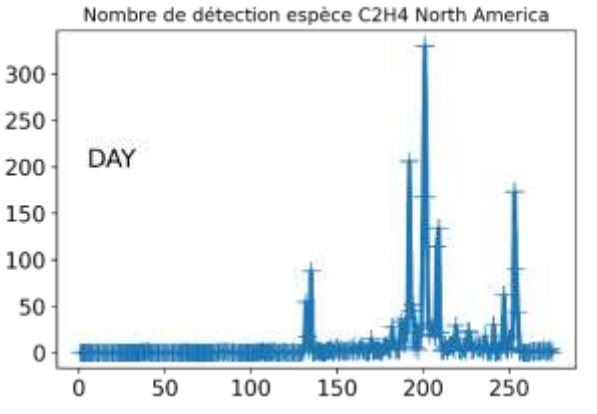








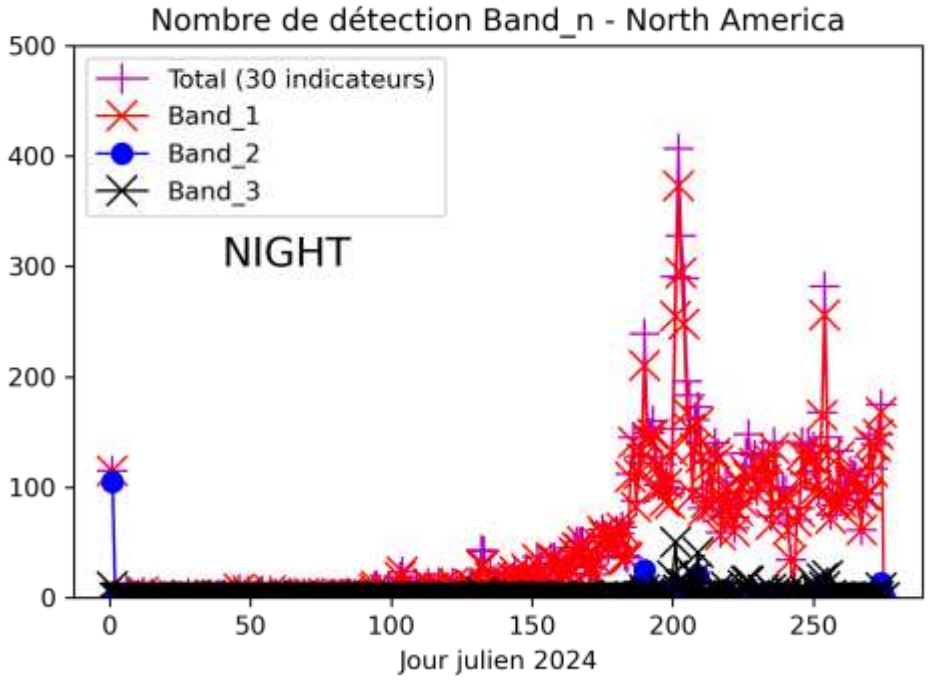


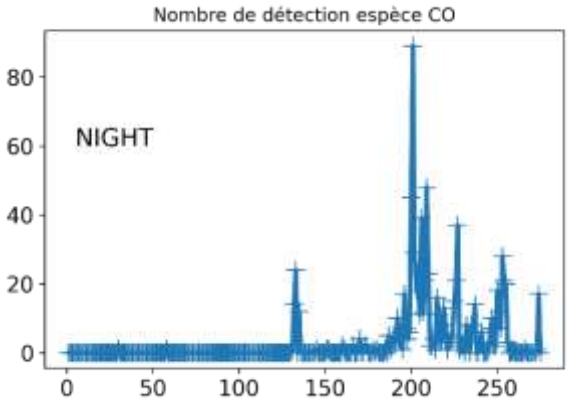


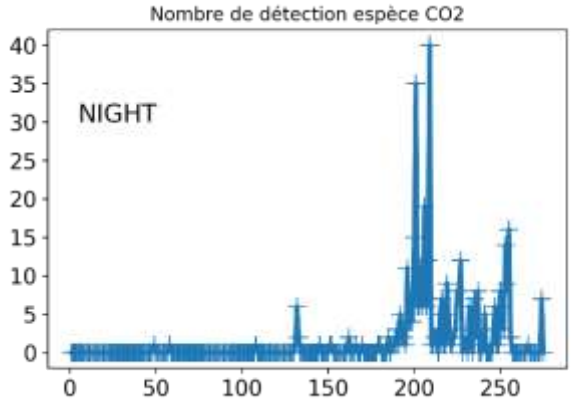
Statistics of detection : North America (NIGHT)

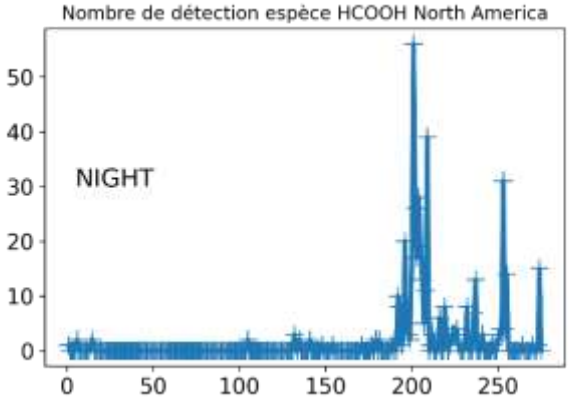
This processing has been evaluated and tested on different datasets for case studies and intercomparisons

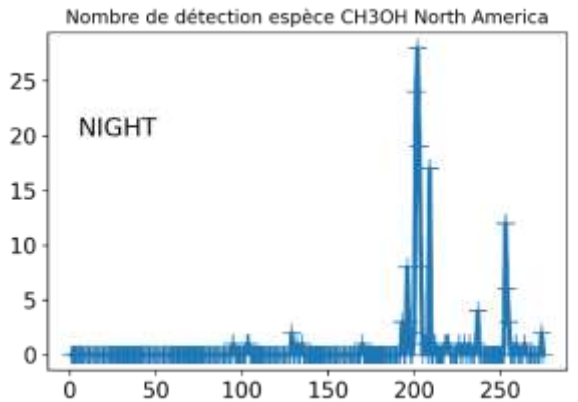
- On going analysis of 9 months of global monitoring of events

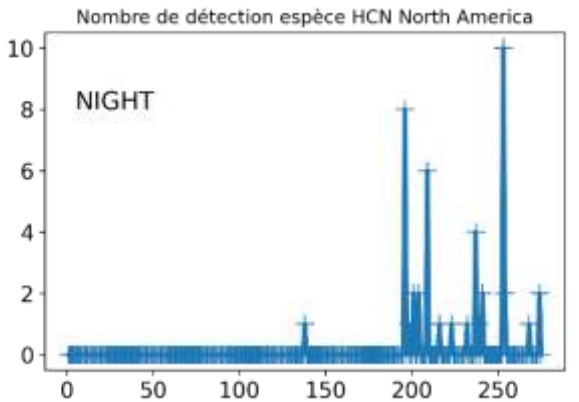


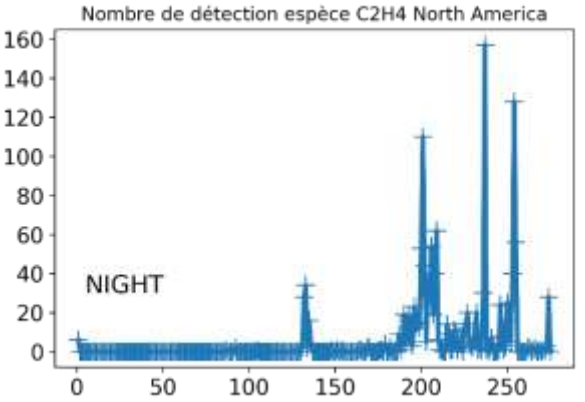










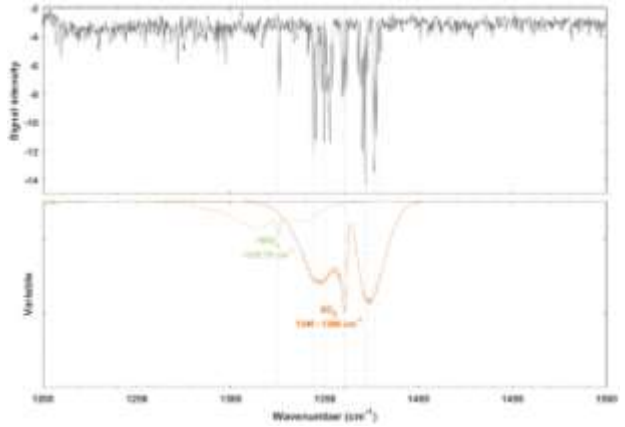


On-going analyses : detection of unexpected events

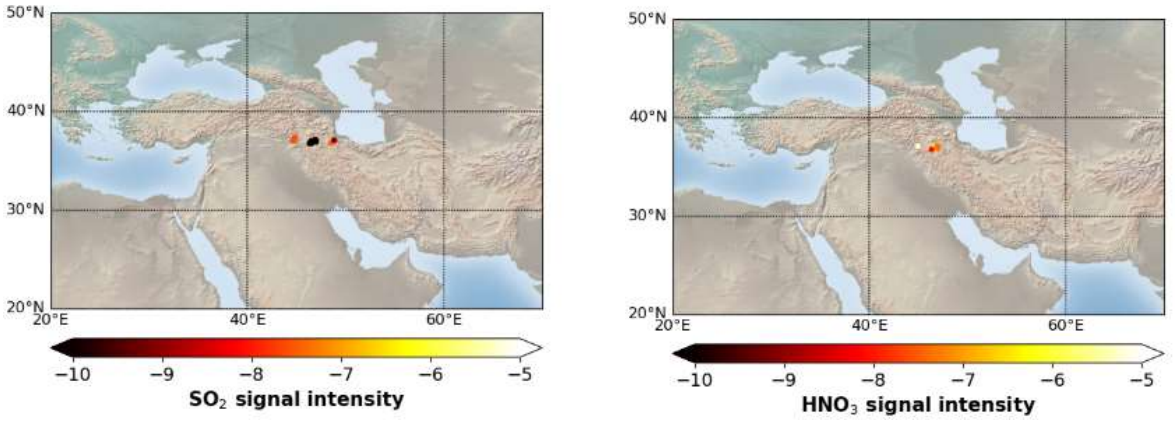
We are exploring the capability of IASI-PCA approach de detect signature of unexpected events

Al-Mishraq (Iraq) sulfur mine burning, October 2016:

This fire on the sulfur plant, which was set by Islamic state, caused a large emission of SO_2 and other sulfured species in the atmosphere, which was observed from several satellite instruments [Björnham et al., 2017]. PCA-based analysis of IASI-B data by Granule Min NRT approach allows to detect SO_2 and HNO_3 plume for this event (Vu Van et al., 2023)

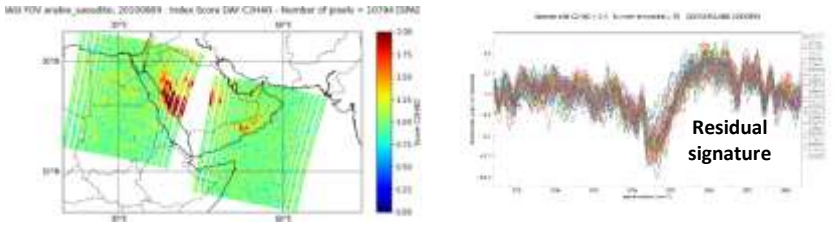


Example of PCA-based Granule Min (GMI) pseudo spectrum on Al Mishraq on 24 October 2016, compared with spectroscopic data. **(Vu Van al al., 2023)**

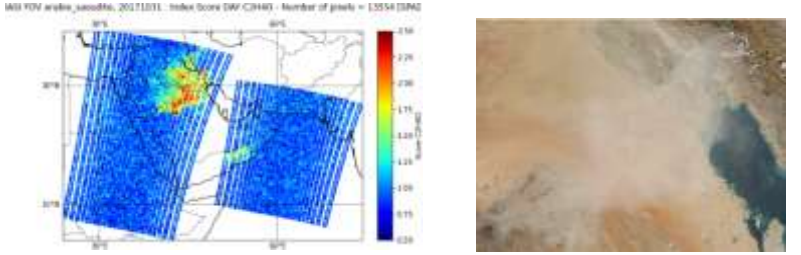


Analysis of sulphur plant fire event in Iraq on 24 October 2016 : spatial distribution of IASI PSA-based residuals residual GMI associated with SO_2, and HNO_3.

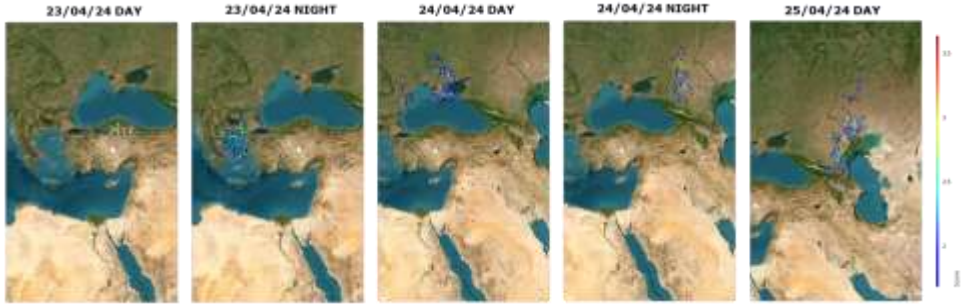
Saudi Arabia, August 2010: the IASI-PCA indicator approach detected an isolated desert dust plume from the calcite spectral signature



This event as well as similar ones have been compared with MODIS datasets, and can be systematically associated with sand dust plumes.



Exemple of monitoring of a (Calcite) desert dust plume over Black Sea, April 2024



Routine, NRT IASI PCA events detection and identification

Implemented on AERIS

In test phase from January 2024, public access current 2025

IASI L1c (format BUFR)

IASI-PCA

log_event_yyyymmdd_hhnnss_Px par numéro d'IFOV, x=1,2,3,4

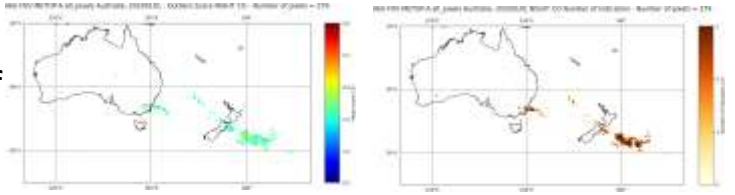
Regroupement des Px + reclassement des IFOVs par *hhnnss*

log_event_yyyymmdd_hhnnss

Accumulation sur 24 h à partir de 0:00 UT + élimination des IFOVs isolés (1) +moyenne pondérée par espèce

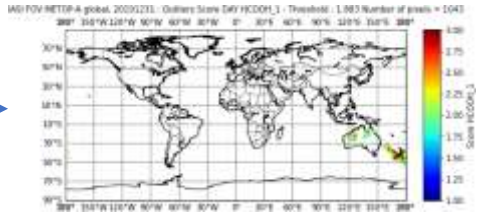
scores_yyyymmdd_hh$_1$nn$_1$ss$_1$-hh$_2$nn$_2$ss$_2$_XX ou scores_yyyymmdd_hhnnss_XX

Hourly detection of 13 species + scores



Fichier complet avec des IFOVs concaténés par jour calendaire + élimination des IFOVs isolés (1)

log_event_yyyymmdd



Daily detection of 30 indicators and scores + FOV detections information

PCA *code_detection*

Diag(Covariance_ *code_detection*)

EV_m et Va_m (m=1,N) + liste des indicateurs pour chaque_m

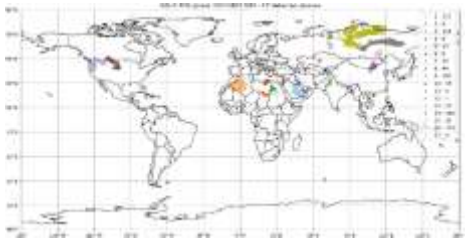
Filtrage + Clustering + analyse des panaches

Attribution du type d'évènement à partir de la liste des indicateurs

Evènements classifiés

Classes of events : Fires (several types), Volvanoes, Others

Map of event + Meta-data

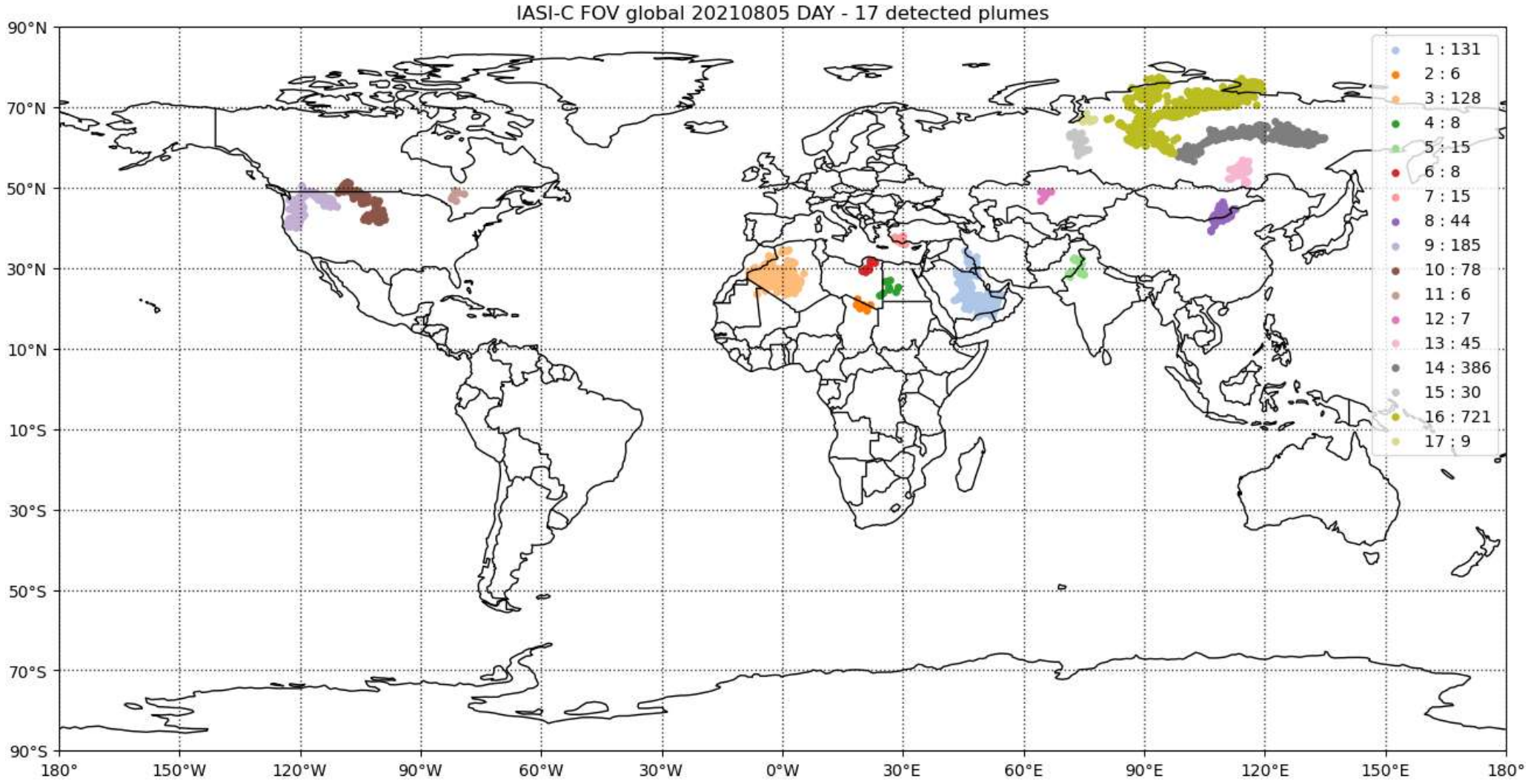


Global daily map of plumes + list of species in each plume + identification of extreme isolated FOVS

Catégorie	Nombre de panaches détectés
Feux type 1	1
Feux type 2	0
Volcans	0
Autres	0

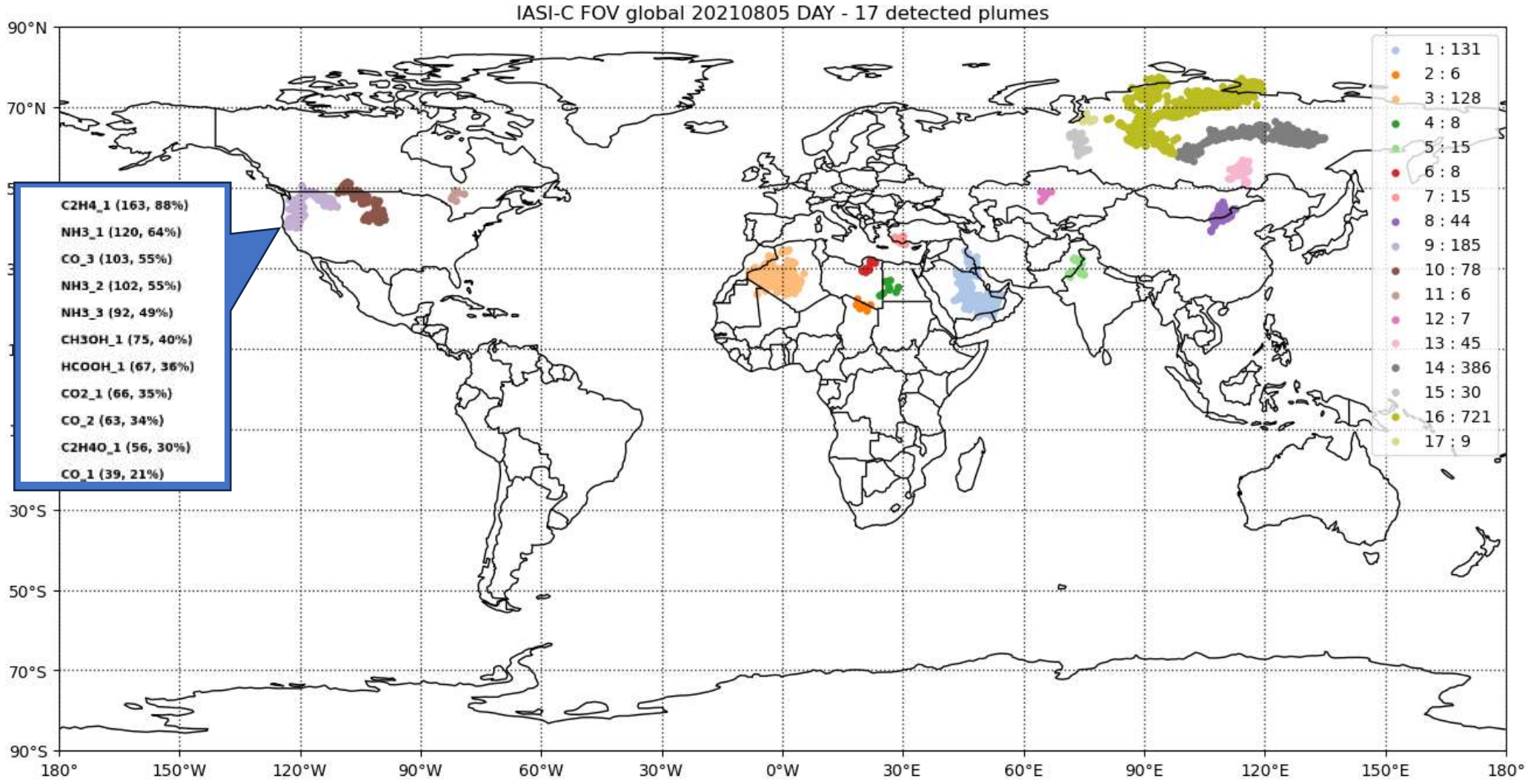
Routine, NRT detection and identification of events

Exemple of L3 product : Identified plume of events for a given day



Routine, NRT detection and identification of events

Exemple of L3 product : Identified plume of events for a given day



IASI PCA3 Visualization Tool

Explore last data avail | Explore historical data

Date

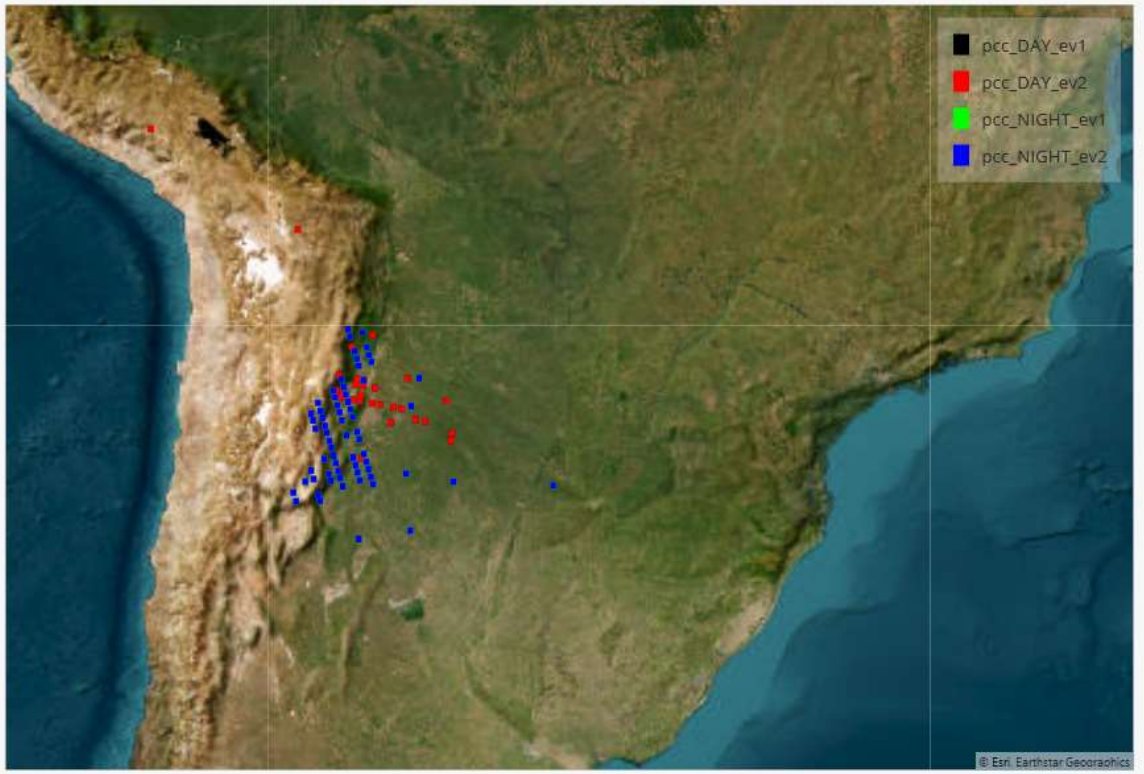
2024-06-13 to 2024-06-18

Date: **14 Jun 2024**

EV	Indicator	Categorie
pcc_DAY_ev1	['CO2_1', 'CO_3']	Feux_type_1
pcc_DAY_ev2	['CO_3', 'HNO3_1', 'CO2_1']	Feux_type_1
pcc_NIGHT_ev1	['C2H4_1', 'HNO3_1', 'C2H2_1']	Autres
pcc_NIGHT_ev2	['HNO3_1', 'NH3_2', 'C2H4_1', '(	Autres

Map of IFOVs classification by event type

|Description of product L3 - Classification event type

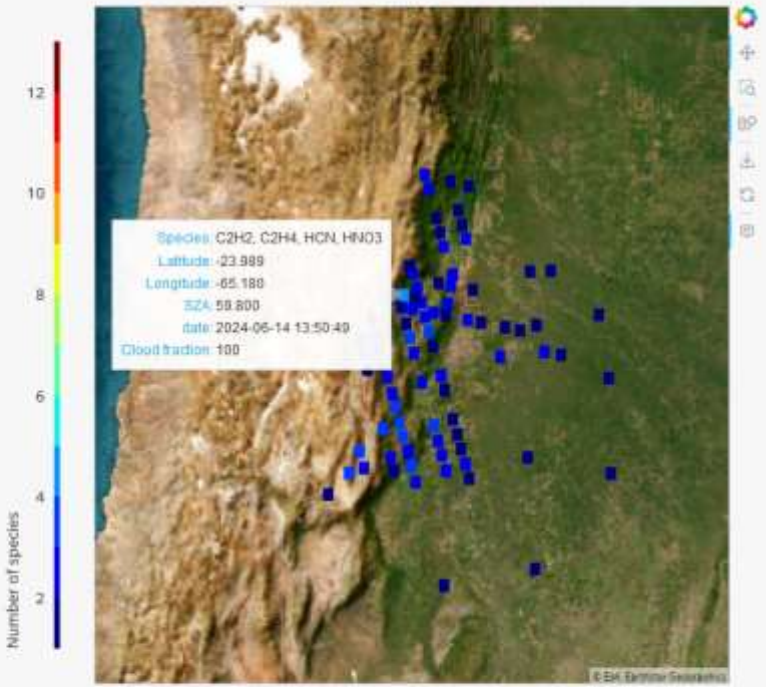


Filter data

Day | Both | Night

Species gases | Parameters

- [x] C2H2
- [x] C2H4
- [ ] C2H4O
- [ ] CH3OH
- [x] HCN
- [ ] HCOOH
- [x] HNO3
- [ ] CO
- [ ] CO2
- [ ] NH3
- [ ] SO2



Map of the number of species detected by IFOV

|Description of product L2 - number of detection|

Summary and perspectives

- **A processing for automatic detection and interpretation of anomalous events from hyperspectral sounding measurements, based on statistical PCA analysis of the L1 spectra, has been implemented and tested on IASI.**
- **Extensive analyses demonstrate the potential of this approach for the detection, classification and monitoring of fire events, volcanoes, and other unexpected events :**
  - ✓ In NRT
  - ✓ In both clear and cloudy/aerosol conditions
  - ✓ Promising potential for the detection of rare events or new spectral signatures, allowing to explore large amount of data for screening useful information related to extreme events
- **With CNES support, the operational processing of IASI-C data is in validation phase at AERIS data center, and will provide systematic, NRT global detection results by 2025 onward. Application to IASI-NG is planned.**

Next steps :

Use 9 months of 2024 global processing for tests and improvements : Verify and evaluate detection and identification of species, identify and solve for false detections, improve dedicated processing for night/day, cloudy/clear FOVs

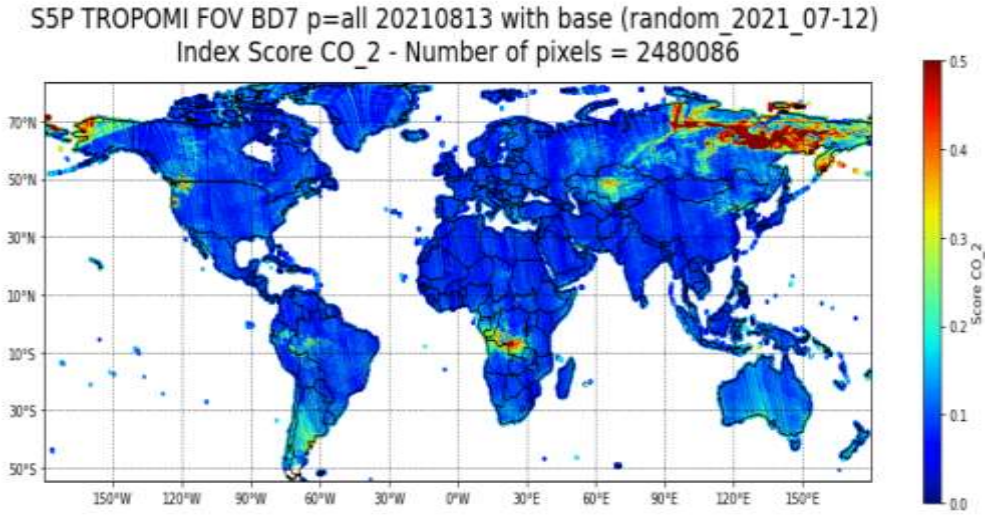
Start Exploiting/exploring systematic processing of IASI-C data with scientists and users : Monitor atmospheric events based on already defined indicators, analyse and explore based on full residual

Perspectives : prepare for the processing of new coming atmospheric sounders : IASI-NG, S5/UVNS, IRS, S4/UVN, CO2M, FORUM

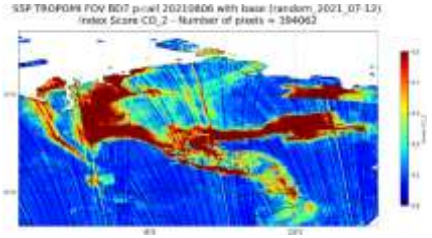
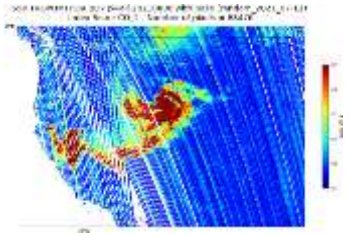
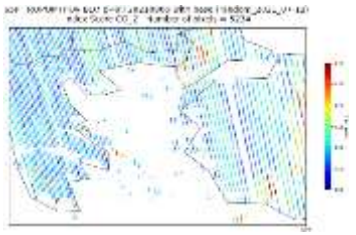
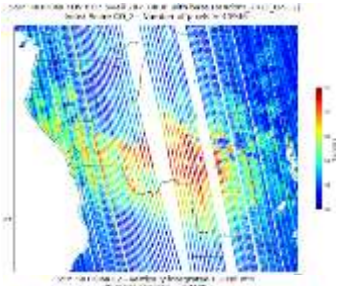
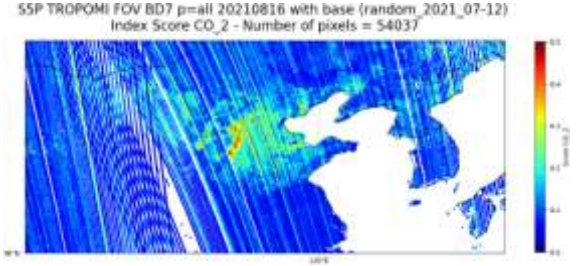
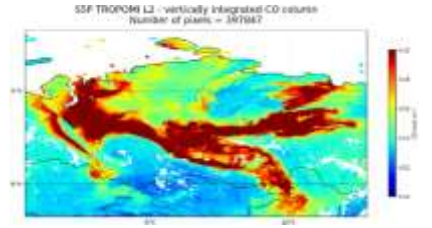
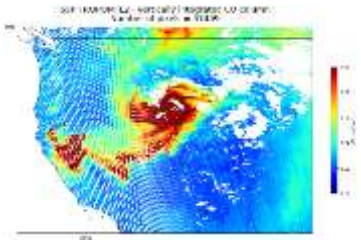
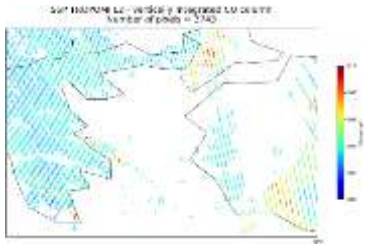
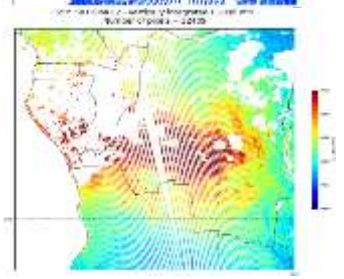
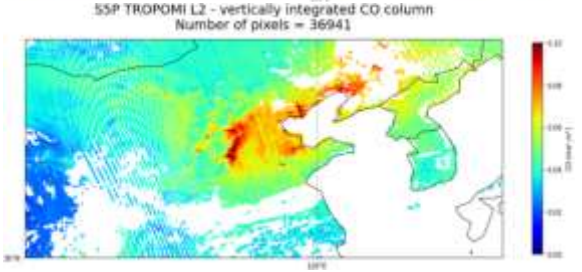
- **S5P-PCA tests on TROPOMI-SWIR data.** Promising complementarity of event detection and characterisation by IASI and S5P, that demonstrates strong potential for wildfire monitoring and analysis, and probably for other atmospheric events (e.g., industry/city pollution)
- **Evaluation and intercomparison of climate models and variability** based on hyperspectral sounders : Tests and analysis of first ranks of IASI-PCA

PCA detection implemented and tested for TROPOMI SWIR spectra (Bands 7 and 8)

Scores map of S5P PCA product
to detect and analyse events
(CO indicators)

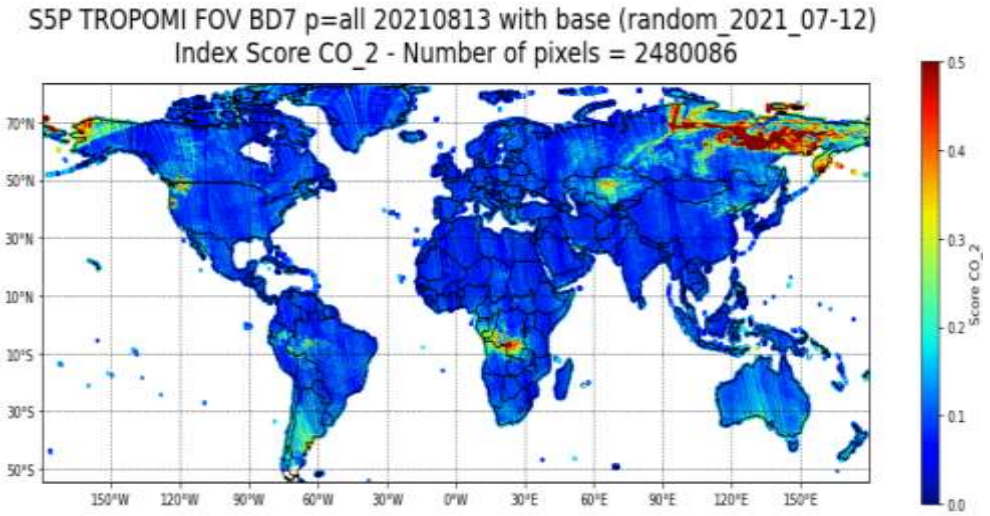


August 2021 : NRT detection of fires and pollution, and validation against operational products

	Siberia fires	*California fires*	*Greece/Turkey fires*	*Africa fires*	*China: Bejin pollution*
S5P PCA CO_2 score maps					
CO S5P L2 maps					

PCA detection implemented and tested for TROPOMI SWIR spectra (Bands 7 and 8)

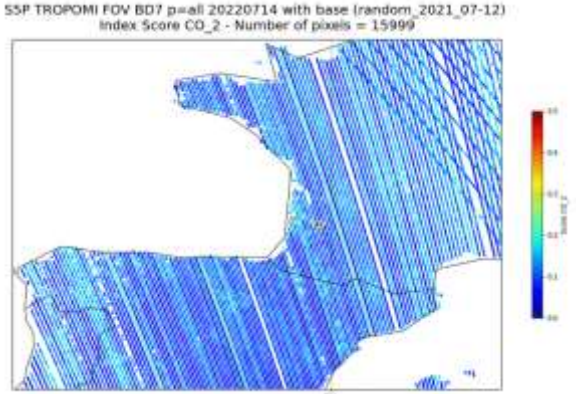
Scores map of S5P PCA product to detect and analyse events (CO indicators)



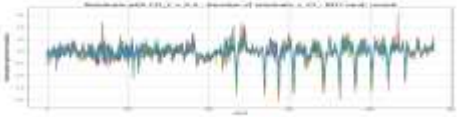
August 2022, South of France fires

Potentiel of combining S5P-PCA and IASI-PCA for fire detections and analysis :

S5P–PCA, Day only : CO_2 outliers
CO signature



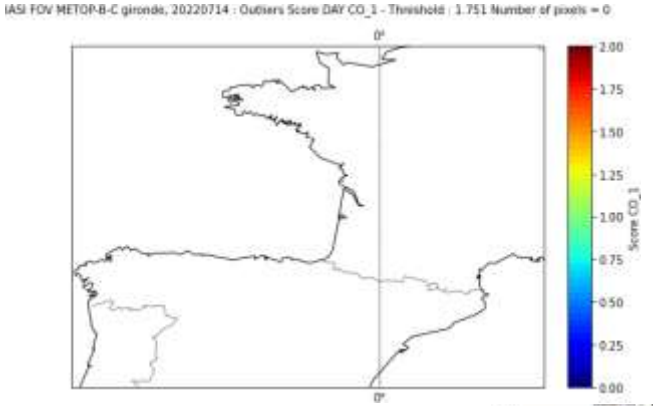
Typical spectral residual in the SWIR (BD7)



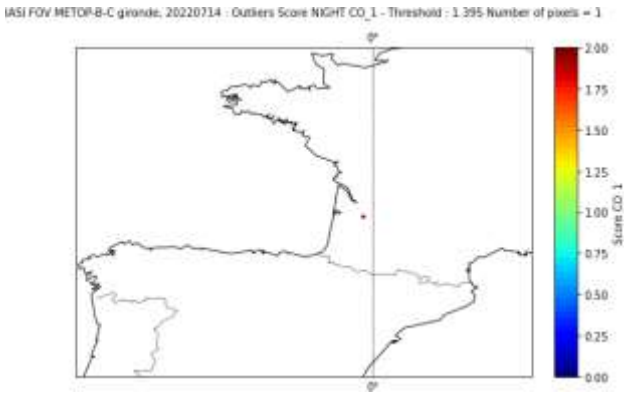
2.30 - 2.34 µm

IASI–PCA (on Metop B and C) CO_1 outliers CO, CO_2, NH_3 signatures

IASI PCA (Day)



IASI-PCA (Night)



2140 – 2180 cm^{-1} (4.58 - 4.67 µm)

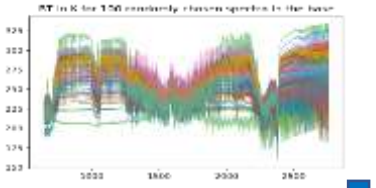
Typical spectral residual in the TIR CO domain

Exploit first PCA for the analysis of the main direction of variability of IASI spectra, for evaluation and intercomparison of climate models :Work initiated in collaboration with CNRM in the frame of FORUM mission

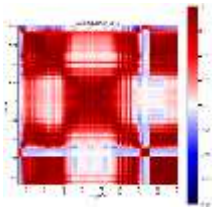
Preliminary result from Lucie Leonarski

Observations

Ensemble d'IFOVs de référence
~120 000 spectres : global / an, plusieurs années



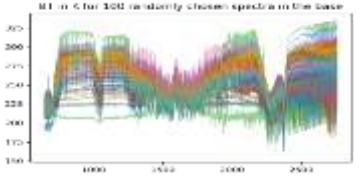
Statistique de la variabilité IASI
Spectre moyen $\bar{y}$ et covariances spectrales



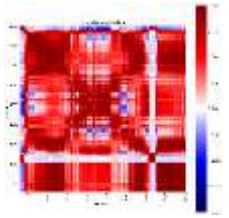
OBS

Models

Simulation ARPEGE-Climat + RTTOV

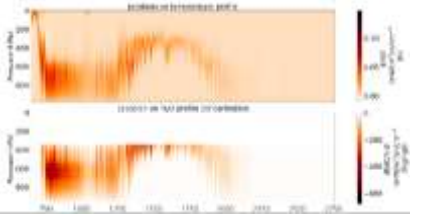


Statistique de la variabilité IASI
Spectre moyen $\bar{y}$ et covariances spectrales

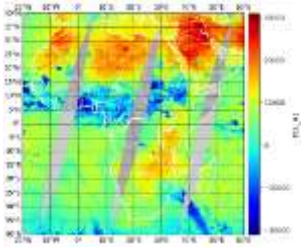


CALC

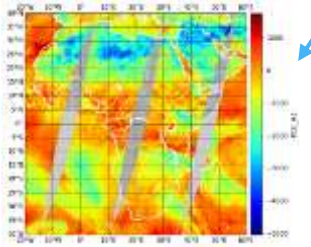
Jacobiens

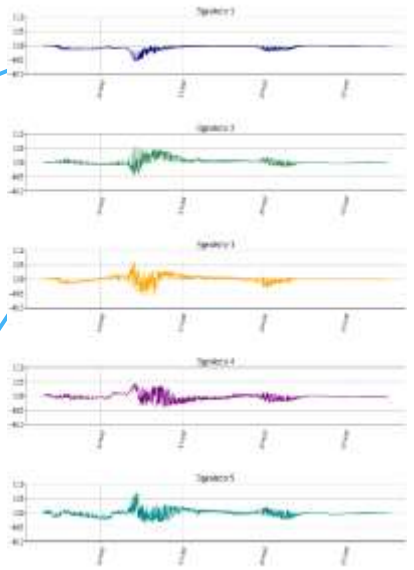


Global daily projection of IASI on PCA 1

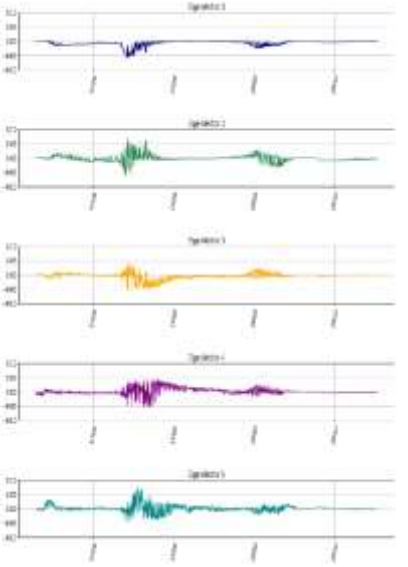


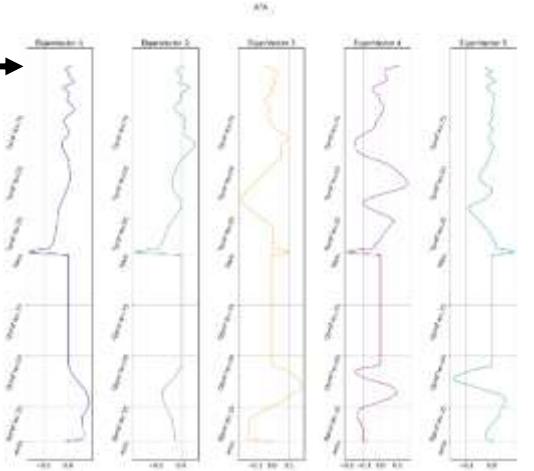
Global daily projection of IASI on PCA 3





Eigenvectors in the spectrum space





Eigenvectors in the geophysical space

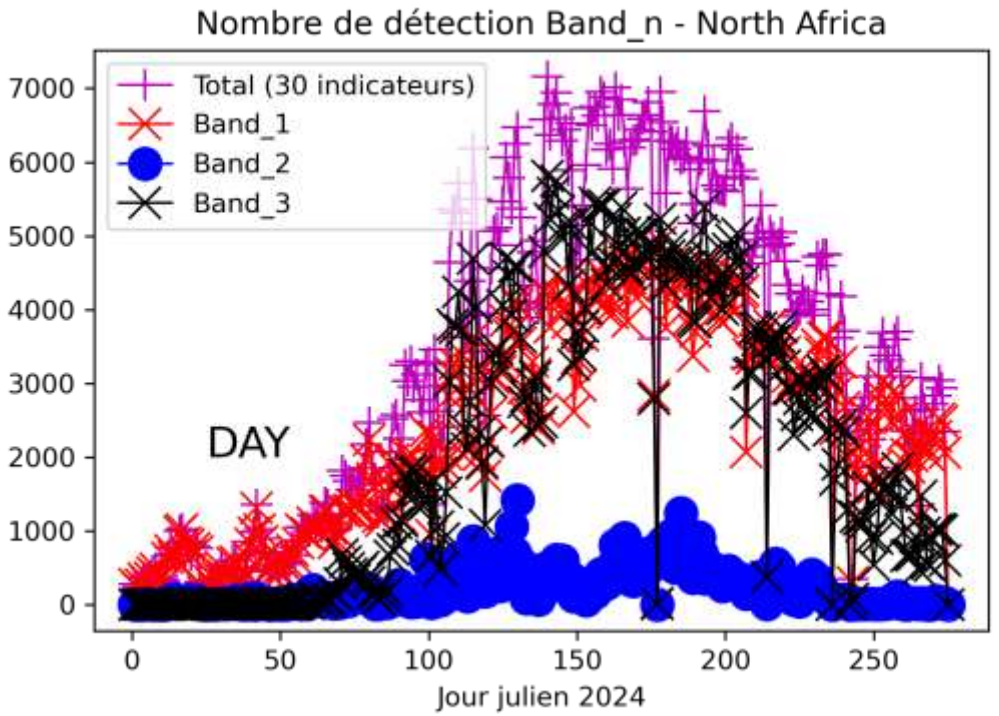
Additional slides

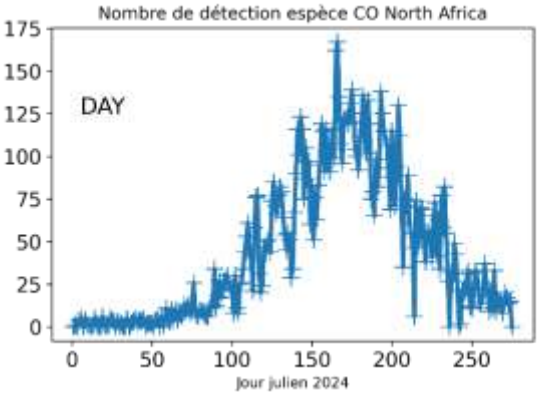
Statistics and monitoring of detections : North Africa (DAY)

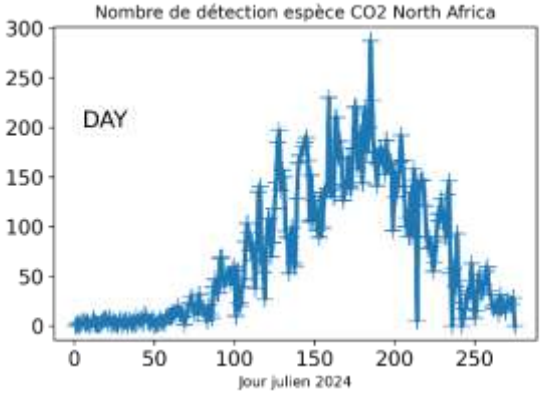
This processing has been evaluated and tested on different datasets for case studies and intercomparisons

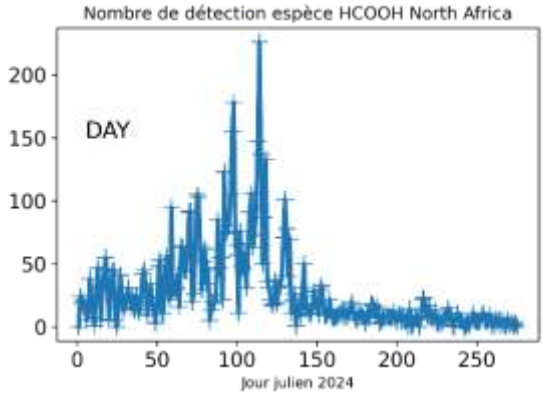
- On going analysis of 9 months of global monitoring of events

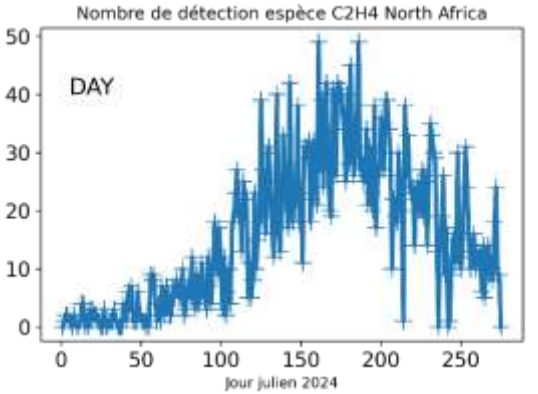
Number of daily detections of events









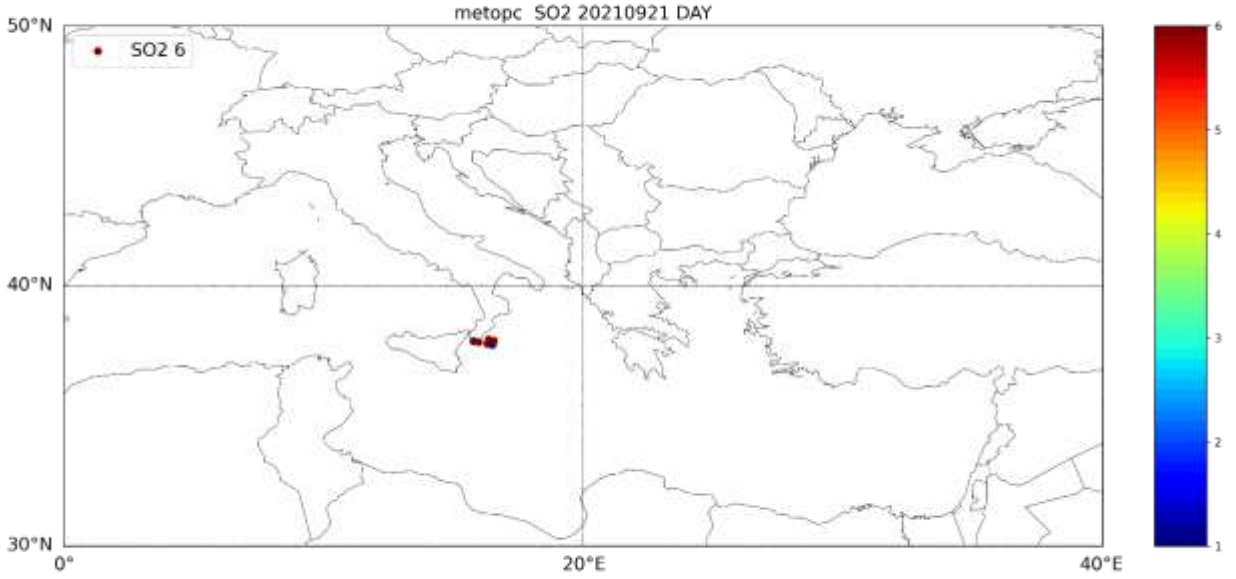


Day and night detection of unexpected events

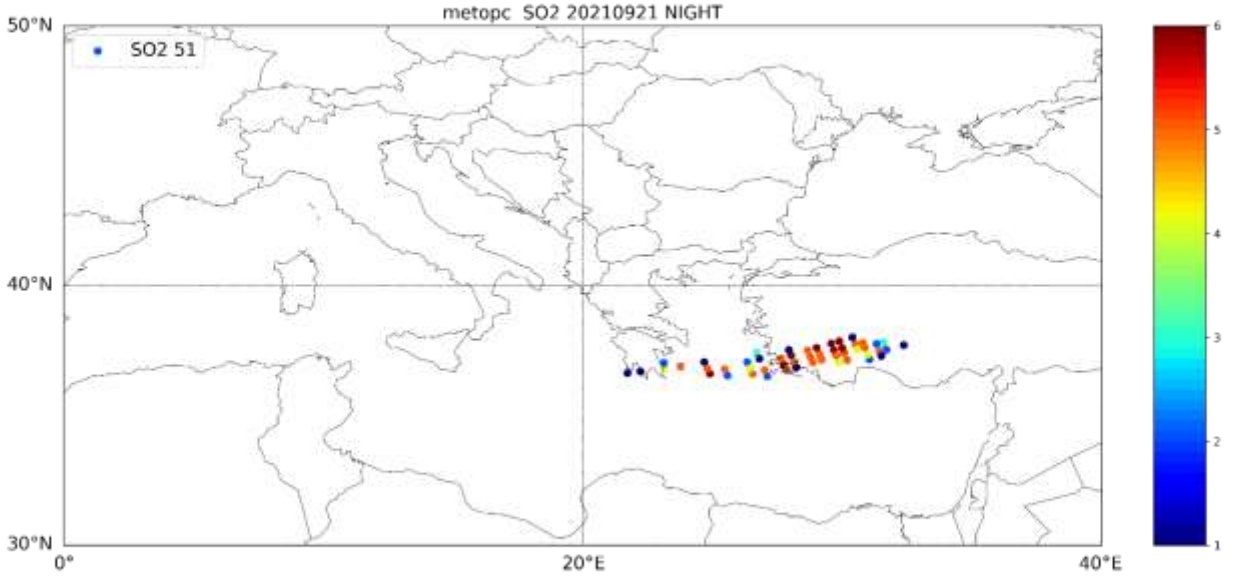
This processing has been evaluated and tested on different datasets for case studies and intercomparisons

- ETNA Volcano eruption, 21/09/2021 : morning and evening maps of the SO_2 plumes

DAY in the morning (6 detected pixels)

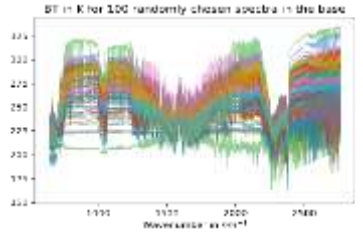


NIGHT in the evening (51 detected pixels)



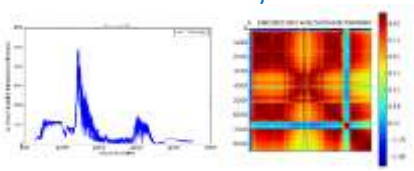
IASI-PCA processing implementation

Reference dataset : ensemble of ~120 000 IFOVs : random selection from global 2013 IASI –A measurements

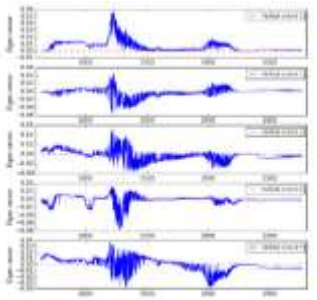


OFFLINE:
Build a truncated Principal Component basis
Truncation : ~150-200

Covariance matrix and averaged spectrum representative of IASI global variability

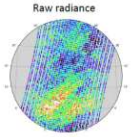


Eigenvectors / Eigenvalues



$E^*, \bar{y}$

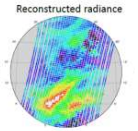
ALL WEATHER input IASI raw spectrum



$$PCC = E^{*T} N^{-1}(y - \bar{y})$$

$$\tilde{y} = \bar{y} + NE^{*}PCC$$

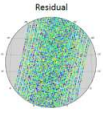
IASI reconstructed spectrum



$$r = N^{-1}(y - \tilde{y})$$

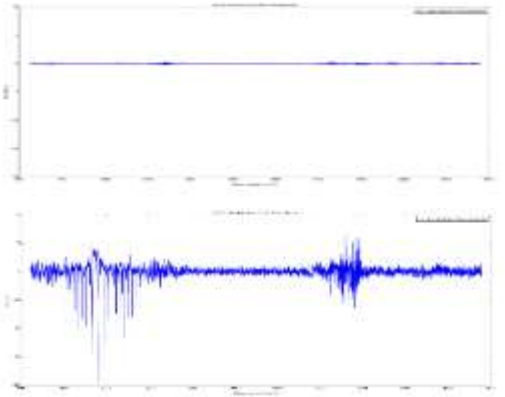
$$RS = \sqrt{\frac{1}{m}\sum_{i=0}^{m} r_i^2}$$

Residual spectrum



~96% nominal reconstructions

~4% Outliers
~50 000 IFOVs / day

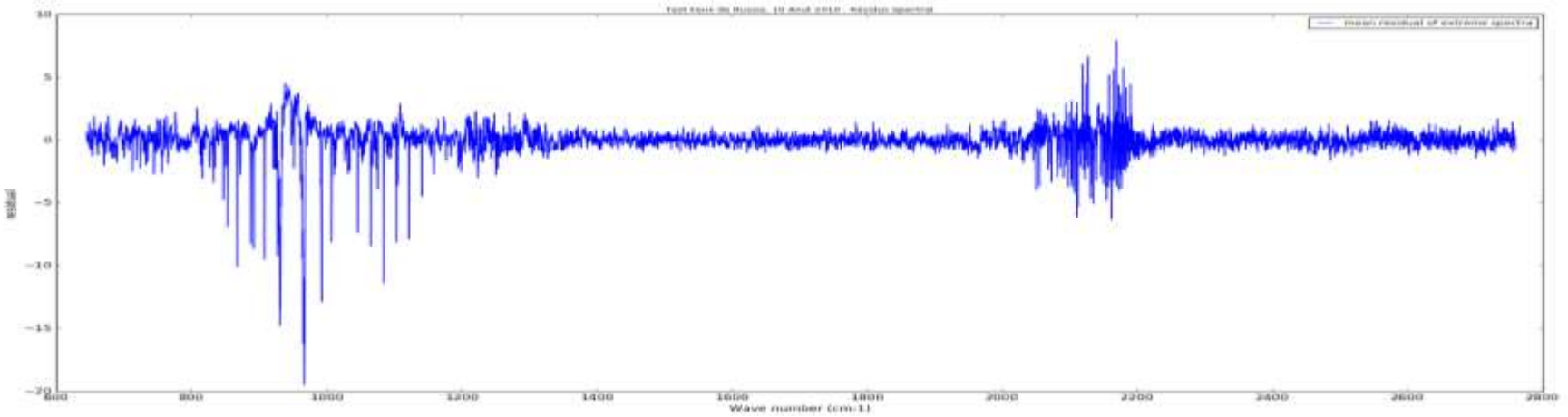


ONLINE :

Project each measured spectrum on the truncated basis, and compute the reconstructed spectrum from the truncated coefficients

Compute the reconstruction residuals (noise normalized)

How to deal with the residuals : reconstruction score and indicators

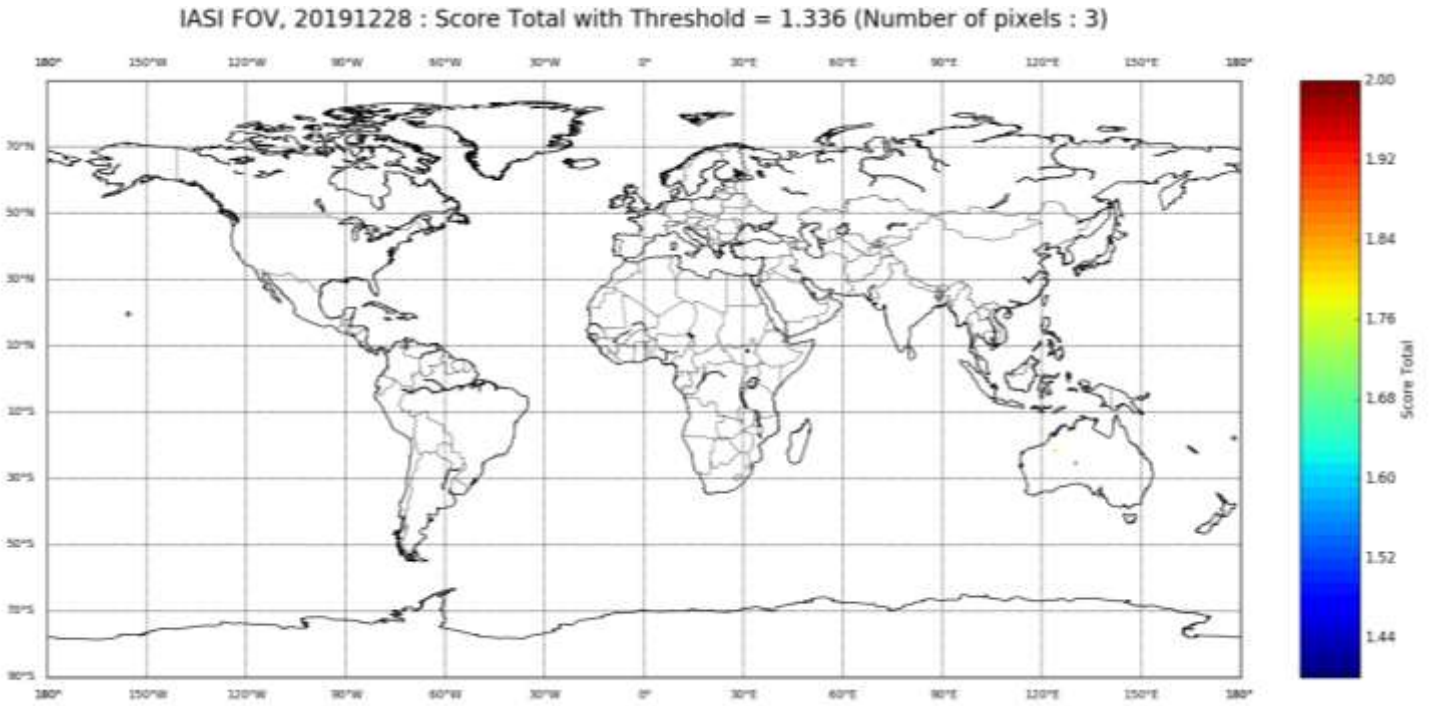


Total reconstruction score :

$$RS = \sqrt{\frac{1}{m}\sum_{i=0}^{m} r_i^2}$$

It allows the detection of outiers : a high score (larger than 1) indicate that the residual spectrum contains signal (and not only noise as in nominal situations).

ILLUSTRATION : Detection of Australian fires and eruption of the Taal volcano



IASI-PCA : a systematic processing of IASI data (A, B and C) is being implemented on AERIS French atmospheric data center : based on indicators

30 indicators has been defined and selected (a spectral interval, and an associated name corresponding to the targeted species) and tested : For each indicator, 2 detection thresholds (night and day) have been optimised, for the automatic detection of outlliers associated to each indicator

```
#date : yyyymmdd Version du 23/05/2023 par C.Camy-Peyret, S.Pipien, P.Prunet
#nombre d'indicateurs (nb_ind) : 30
#nom_ind        wn1        wn2  seuil_jour  seuil_nuit   score   diagn
C2H2_1      728.000    732.000       2.076       2.008    stdv    mean
C2H4_1      940.000    958.000       1.490       1.419    stdv    mean
C2H4O_1     866.500    877.250       1.856       1.665    stdv    mean
CH3OH_1    1030.000   1040.000       1.437       1.380    stdv    mean
CO_1       2153.500   2155.750       1.751       1.395    stdv    mean
CO_2       2157.250   2159.250       1.842       1.538    stdv    mean
CO_3       2178.500   2191.000       1.698       1.439    stdv    mean
CO2_1      2048.250   2059.000       1.883       1.543    stdv    mean
CO2_2      2064.500   2065.500       2.408       2.098    stdv    mean
CO2_3      2077.250   2078.250       2.307       2.069    stdv    mean
HCN_1       711.000    715.000       2.039       1.978    stdv    mean
HCOOH_1    1103.000   1109.000       1.883       1.588    stdv    mean
HNO3_1      878.000    880.000       2.267       2.086    stdv    mean
HNO3_2      895.000    897.000       1.965       1.933    stdv    mean
HNO3_3     1313.000   1332.000       1.457       1.316    stdv    mean
NH3_1       961.000    971.000       2.350       1.640    stdv    mean
NH3_2       925.000    935.000       2.090       1.633    stdv    mean
SO2_1      1138.500   1148.000       1.558       1.507    stdv    mean
SO2_2      1320.000   1324.000       1.471       1.323    stdv    mean
SO2_3      1327.000   1338.000       1.471       1.323    stdv    mean
SO2_4      1371.000   1371.750       2.224       2.120    stdv    mean
SO2_5      1376.000   1376.750       2.102       2.074    stdv    mean
SO2_6      1344.750   1345.250       2.313       2.189    stdv    mean
Surf_1      833.250    834.000       2.396       2.309    mean    stdv
Surf_2      861.500    863.500       1.624       1.461    mean    stdv
Surf_3     1234.250   1235.250       1.790       1.602    mean    stdv
Surf_4     2132.500   2134.000       1.872       1.525    mean    stdv
Total_1     646.000   1149.500       1.468       1.293    stdv    mean
Total_2    1225.000   1934.500       1.205       1.149    stdv    mean
Total_3    2025.000   2759.000       1.633       1.261    stdv    mean
```

Each day, global detection results are automatically generated in a log file : *log_event_yyyymmdd.txt*

- Each line of the file is associated to one IFOV outlier (i.e., when at leat 1 indicator passes over the threshold)
- For each outlier IFOV, the line contains the score value of all the indicators, as well as **a "*code_detection*" binary vector providing detection status** (1 for detection, or 0 for non detection) of each indicator

ILLUSTRATION :
Maps of 4 indicator scores (when higher than the threshold) produced from the log file (Australian fires)