Principal Component Analysis of hyperspectral sounding measurements for the detection of extreme atmospheric events

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

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

CNES contracts : n°201128/00, EU Grant : INEA/CEF/ICT/A2018/1810854

Objectives

The European operational observation of atmospheric composition from space currently benefits from data provided by IASI and GOME onboard Metop, as well as Sentinel 5 precursor (S5P/TROPOMI), and will be extended and consolidated soon by Copernicus Sentinel 4 (MTG IRS and UVN), Sentinel 5 (Metop-SG, with IASI-NG and UVNS) and CO2M.

In the next few years, hyperspectral sounding missions will thus deliver each day several Terabytes of raw data cubes representing an unprecedented amount of atmospheric data. A key requirement is to develop capacities of intelligent screening of large amounts of satellite data for targeting scenes of interest such as extreme atmospheric events, in view of their specific processing or exploitation.

We have implemented a specific approach developed for hyperspectral atmospheric sounding measurements, i.e. the Principal Component Analysis (PCA) of L1 spectra for detection of extreme events. Based on extensive experience of the PCA method for compression and noise-reduction of IASI L1C data, dedicated algorithms and tools have been developed for the systematic detection of fires, volcanoes, pollution plumes and other events, and implemented for the processing of atmospheric spectra recorded by IASI-A, -B, -C in the Thermal InfraRed (TIR) domain and by TROPOMI, focusing on bands 7 and 8 in the Short-Wavelength InfraRed (SWIR) domain.

Summary

We have tested the feasibility and performances of a near-real-time processing of L1 IASI spectra for the detection, identification and monitoring of atmospheric extreme events. The detected extreme events are defined as data outliers with respect to the representative global variability of IASI spectra (1), and thus the detection is driven by specific metric of the rarity (infrequency) of the spectrum behaviour (2a, 2b). This approach has been also adapted and applied to S5P/TROPOMI measurements (4).

The ability of this processing to deal with both clear and cloudy data has been demonstrated (3a), and multi-years results have been analysed for evaluating performances and limitations of this approach (Vu Van et al., 2023). Our results explore the potential for characterizing and classifying fires and other atmospheric events from multi-species detection (3a) and prepare the systematic and automatic processing of these data (3b) for large scale and long period analysis and statistics (3c).

The potential of combining IASI and S5P/TROPOMI for the detection, characterisation and classification of extreme events such as fires and volcanoes is analysed (4). Applications on the next atmospheric sounding missions on polar and geostationary orbits (IASI-NG, MTG/IRS, Sentinel 4 and 5, CO3M) are promising.

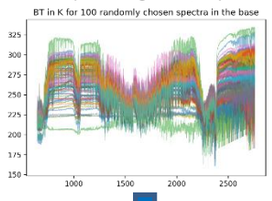
Vu Van, A., Boynard, A., Prunet, P., Jolivet, D., Lezeaux, O., Henry, P., Camy-Peyret, C., Clarisse, L., Franco, B., Coheur, P.-F., and Clerbaux, C., Atmos. Meas. Tech., 16, 2107–2127, https://doi.org/10.5194/amt-16-2107-2023, 2023

1. Principle of the method and processing

1. Build a global training set, i.e. a representative set of normalised IASI L1C spectra,
2. Compute the corresponding covariance matrix ***C*** (dimension of 8461 x 8461),
3. Eigenvector (***E***) decomposition of the covariance matrix ***C***,
4. Choice of the optimum number of PC (truncation by thresholding eigenvalues as a function of the PC rank)
5. Compute the principal components by projecting one spectrum (vector ***y***) on the truncated eigenbasis,
6. Compute the spectral residual ***r*** (difference between original and reconstructed radiances, normalised by IASI instrumental noise ***N***). Compute the reconstruction score ***R*** over ***m*** spectral channels to efficiently identify outliers.

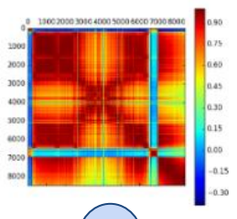
From a reference database of IFOVS

~120 000 spectra : global / 1 year

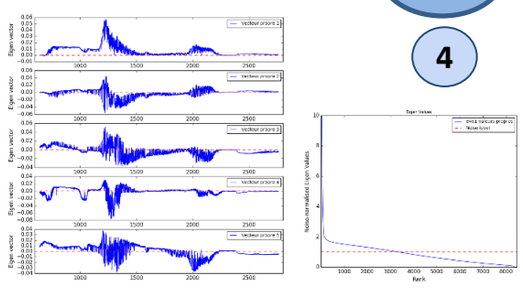


Statistics of IASI global variability …

Mean spectrum $\bar{y}$ and spectral variances/covariances

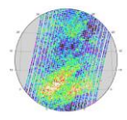


… and eigenvectors database E



E, $\bar{y}$

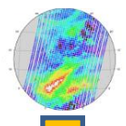
Raw radiance



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

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

Reconstructed radiance

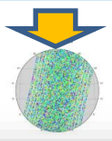


$$r = N^{-1} (y - \tilde{y})$$

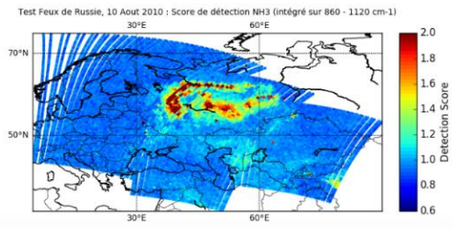
$$R = \sqrt{\frac{1}{m} \sum_{i=0}^{m} r_i^2}$$

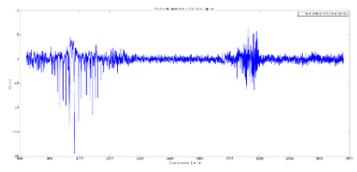
Here R is the global reconstruction score

Residual



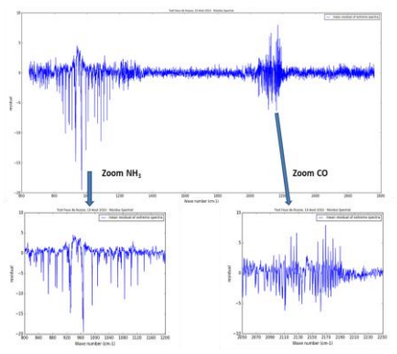
Detect outliers, map extreme events





2a. Indicators

Extreme events are detected using indicators for which reconstruction scores are computed over reduced spectral intervals targeted on species of interest

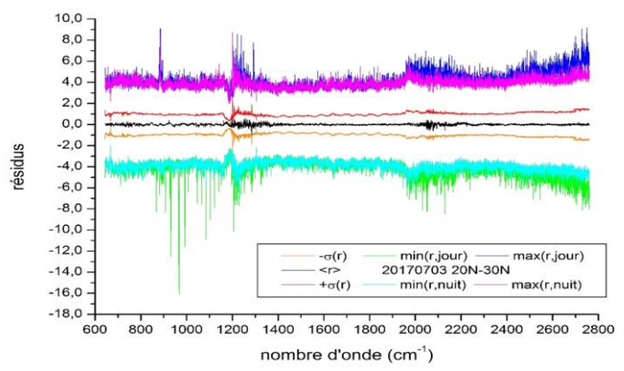


Indicator name	wn_1	wn_2	Threshold (day)
C2H2_1	728.00	732.00	2.035
C2H4_1	940.00	958.00	1.614
C2H4O_1	866.50	877.25	1.778
CH3OH_1	1030.00	1040.00	1.627
CO_1	2153.50	2155.75	2.051
CO_2	2157.25	2159.25	2.083
CO2_1	2048.25	2059.00	1.492
CO2_3	2077.25	2078.25	2.087
HCN_1	711.00	715.00	2.079
HCOOH_1	1103.00	1109.00	1.800
HNO3_1	878.00	880.00	2.325
HNO3_3	1313.00	1332.00	1.332
NH3_1	961.00	971.00	1.666
NH3_2	925.00	935.00	1.688
SO2_1	1138.50	1148.00	1.780
SO2_2	1320.00	1324.00	1.820
SO2_6	1344.75	1345.25	1.968

Illustration and list of the indicators used to detect and characterize extreme events. They have been defined using known absorption features. Detection occurs when the score is above the threshold.

2b. Spectral residual statistics

Extreme events are carefully analysed and interpreted by identifying anomalies (or outliers) in the spectral residual statistics i.e. cases for which the scores (based on the residuals) fall outside a given "confidence interval".



Statistical characteristics of the reconstruction residuals of all the IFOVs for a given period and geographical area. The average (black), the dispersion around the mean (red and orange), the two extremes (max in purple and blue, min in green and light blue).

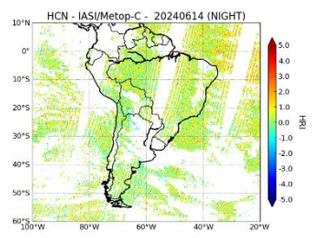
3. Results : scientific analysis and routine detection

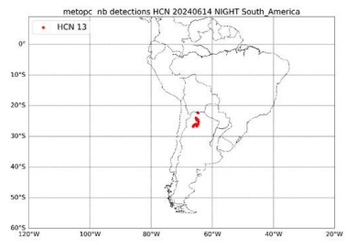
3a. Detection over could, multi-species monitoring

Predefined indicators allow the automatic detection, mapping and characterization of fire plumes, even in the presence of clouds. Example of detection over Argentine, June 2024.

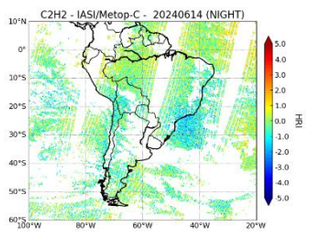
IASI L2 product | ***IASI-PCA outliers***

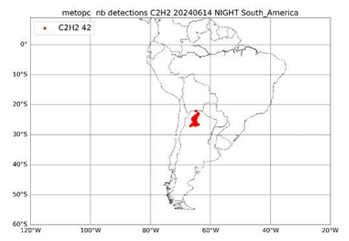
HCN



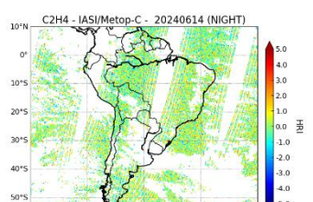


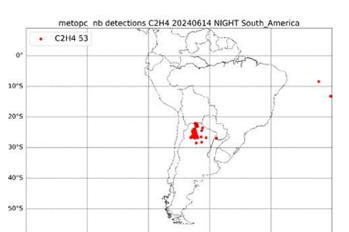
C_2H_2





C_2H_4

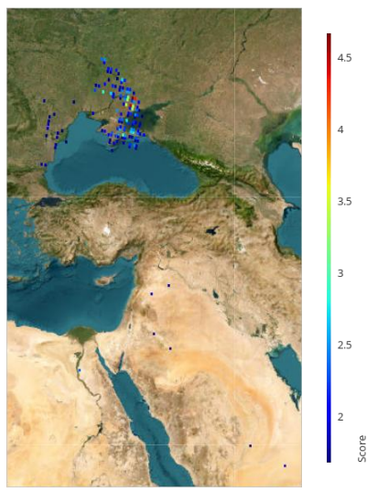




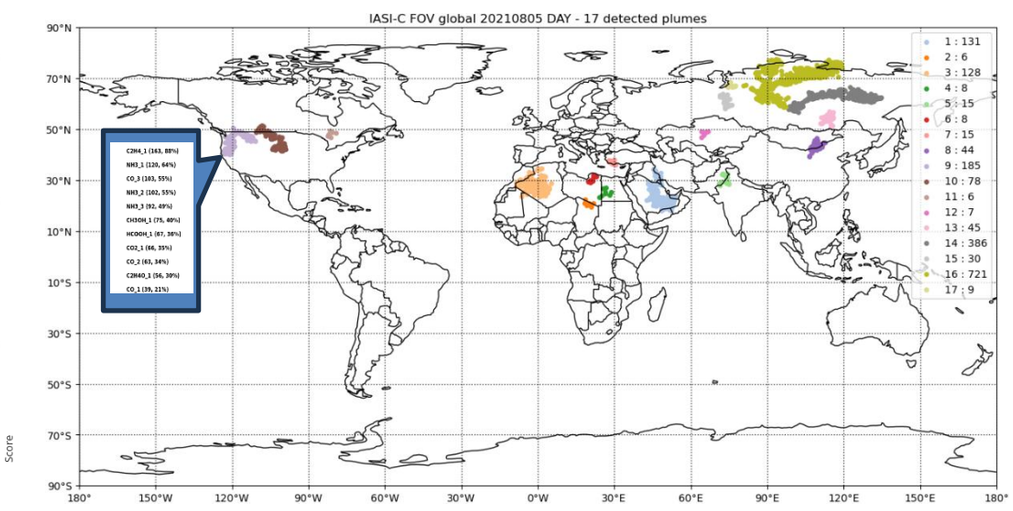
3b. Operation : Toward a NRT processing of IASI for routine monitoring and classification of events

Implementation of near-real-time, routine processing of IASI-C data for providing interpreted products, visualization facilities and open-access database available to the community from AERIS French data centre. In test phase from January 2024, public access current 2025.

Automatic detection and classification of events (“Level 3” products : plumes or isolated events) according their most significant associated indicators, among the categories : “type 1 fires” (intense fires with at least CO_n and CO2_n indicators), “type 2 fires” (low temperature fires with at least C2H2_1, C2H4_1 and HCOOH_1), “fires” (more general category of fires) , “volcanic eruptions” (with at least 3 SO2_n indicators) and “other events”.

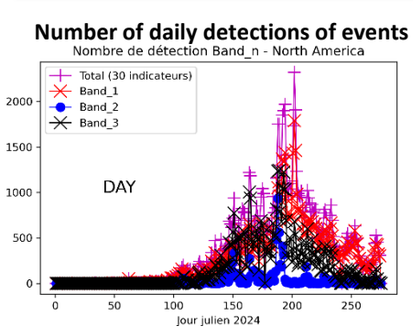


Example of detection and visualisaion of a desert dust plume around Black Sea, April 2024

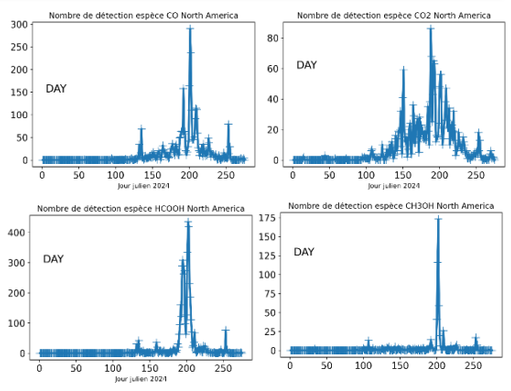


Example of Level 3 product : Identified plumes of events for a given day

3c. Statistics and monitoring of detections : North America



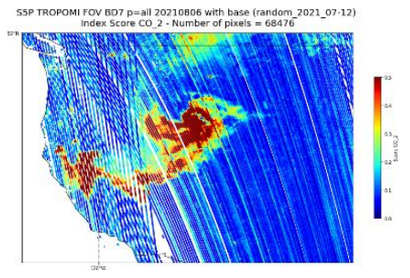
On going analysis of 9 months of global monitoring of events



4. Next steps : synergy multi-instruments (SWIR + TIR in synergy)

Multi-instrument detection and monitoring of California fires (2021) : Combined PCA from TROPOMI (SWIR) and IASI

TROPOMI : *CO detection*



IASI : *C_2H_4 detection*