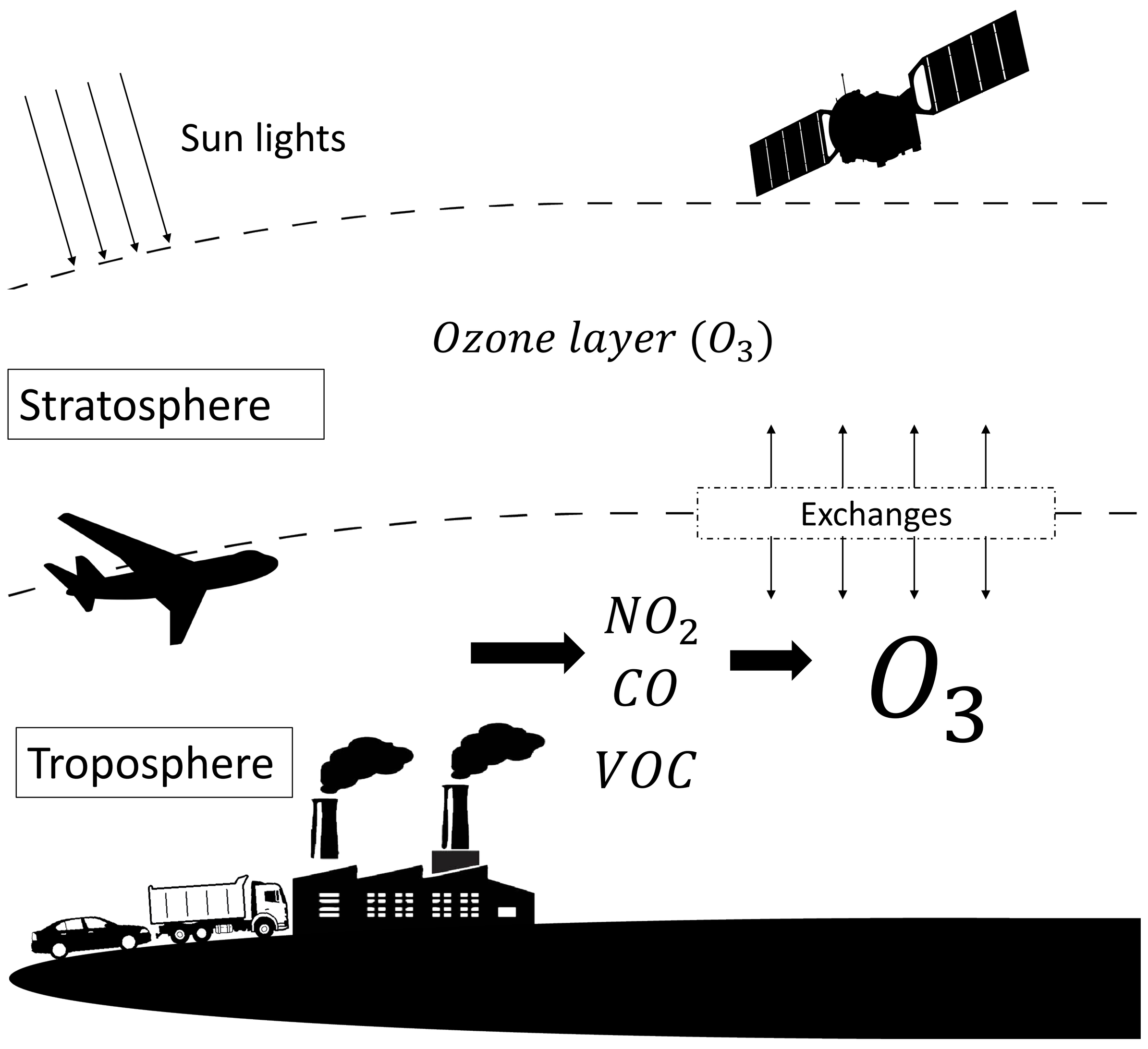
Université de Paris · lisa Laboratoire Inter-Universitaire des Systèmes Atmosphériques · LIPADE Laboratoire d'Informatique PAris DEscartes · diip data intelligence institute of Paris

Smoothing of incomplete air pollution regions of interest from satellite observations

Intern: Camilo ARGÜELLO **Supervised by:** Laurent Wendling, Camille Kurtz, Nicole Vincent (UP), Gaëlle Dufour (LISA)

Problem statement:

- The Ozone (O_3) is well known to be present in different layers at the atmosphere.
- In the Stratosphere, the ozone layer acts as a shield that absorbs most of the sun's ultraviolet radiation [2].
- The O_3 it is present in the troposphere due the human activities, when the emissions interact with the sunlight and other gasses. In lower altitudes the O_3 turns to be a toxic atmospheric pollutant [1,4,5].

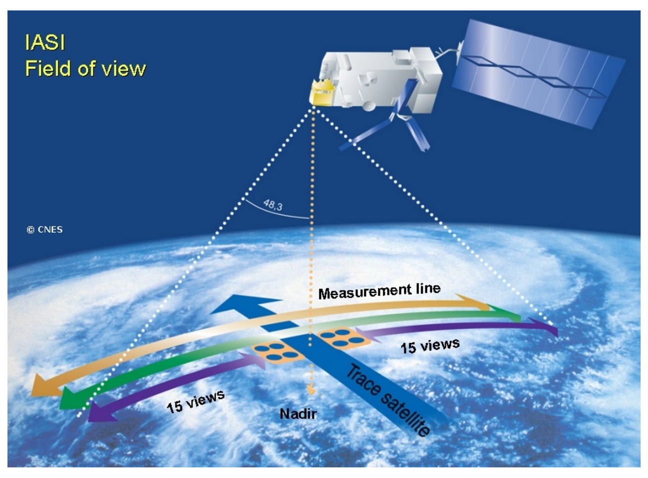


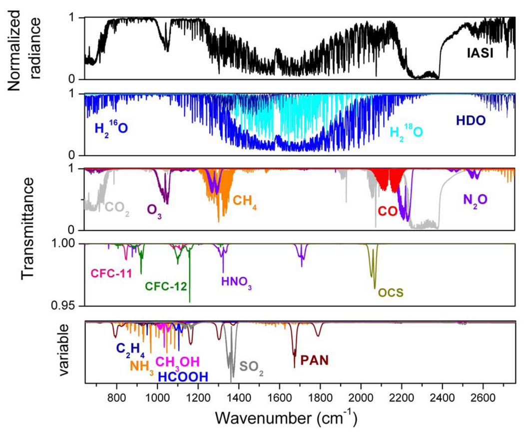
Objectives

To provide ***spatial*** and ***temporal*** descriptions of the ***various pollutants*** from the urban to the continental and the global scales based on ***satellite observations***

- Manage ***missing*** and ***uncertain*** data to ***restore*** plumes
- ***Tracking*** and ***visualization*** over time of the plumes
- Study the interest of adding ***exogen data*** to build plumes

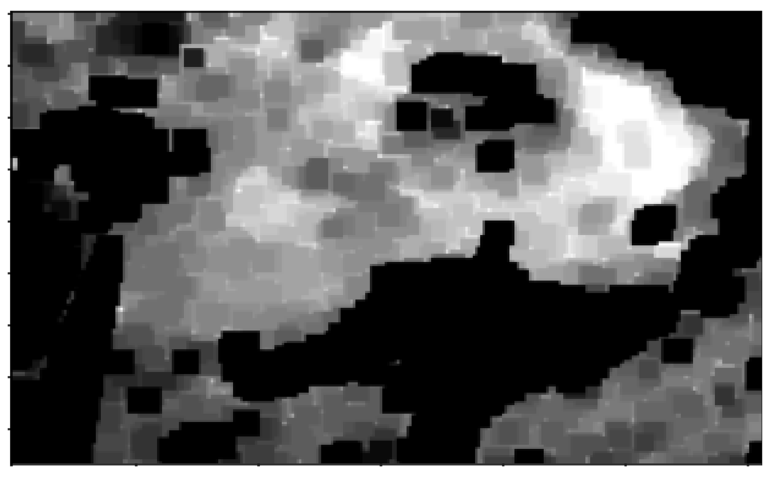
1. Data acquisition





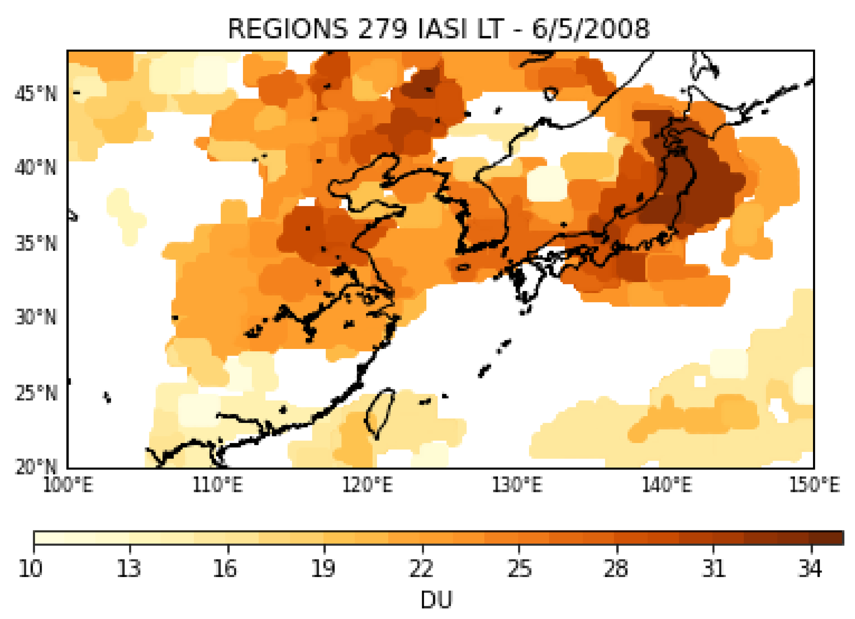
- We were given data from the pollution in Beijing-China for 2 different seasons.
- Each observation is characterized by a resolution (0.25 degree ≈ 25 km), latitude, longitude, pixel values in Dobson Units (DU), quality flag, and an image type (UT/LT).
- We consider the white zones as cloudy areas.

2. Image pre-processing



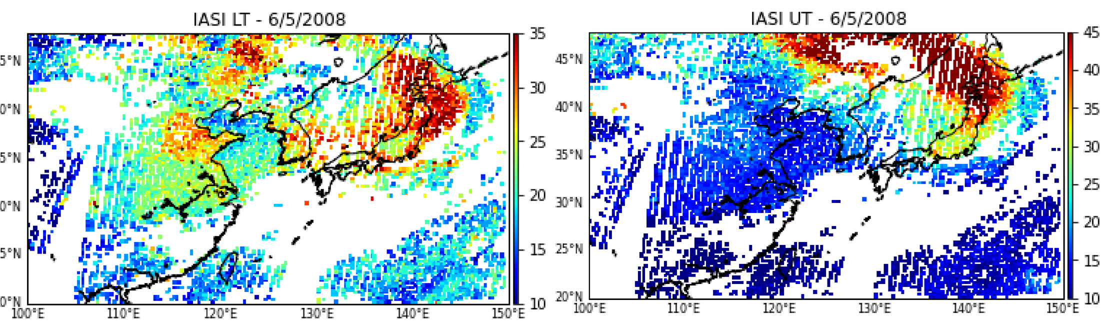
- A set of morphological filters such as a closing, and an image inpainting is applied to fill some small zones missed in the observation.

3. Region extraction

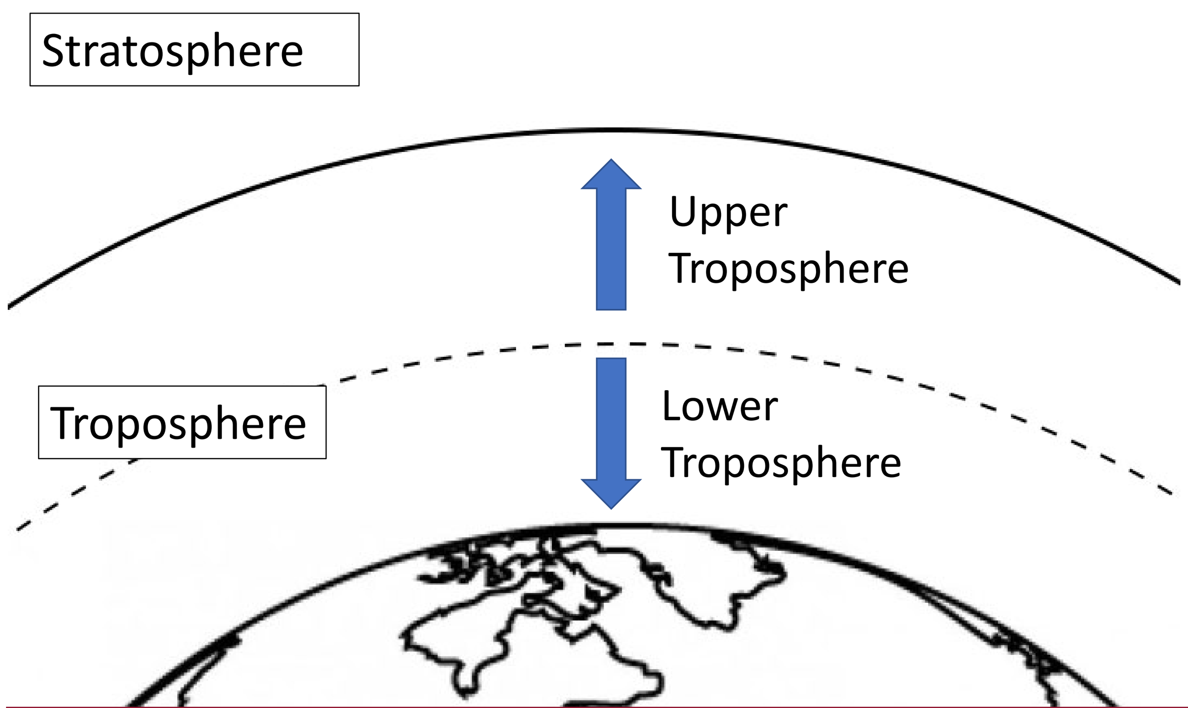


Regions found. The darker colors are the highest values in DU.

- We aim to find a set of Maximally Stable and Extremal Regions (**MSER**) [3].
- This algorithm scans the image in all its possible thresholds to find a set of maximal and stable region.
- A MSER region is composed by a set of points constrained to an adjacency in (4 or 8-neighbourhood).

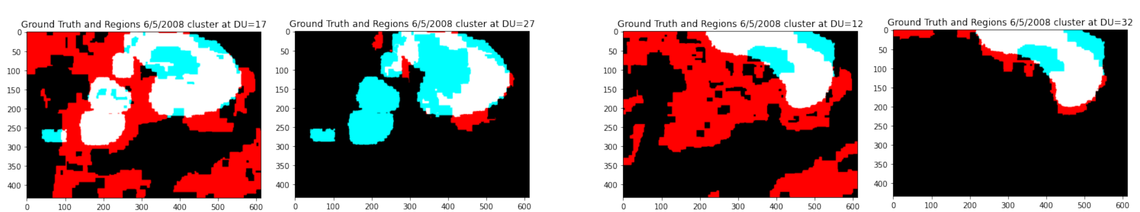


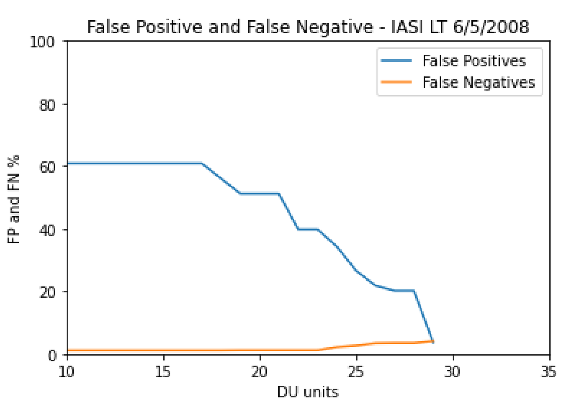
Lower Troposphere (LT) Upper Troposphere (UT)

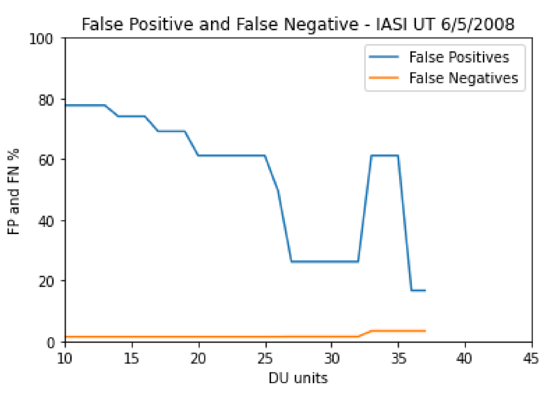


Preliminary Results

The evaluation aims to compare the group of regions found with the ground truth. For the tests, an expert mark out the regions where there exists a higher concentration of O_3 in the LT and UT images. Then, we compare the intersection between each cluster and the ground truth. Our goal is to minimize the number of false positives and false negatives. We make a comparison between both percentages of errors compared with the DU scale.





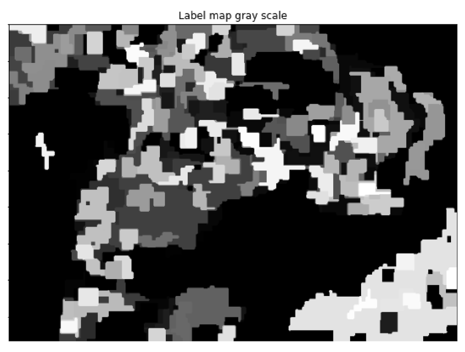


In this case based on the Dobson Units measurements, a progressive limit was defined. For this day, starting from the value 16 over 35 DU, the group of the highest candidate regions to be pollution can be obtained. Where the highest group is located at 29 over 35 DU.

The same test was made for the UT images, where values starting at 25 DU are the candidate regions to be pollutant.

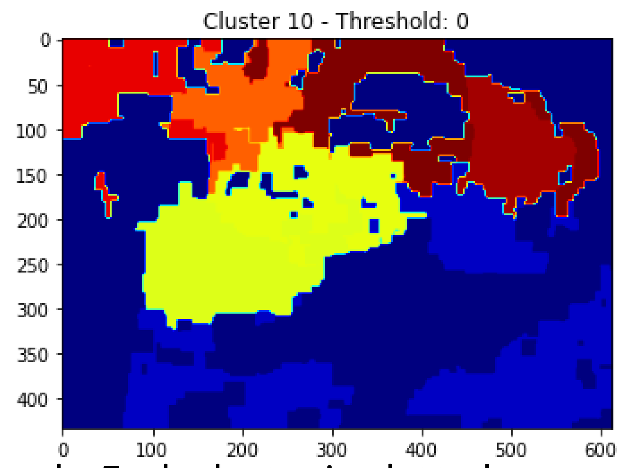
Some regions cause an increase in false positives at very high values.

4. Region projection



- The MSER regions found are projected in a label map to create a partition.
- Each partition corresponds to a region.

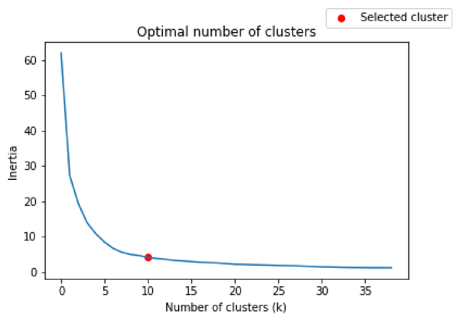
5. Clustering



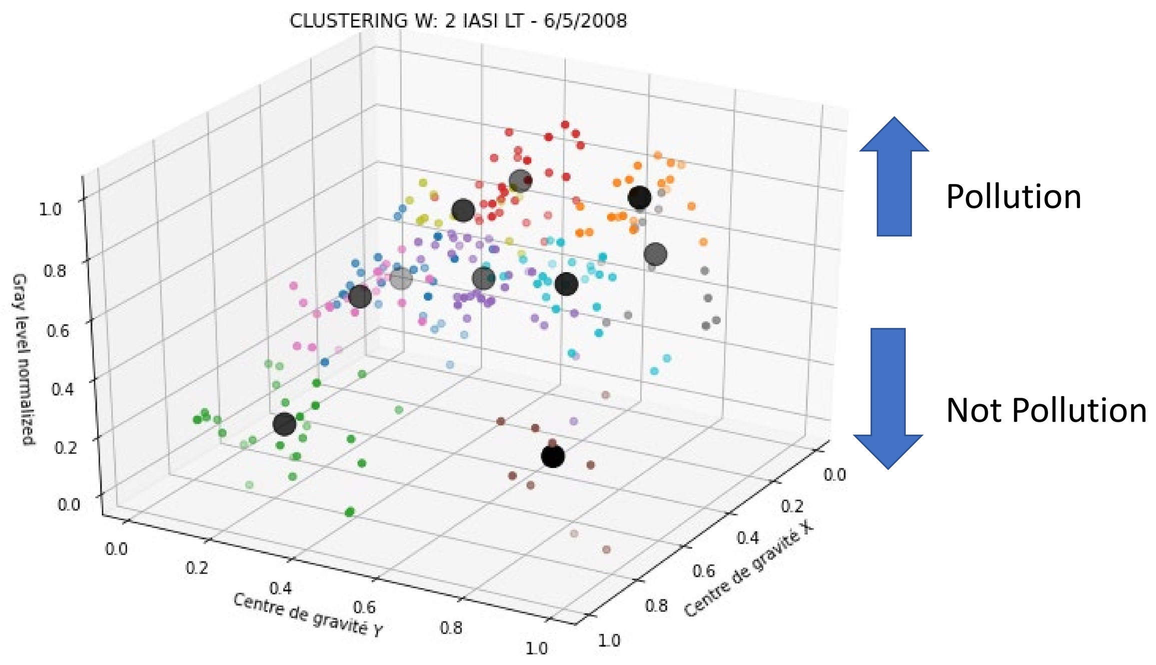
Clustering result. Each cluster is plotted as a unique color region.

- The *kmeans* cluster algorithm is used to determine the candidate regions and groups with pollution.
- The input values are the coordinates of the center of gravity in **X** and **Y** positions of each reconstructed region and the DU value to perform the clustering.

$$X = [x_m, y_m, DU_m]$$



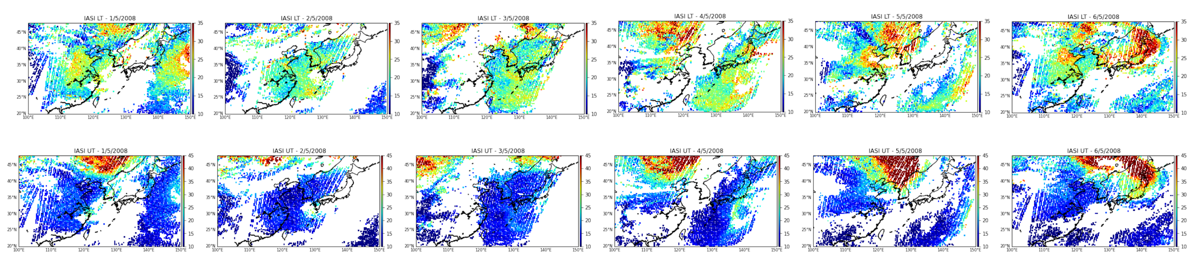
Optimal number of clustering by using the "Elbow" rule.



Clusters found with Kmeans
For this example, k-clusters = 10

Future work

- To track a selected region/group to know the pollution that is moved from one place to another. For example, pollution detected in China and moved out to Korea.



- Classification of pollution based in their type. Natural or anthropogenic.
- Manage ***missing*** and ***uncertain*** data to ***restore*** plumes.
- To compare with other seasons to verify whether we are detecting only pollution areas.
- To add new information for example of CO, to find a correlation between the regions found and another gas observations.

References

1. Doche C., G. Dufour, G. Foret, M. Eremenko, J. Cuesta, M. Beekmann, P. Kalabokas, Summertime tropospheric ozone variability over the Meditterranean basin observed with IASI, Atmos. Chem. Phys., 14, 10589-10600, 2014.
2. European Environment Agency (EEA). Protection of the Ozone Layer.
3. J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," in Proc. BMVC, 2002.
4. Dufour, G., M. Eremenko, J. Cuesta, C. Doche, G. Foret, M. Beekmann, A. Cheiney, Y. Wang, Z. Cai, et al., Springtime daily variations in lower-tropospheric ozone over east Asia : the role of cyclonic activity and pollution as observed from space with IASI, Atmos. Chem. Phys., 15, 10839-1085, 2015.
5. Dufour, G., Eremenko, M., Beekmann, M., Cuesta, J., Foret, G., Fortems-Cheiney, A., Lachâtre, M., Lin, et al. : Lower tropospheric ozone over the North China Plain : variability and trends revealed by IASI satellite observations for 2008–2016, Atmos. Chem. Phys., 18, 16439-16459, 2018.
6. Eremenko, M., Dufour, G., Foret, G., Keim, C., Orphal, J., Beekmann, M., Bergametti, G., and Flaud, J.-M. (2008), Tropospheric ozone distributions over Europe during the heat wave in July 2007 observed from infrared nadir spectra recorded by IASI, Geophys. Res. Lett., 35, L18805
7. Carsten Rother, Vladimir Kolmogorov, Andrew Blake : "GrabCut" : interactive foreground extraction using iterated graph cuts. ACM Trans. Graph. 23(3) : 309-314 (2004)
8. Michaël Clément, Camille Kurtz, Laurent Wendling : Learning spatial relations and shapes for structural object description and scene recognition. Pattern Recognit. 84 : 197-210 (2018)
9. Michaël Clément, Adrien Poulenard, Camille Kurtz, Laurent Wendling : Directional Enlacement Histograms for the Description of Complex Spatial Configurations between Objects. IEEE Trans. Pattern Anal. Mach. Intell. 39(12) : 2366-2380 (2017)
10. F. Solera, S. (2015) Calderara, R. Cucchiara. Socially constrained structural learning for groups detection in crowd. IEEE. Pattern Analysis and Machine Learning.

With the financial support of diiP, IdEx Université de Paris, ANR-18-IDEX-0001