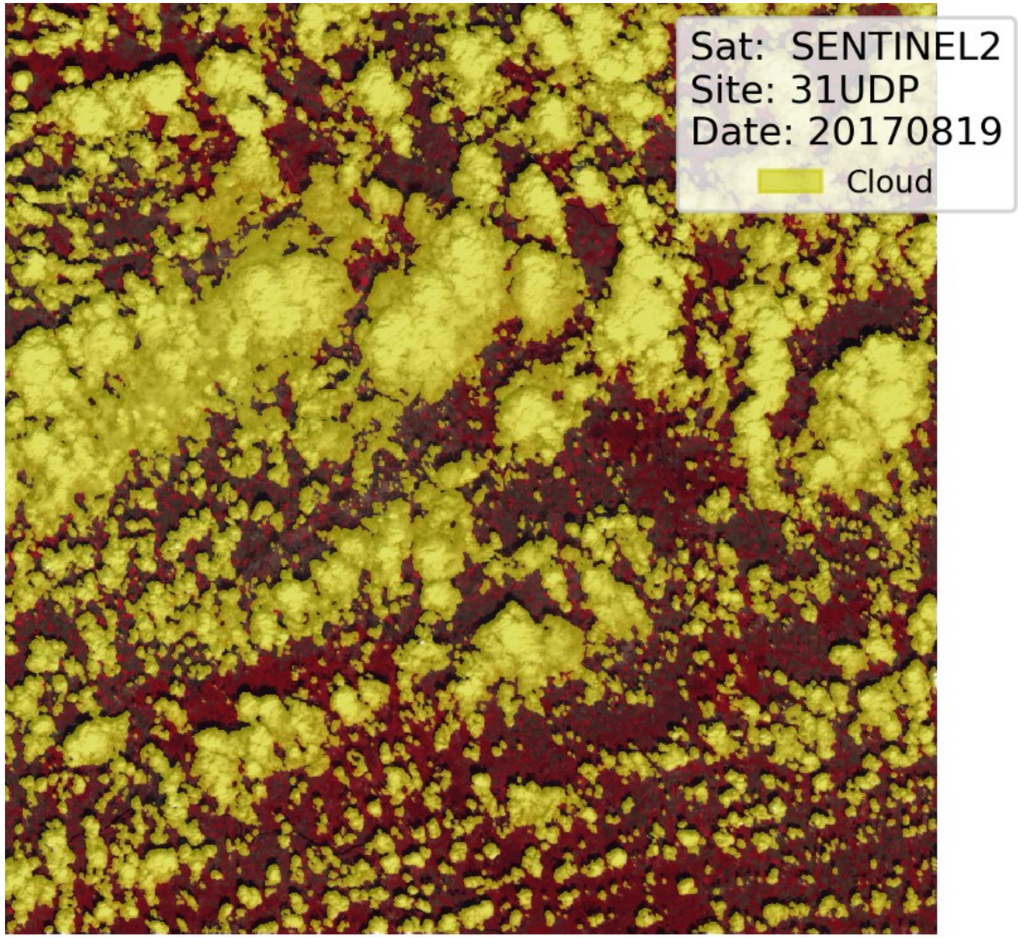
Improving future optical Earth Observation products using transfer learning

Peter Kettig, Eduardo Sanchez-Diaz, Simon Baillarin, Olivier Hagolle, Jean-Marc Delvit, Pierre Lassalle, and Romain Hugues

CNES, CESBIO, IRT – Toulouse, France

Table of contents



Context

- CNES has access to a multitude of optical image products
  - Sentinel2
  - Venus
  - Pleiades
  - …
- Database is only partially labeled today

→ Cloud detection first step for labeling process

- Cloud screening modules specifically designed for each sensor:
  - Maja
  - SVM
  - …

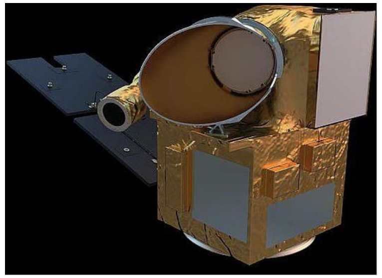
Source: CESBIO

Context

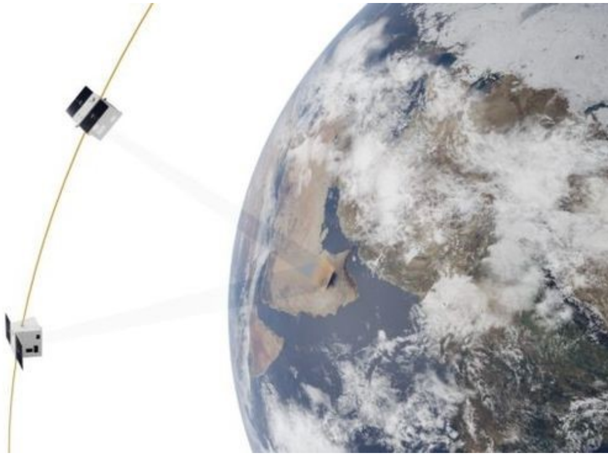
Database is still growing:

- New platforms to be equipped with cloud detection modules:
  - MicroCarb
  - CO3D
  - …

- Specifications for those still TBD; no simulated products
- Existing modules highly dependent on geometric and radiometric attributes



Microcarb

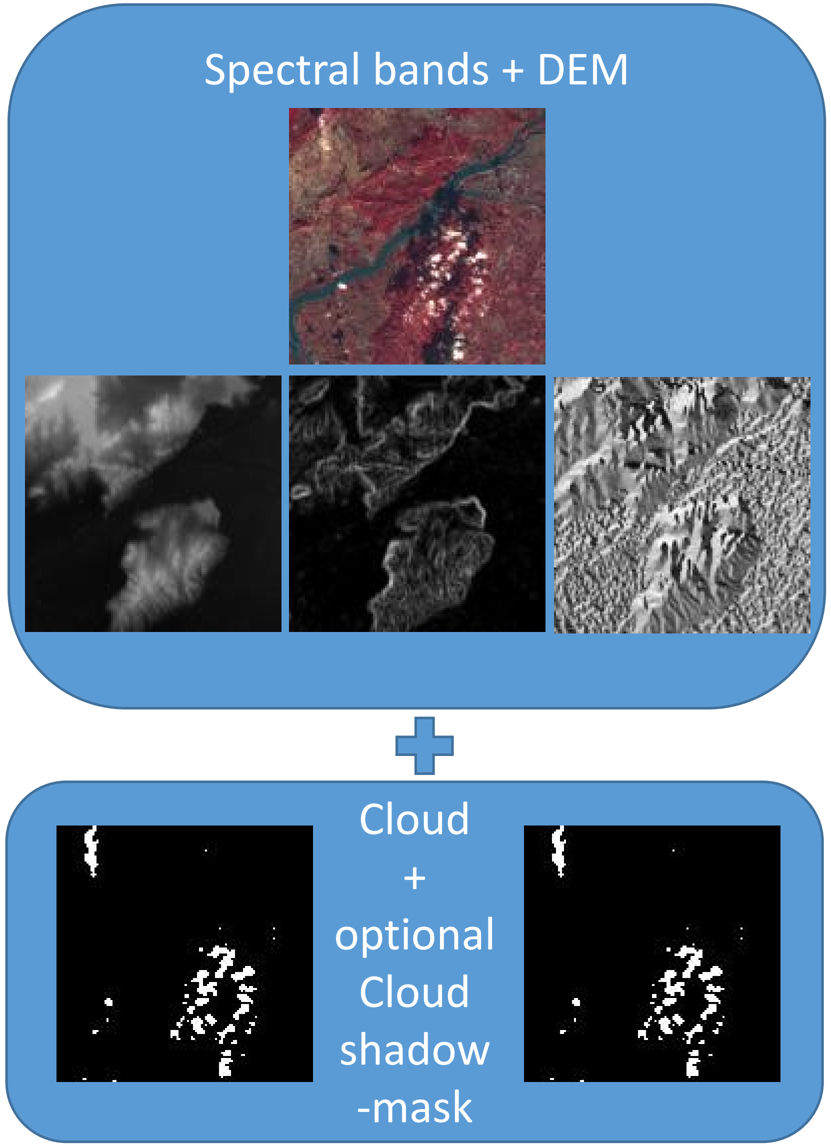


CO3D Constellation

→ Can we use the existing database and processing chains as 'a priori' for future platforms for this labeling task?

Methods used

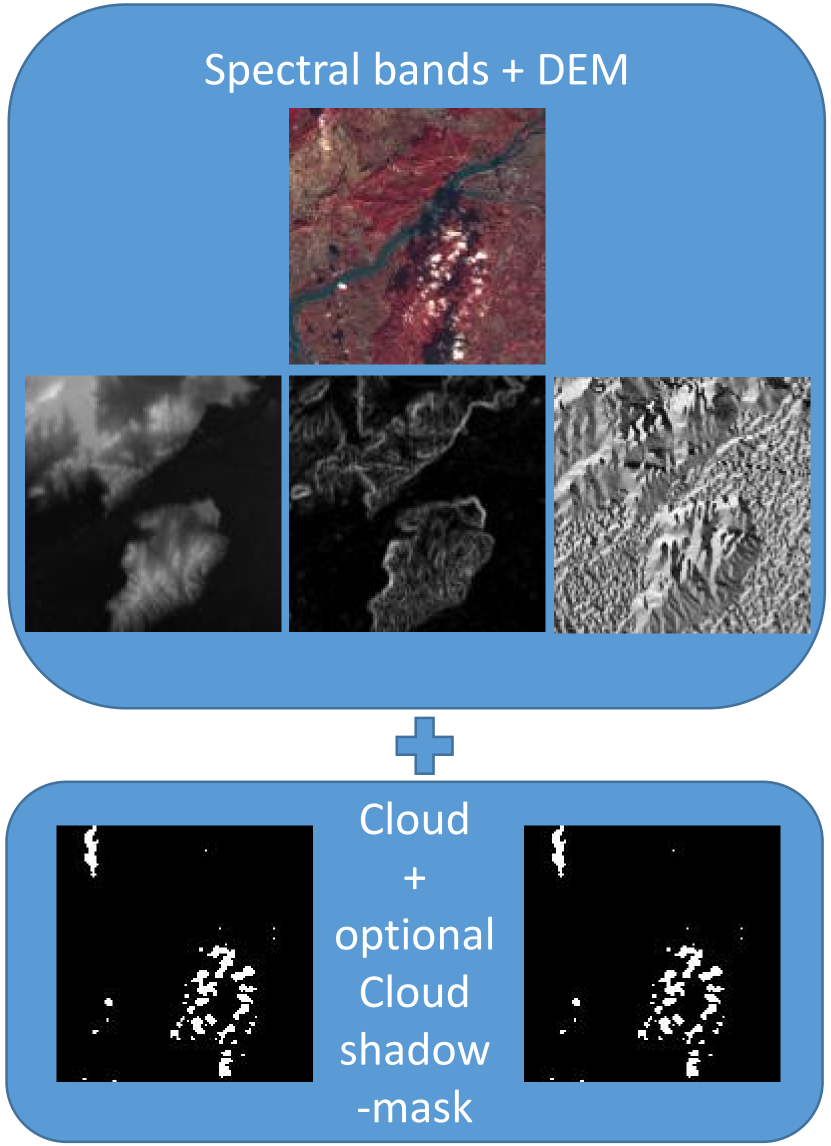
- Training database:
  - Sentinel-2 scenes (tiles) and binary cloud-mask from existing processing chains :
    - Maja
    - Fmask
  - DEM (Terrain model)
- Validation database:
  - Semi-automatically labeled cloud masks (using ALCD)
- Inference: Transfer of Spot-4/5 scenes with DEM and same band configuration



Methods used

- Selection of bands that exist for other platforms (such as SPOT/Pleiades):
  - Red (B4)
  - Green (B3)
  - Blue (B2)
  - NIR (B8)
  - SWIR (B11)
- Using lowered resolution (Original processors also work at reduced resolution)

→ Comparison on equal basis

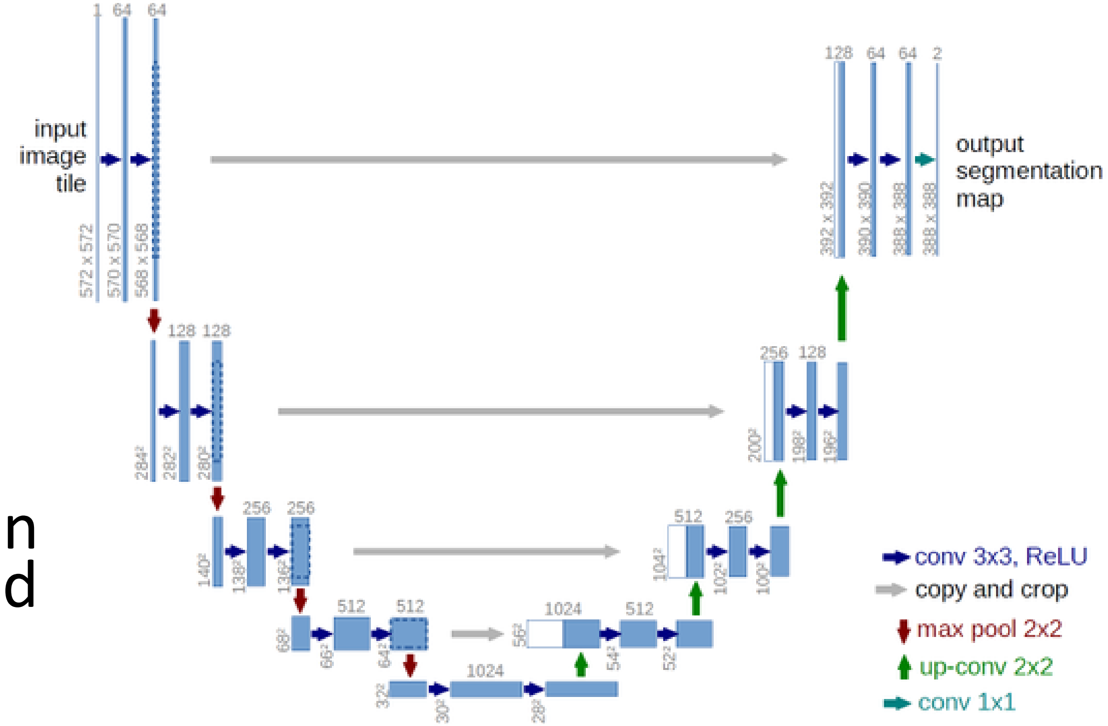


Methods used

Network characteristics:

- Unet with batchnorm and dropout
- Loss: BCE with Dice
  - Highly imbalanced dataset (only 20% of all training pixels considered ‘cloudy’)
- Optimizer: Adam
- Automatic hyperparameter optimisation using [HyperOpt](#) for dropout fraction and filter size

→ Use trained model to infer cloud masks on pre-formatted SPOT and Pleiades imagery



Source

Results

Training: Sentinel-2

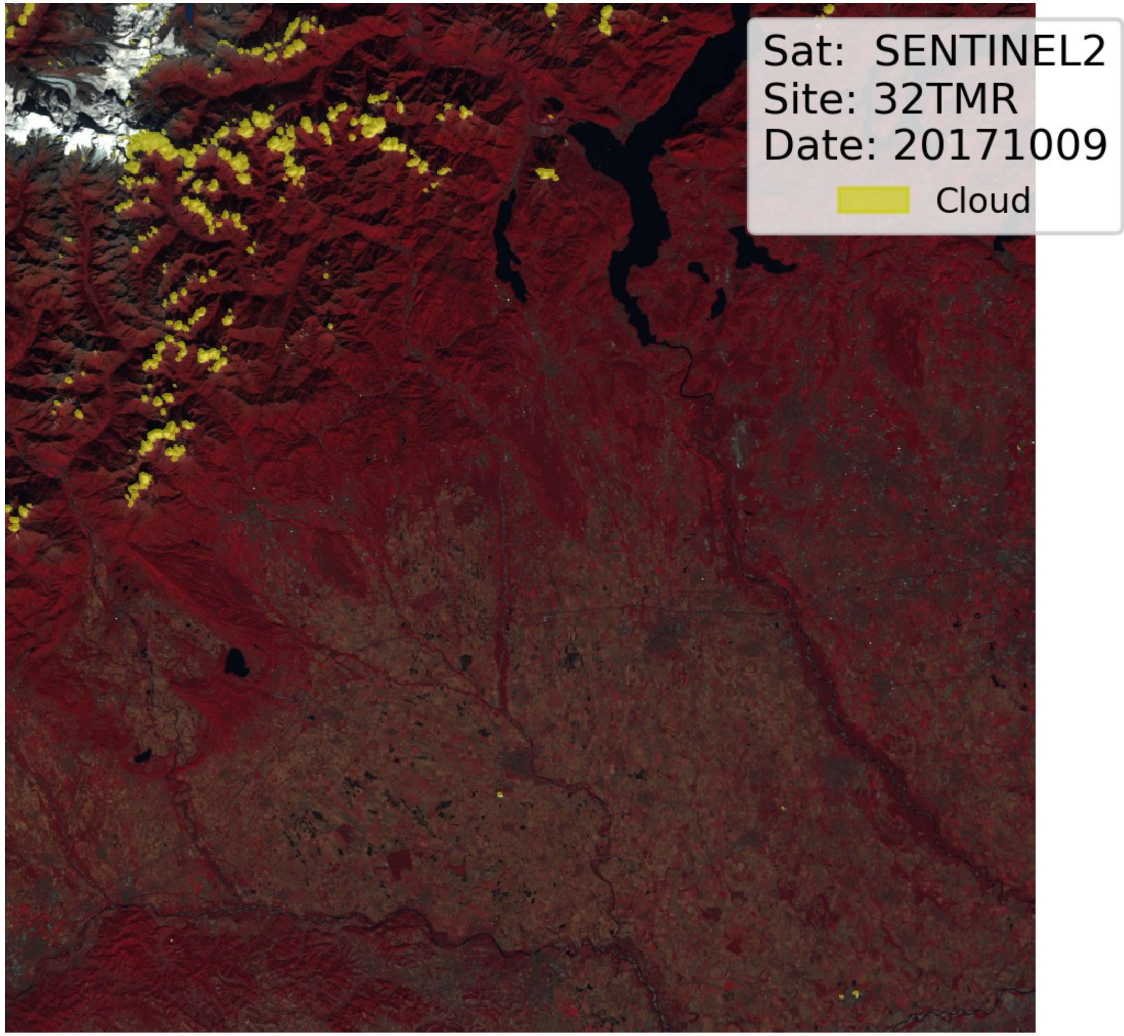
Inference: Sentinel-2

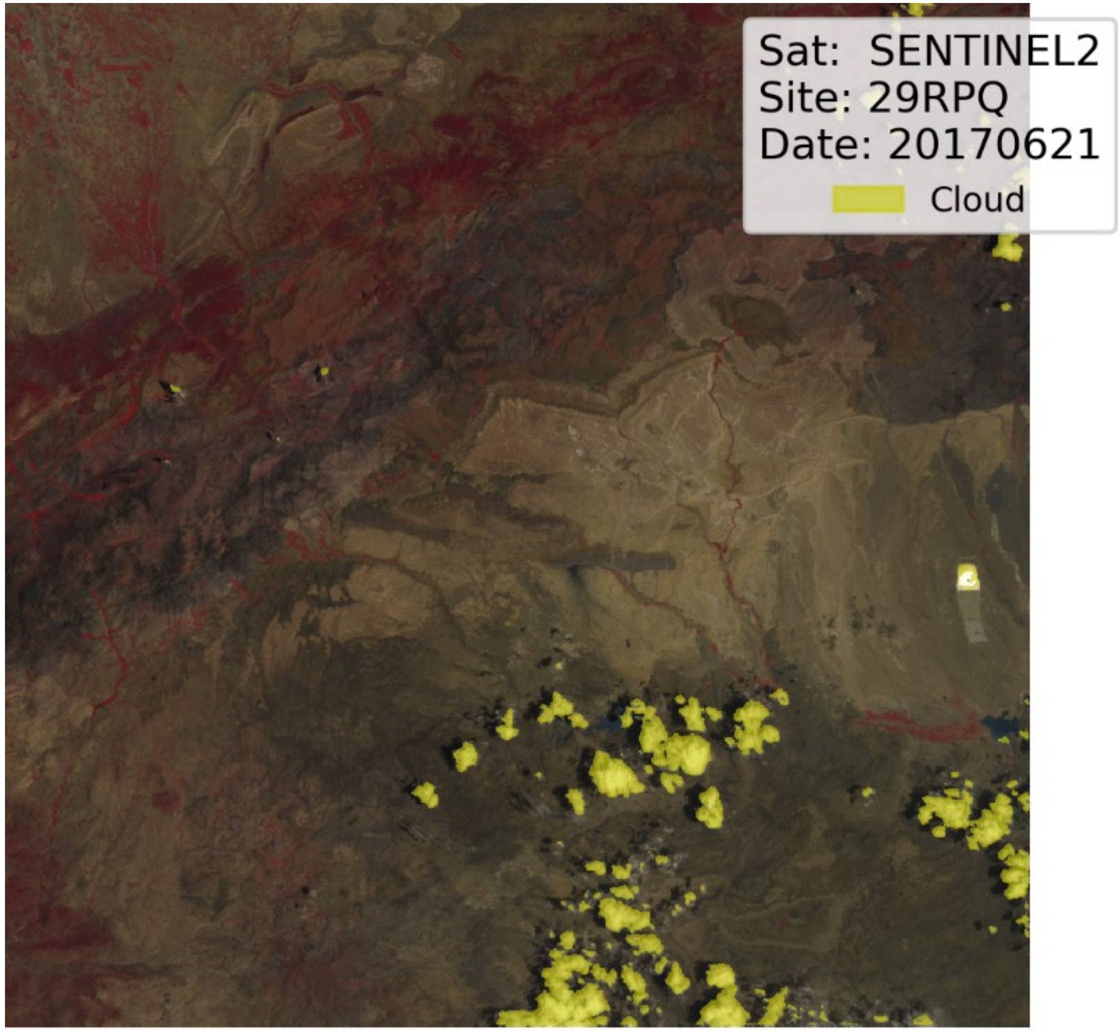
Validation – Sentinel-2

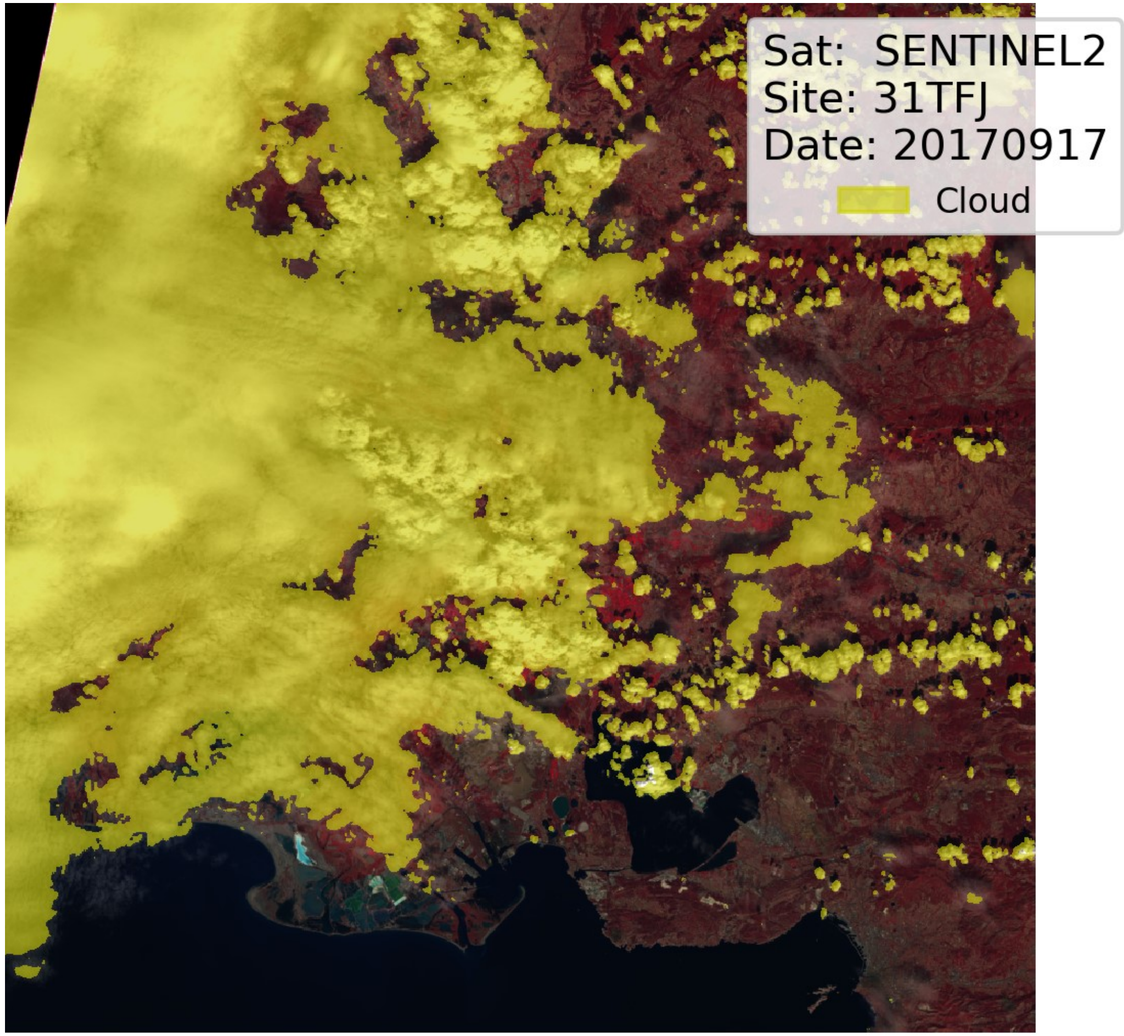
$$F_\beta = (1+\beta^2) * \frac{(precision*recall)}{(\beta^2*precision)+recall} \; with \; \beta = 2$$

- Different cloud classes are unified into one
- Comparison using masks provided by existing cloud-screening processors

Algorithm	F2-score in %
Maja	**91**
FMask	**90**
CloudDL	88
CloudDL with DEM	**90**




Sat: SENTINEL2
Site: 29RPQ
Date: 20170621
Cloud


Sat: SENTINEL2
Site: 31TFJ
Date: 20170917
Cloud

Results

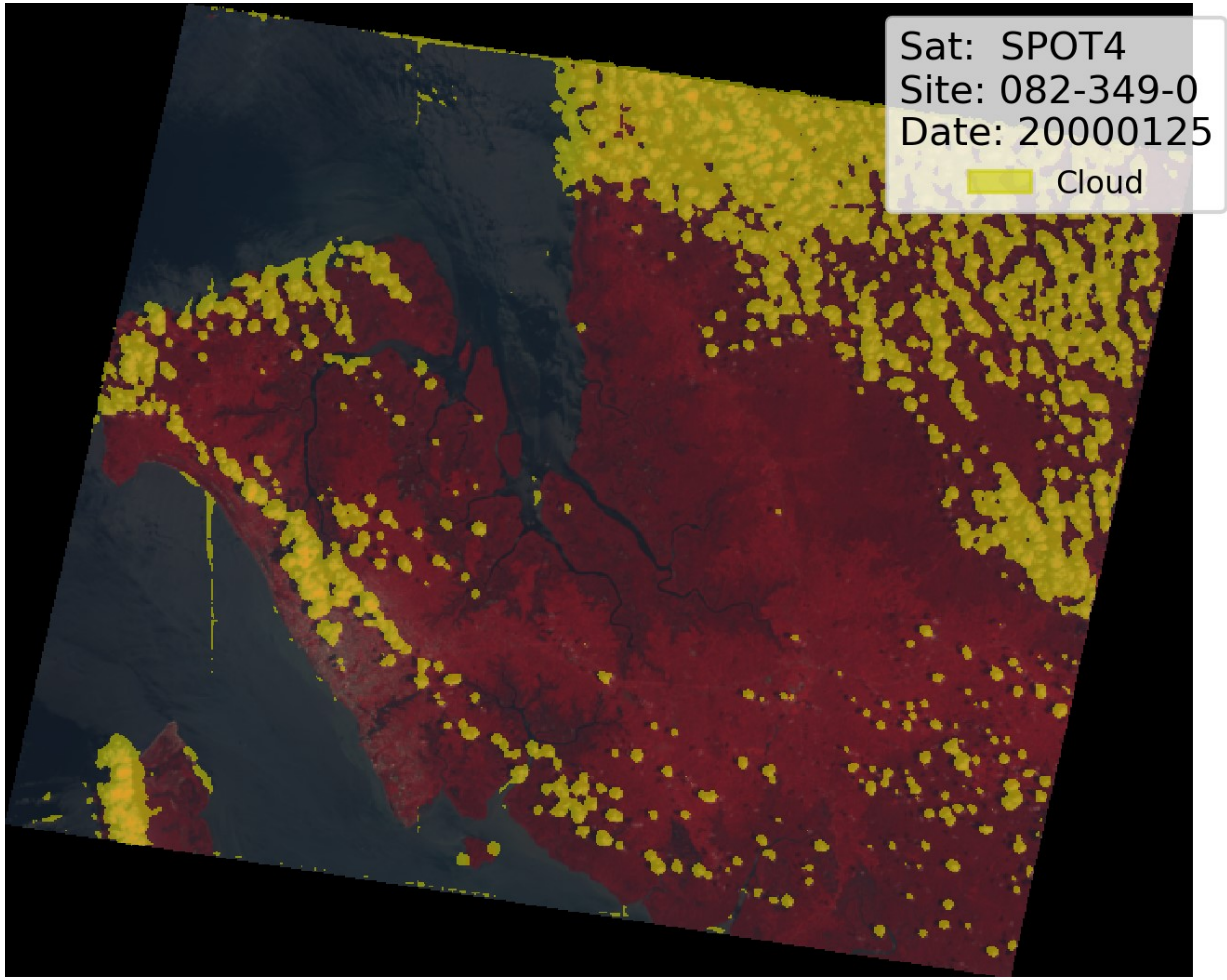
Training: Sentinel-2

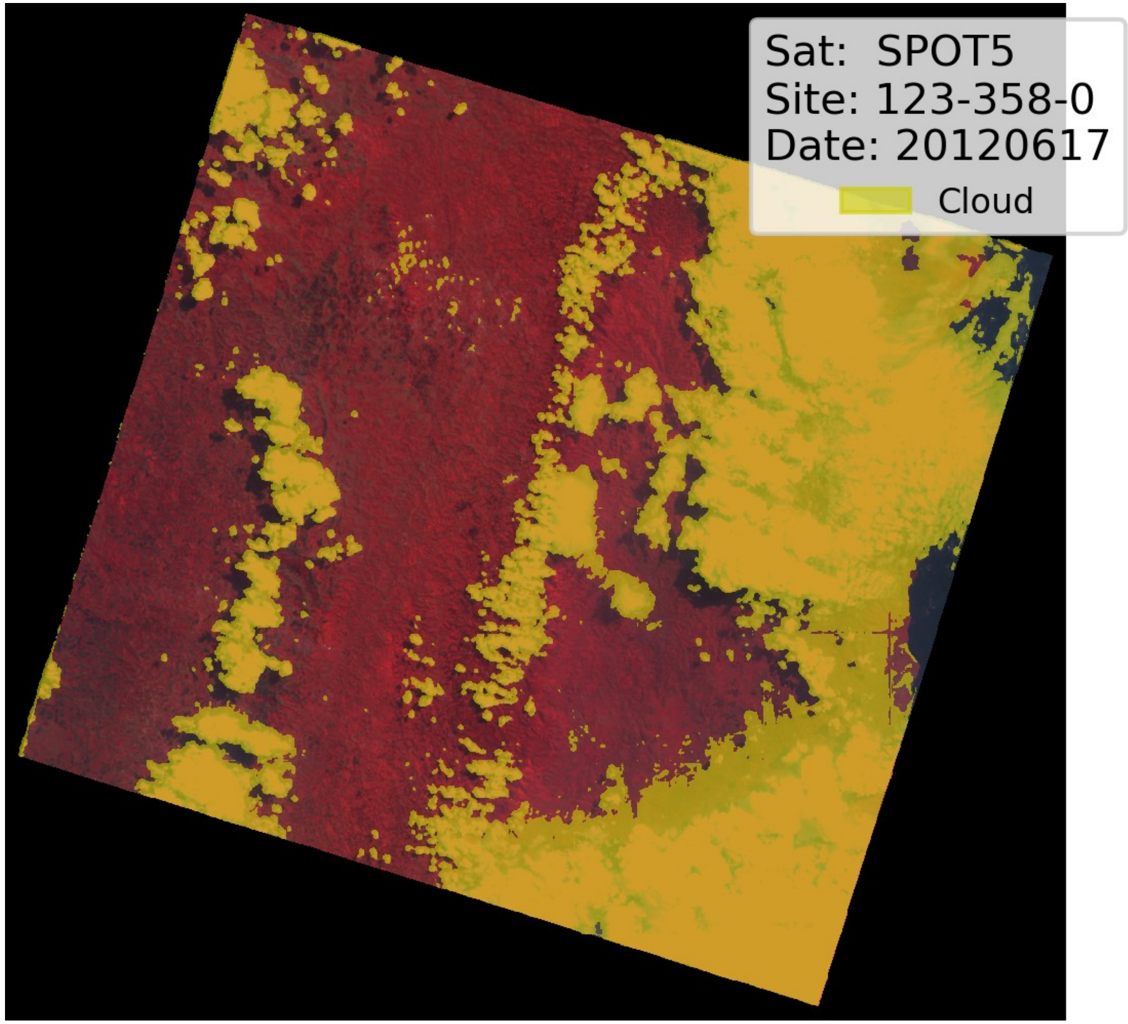
Validation: SPOT-4 and -5

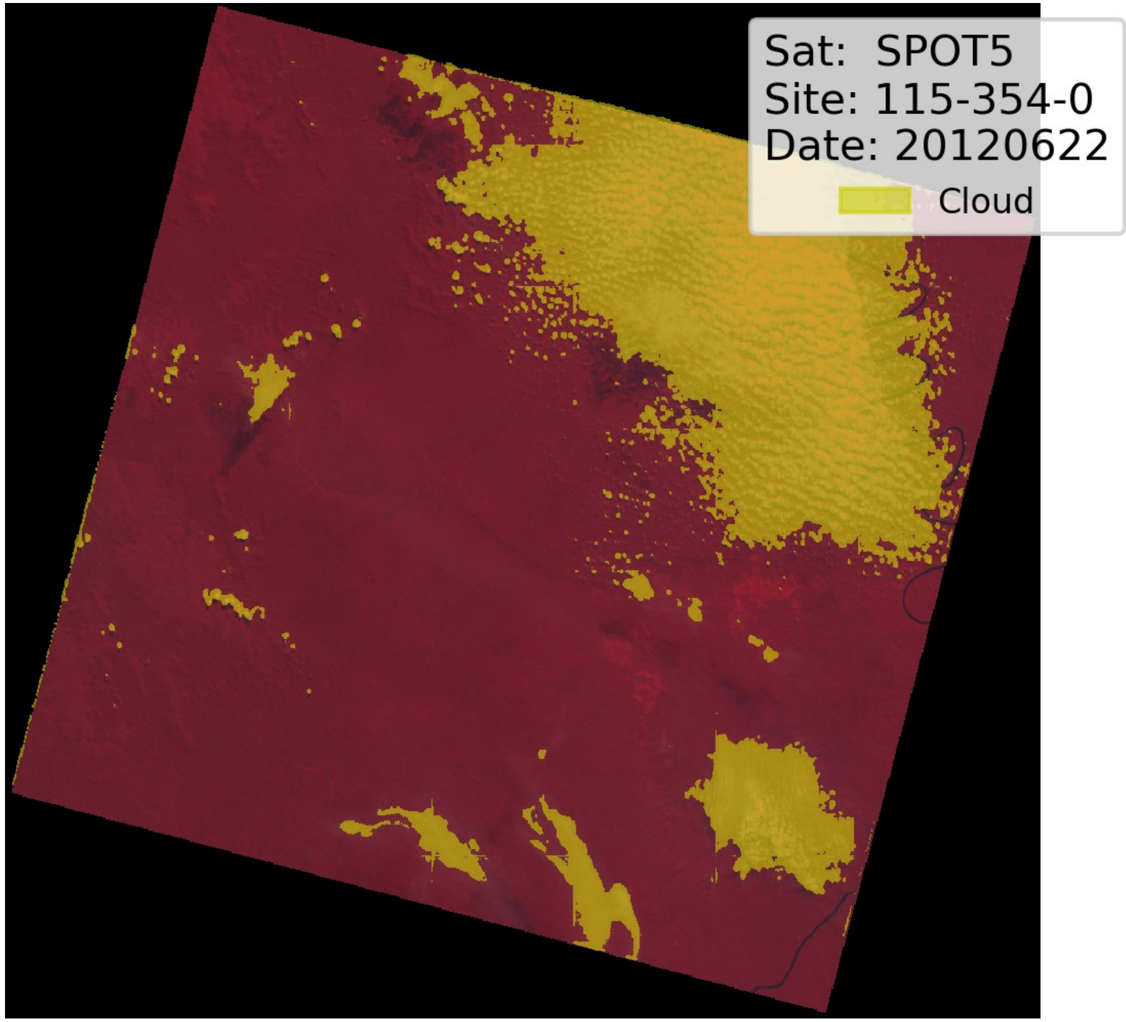
Validation – Spot4/5

- Different cloud classes are unified into one
- Comparison using masks provided by existing cloud-screening processors

Algorithm	Fscore in %
Maja	**91**
CloudDL	85
CloudDL (with DEM)	**89**




Sat: SPOT5
Site: 123-358-0
Date: 20120617
Cloud


Sat: SPOT5
Site: 115-354-0
Date: 20120622
Cloud

Conclusion

- Transfer learning using band adaptation possible
- Results are robust to image quality variations (radiometrically and geometrically)
- Adaptation to other sensors (Pleiades) ongoing

→ Future EO products can be improved using transfer learning and existing databases

Conclusion

- Method ready to be adapted to future platforms, following these steps:
  - Coarse pre-flight model for development purposes
  - Once first ‘real’ images are acquired during LEOP, a dedicated validation database needs to be built
  - Using real images and the new validation database, a refined model for production purposes is trained
  - Throughout instrument lifetime: Regular quality-assurance tests and possible re-training using new samples if quality-gate not met

Thank you

Contact:

Peter.kettig@cnes.fr