

FG-Cloud peut-il aider à traiter de gros volumes de données satellites Sentinel 2 ?

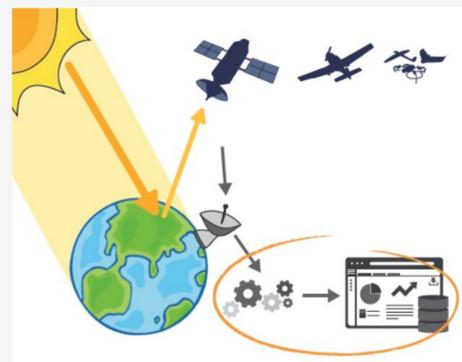
Cas de l'identification et du suivi des peupleraies à large échelle

Hamrouni, Y. ^{1,2} ; Sheeren, D. ² ; Heintz, W. ² ; Blaise, F. ² ; Ladet, S. ²

1- Conseil National Du Peuplier, 47 Rue de Chaillot, France.

2- DYNAFOR, Université de Toulouse, INRA, Castanet-Tolosan, France.

Domaine de recherche: la télédétection



- ✓ Son principe = Une source d'énergie + Une cible + Un rayonnement émis + Un rayonnement réfléchi + Enregistré par un capteur
- ✓ La télédétection englobe le processus d'acquisition de données (images) ET le processus de traitement de ces dernières pour en extraire des informations pertinentes au regard d'une question, d'un besoin d'un utilisateur sur un territoire donné
- ✓ Une révolution récente: lancement des satellites Sentinel-1 et 2 avec une triple haute résolution : spatiale (10m à 20m), spectrale (10 bandes pour Sentinel-2), temporelle (fréquence de revisite tous les 6 jours) et accessible en OpenData
- ✓ Des méthodes et des outils disponibles, libres, en ligne : OTB, Snap, EnMAP Box, Google Engine... Et codes source Python...

→ Rapprochement des compétences/ spécialistes entre Observation de la terre et Numérique (Big Data)

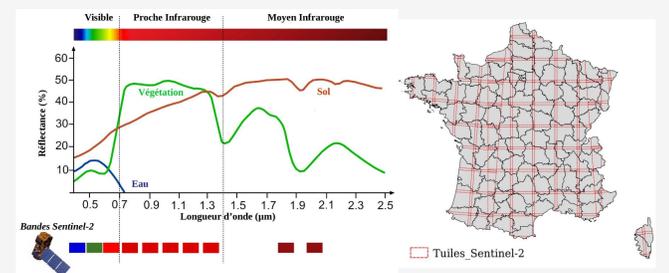
→ Besoin de changer de paradigme dans le traitement des images en y associant de nouveaux moyens d'infrastructures

Source : Marie Jagaille, GéoBretagne



Zoom sur le satellite Sentinel-2

- ✓ Mis en orbite en juin 2015 par l'Agence spatiale européenne (ESA) dans le cadre de Copernicus
- ✓ Equipés de capteurs optiques passifs qui mesurent l'énergie solaire réfléchi par les objets situés en surface de la terre sur 10 bandes spectrales le long du spectre électromagnétique
- ✓ À chaque passage du satellite, 10 images sont générées. Donc volume données /année/France entière = 8 To

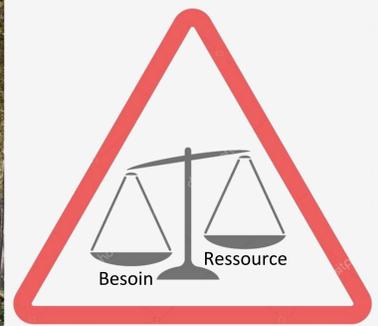


Champ d'application:

- ✓ Bourse Cifre de Yousra Hamrouni (2017-2020)

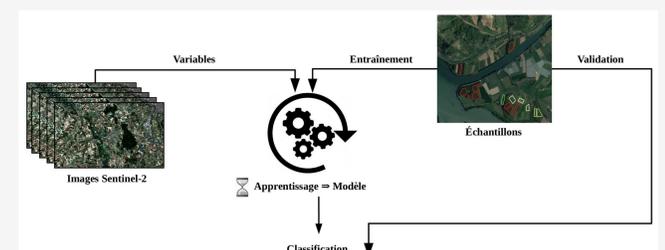
Etude : étudier les capacités de la télédétection à localiser annuellement les peupleraies et suivre leur évolution sur une large étendue

- ✓ Le peuplier est la 1ère essence de feuillus plantée en France pour des usages multiples mais difficulté à connaître la ressource annuelle disponible
- ✓ Une alerte: Besoin en bois > Ressource disponible avec déficit estimé à 500 000 m3/an en 2025



Chaîne de traitement

- ✓ Extraction de variables spectrales issues des images satellitaires (bandes disponibles dans les différentes longueurs d'onde) aux différentes dates d'acquisition => entre 350 et 400 variables.
- ✓ Apprentissage automatique à partir d'un jeu de données d'entraînement extrait des échantillons de référence (50%) et prédiction spatiale sur les tuiles Sentinel-2 à partir du modèle appris.
- ✓ Validation de la classification à partir des échantillons de référence indépendants (50% restants).
- ✓ Pour le passage à l'échelle : apprentissage par transfert (actif) pour adapter le modèle local initial à tout le territoire (travail en cours)



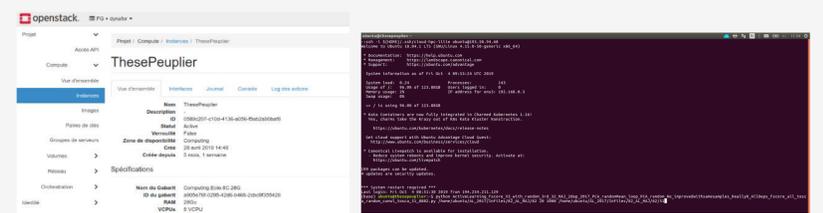
Infrastructure de calcul utilisée à Dynafor

Test des services de calcul dont la solution déportée sur le cloud de France Grilles dans le domaine de la télédétection pour aider ...

... à développer un algorithme informatique pour analyser les images satellites, pour en déduire annuellement une couche cartographique des peupleraies sur le territoire national

- ✓ Existant à Dynafor: un serveur de calcul hors garantie (TOSCA), à 95% saturé: actuellement serveur de production
- ✓ Prise de contact avec France Grilles en décembre 2018 → décision d'un bêta-test
- ✓ Demande validée en avril 2019 avec rédaction de Plan de Gestion de Données et formulaire utilisateur
- ✓ Été-Automne 2019: mise en œuvre de tests unitaires:
 - de téléchargement de fichiers
 - de calcul d'apprentissage actif
 - d'envoi

	TOSCA DELL C6145	FG-Cloud
VM	4*12 =48 cœurs et 300 Go RAM	8 cœurs et 28 Go RAM
Stockage	10 To	125 Go
Logiciels	Orfeo ToolBox 6.4, Python 2 & 3, QGIS2.18	
Bibliothèques	GDAL2, librairies Python	



Conclusions

- ✓ Premiers tests réussis
- ✓ Mais espace de stockage EG-Cloud limité ne permet pas de passer en mode production
- ✓ Bon accompagnement et réactivité du support FG

Perspectives

- ✓ Interfacer le service FG-iRODS de stockage haute disponibilité pour accroître l'espace de stockage
- ✓ Echange à faire avec le Centre d'Expertise Scientifique CES Occupation des sols – THEIA-LAND qui produit carte OSO annuelle

