**Hugo SENGES**

**SOAD**  
Variables numériques en appui aux métiers de L’OFB

Vincennes, le 26 novembre 2021

Automatisation de la fourniture d’images mosaïques recombinées Sentinel 2

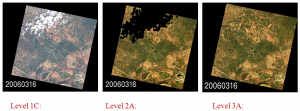
Note de pré-développement concernant une chaîne algorithmique visant la l’encapsulation d’une fonctionnalité de téléchargement d’images mosaïques recombinées.

1. Périmètre et objectif

Dans le cadre de l’activité d’ajout d’une chaîne de traitement dans le package INRAE SISPPEO (dédié à la télédétection des CIPAN[[1]](#footnote-1)), il est nécessaire de disposer d’un process d’acquisition automatisée de mosaïque d’image recombinée, dans laquelle les nuages sont masqué et l’étalonnement réalisé. Cela correspond aux produits délivrés par le pôle Theia pour les données issues de la mission Sentinel 2, produits désignés par leur niveau : 2A et 3A.

Cette fonctionnalité pourrait prendre la forme d’un outil développé en python, associant plusieurs librairies développées à cet effet, éventuellement assorti d’une cartouche CLI permettant l’exécution de commande via un terminal directement. En outre, les méthodes de tuilage et traitement des images Sentinel sont déjà implémentées et validées au sein du package python SISSPEO, hébergé sur les serveurs de l’INRAE. Il s’agit donc de créer, à destination des ingénieurs, techniciens et géomaticiens de l’OFB, en directions nationale et régionales, un outil de téléchargement de la donnée finale, qui limite les étapes intermédiaires, le stockage convoqué, et se concentre sur la recherche de la ressource la plus adaptée.

Deux alternatives se posent alors : développer en interne une solution algorithmique de traitement d’images Sentinel 2 de niveau 1C (brutes ortorectifiées), au temps de revisite de 5 jours, afin de proposer, en continu, l’image composite, désenuagée la plus à jour au regard d’un seuil de couverture effective du territoire considéré ; ou bien se reporter aux produits déjà produits par le pôle Theia, connus sous la dénomination 2A et 3A, et en automatisée l’acquisition par les agents lorsqu’ils en ont besoin, selon une emprise géographique et un pas de temps défini. Ces données sont détaillées ci-après.



Le niveau 3A est produit grâce à la chaîne WASP développée par les équipes du CNES (<https://github.com/CNES/WASP>). C’est un produit composite, une mosaïque d’images acquises durant 45 jours (publié le 15 de chaque mois), ce qui permet d’obtenir une orthophotographie sans nuages, distorsions atmosphériques et avec un étalonnage permettant de limiter les discontinuités entre pixels natifs de clichés différents. Dans les faits, cette ressource est une composition des images de niveau 2A.

Ce produit satisfait le besoin interne pour les ingénieurs et géomaticiens de l’OFB quant à la qualité de la donnée (résolution spatiale, étalonnage, multispectre). Son seul défaut est lié à sa résolution temporelle, qui dépend des routines de calculs du pôle Theia, et n’est donc pas modulable selon les besoins des services.

Le niveau 2A est quant à lui produit via la chaîne MAJA (<https://github.com/olivierhagolle/Start_maja>) , sur une seule date, donc un seul cliché duquel sont enlevé nuages et distorsions dues au relief, aux aérosols et aux conditions atmosphériques. La chaîne de traitement sous-jacente est disponible en annexe. Ce niveau répond à la problématique de standardisation des images Sentinel optiques mais reste un produit potentiellement lacunaire en termes de donnée affichée :

« Les chaînes qui utilisent des produits de niveau 1C ou de niveau 2A de Sentinel-2 doivent donc être résistantes à la présence de données manquantes dans la couverture spatiale ou temporelle. »[[2]](#footnote-2)

Dans les deux cas, la résolution spatiale native de 10m est inchangée. L’accès à ces données est possible via la plateforme de téléchargement Theia dédié[[3]](#footnote-3). Celle-ci permet d’accéder aux données fournies par Theia[[4]](#footnote-4), à savoir :

* Sentinel-2 N2A et N3A (qui nous intéressent ici)

Mais également :

* Landsat-8
* Landsat 5-7-8 N2A ancien format
* Occupation des sols
* Pléiades RTU
* Snow N2B et N3B
* SPOT 5 stereoscopic survey of Polar Ice
* Spot World Heritage
* Spot world heritage ancien format
* Venµs
* Qualite des eaux continentales

L’obtention des données de réflectance Sentinel-2 par le biais de la plateforme développée par Theia requiert cependant plusieurs éléments.

Tout d’abord un environnement Pyhton compatible (et pour la cartouche .CLI, les invites de commandes adéquats). La documentation à ce sujet est fournie dans le répertoire github pointé précédemment. Ensuite il faut disposer des références du produit à télécharger. On peut résumer ici ces références selon quatre grands domaines :

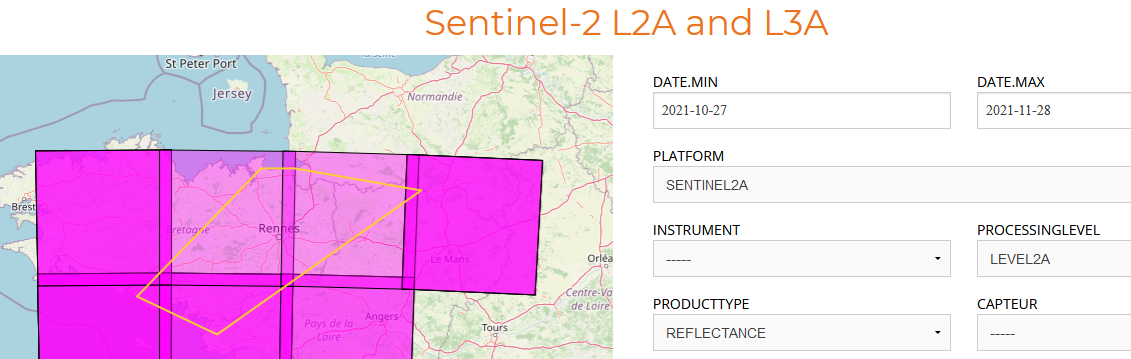
* la définition temporelle (période d’acquisition considérée),
* le produit visé (senseur et niveau de traitement ; ex : « Sentinel-2 N2A »)
* l’emprise géographique considérée (cf. infra)
* le format de données délivré (archive ou GeoTiff)

Enfin, il faut également disposer d’un compte Theia permettant d‘accéder aux données. Les paramètres de connexion de celui-ci sont alors encapsulés dans le fichier config\_theia.cfg.

Bien que la création de compte Theia pour les organismes publics semble gratuit et de droit, la gestion des mail de validation de création de compte semble sous-optimale. Un contact au sein des équipes de ce pôle pourrait ainsi faciliter l’obtention de comptes pour les agents de l’OFB.

**Sur l’emprise géographique :**

Le référencement des ressources images à télécharger en fonction de l’emprise au sol considéré est une opération qui passe *in fine* par le recoupage de cette zone avec le dallage de Sentinel 2. Les différentes tuiles qui découpent le l’emprise au sol sont référencées par un code (« T31TCJ» par exemple) et, si l’outil de téléchargement proposé par Theia autorise d’autres paramètres géographiques que la référence des tuiles visées, il peut être utile de connaître les références des tuiles qui couvrent l’espace de travail considéré. Pour cela il existe une librairie Python, EOTile[[5]](#footnote-5) qui permet à travers la saisie d’une aire WKT, d’une chaîne de caractère géoréférencée ou bien d’une BBox, d’obtenir la liste des tuiles concernées.



**Point sur le développement d’une chaîne de traitement équivalant 3A en interne OFB :**

La question de savoir d’il est opportun de développer en interne un outil de production automatisé de mosaïque d’images composées sur la base de ressources de niveau 2A est difficile à trancher. Au regard de l’importance relative d’un tel projet de développement, les plus-values vis-à-vis de l’adoption des outils de téléchargement livrés par Theia doivent être clairement identifiés. Deux pistes principales sont à l’étude à ce propos. D’une part, le développement d’un outil interne permettrait d’accroître l’indépendance des services de l’oFB et leur résilience vis-à-vis d’une plateforme (Theia) dont on a pas la main sur les conditions de maintenance, les changements éventuels de standards…etc. D’autre part un tel développement pourrait permettre de proposer aux agents un environnement plus modulaire et une ressource de niveau 3A générée en continu.

Autrement dit, il est envisageable de produire une chaîne de traitement qui produise, sur la base de paramètres similaires à ceux évoqués dans le cas de l’outil Theia, un produit similaire au niveau 3A de Theia, mais dont la génération sur le moment permettrait d’assurer que le produit soit le plus à jour possible.

Produire une mosaïque d’image sur le moment, cela consisterait en la fusion successive des dernières ressources publiées au niveau 2A, jusqu’à atteindre un certain seuil de couverture du territoire (mettons 90%). Dès lors, chaque pixel doit être rattachable à la date de prise de vue la plus récente sur laquelle l’apparaît (du moins si la ressource ainsi produite vise à être intégrée dans des protocoles de télédétection).

La question principale est alors de savoir si le temps de développement d’une telle application, ainsi que les temps de traitement subséquents (le téléchargement de toutes les données de niveau 2A ainsi que le calcul là la volée jusqu’à la satisfaction du critère de couverture du territoire), qui peuvent être relativement importants, est pertinente au regard de du gain apporté par cette solution interne. Il y a principalement à gagner sur l’autonomie de l’OFB et sur la résolution temporelle du produit final. Sachant que les ressources de niveau 3A sont publiées mensuellement par Theia, la question est de savoir si l’on peut espérer gagner suffisamment de temps via un outil propre, relativement au critère de couverture du territoire. Cette question ne pourrait toutefois être tranchées qu’en menant plus avant une analyse sur la base des chroniques 2A existantes.

1. Logigrammes

Saisie via CLI des paramètres images requis

Obtention d’un compte auprès de Theia :   
<https://sso.theia-land.fr/theia/register/register.xhtml>

Téléchargement de la suite pyhton dédiée :

Deploiement d’un outil de téléchargement des données 2A via Theia

s

s

Téléchargement de la suite pyhton dédiée :

Saisie via CLI des paramètres géographiques requis

Adoption de la suite EOTile pour référencement des Tuiles Sentinel

Temps nécessaire estimé : 0.5 H/Jeq pour la construction/adaptation de l’environnement Python

s

Production d’une image de niveau 3A à jour

Contrôle du critère de remplissage de l’emprise et relance éventuelle de l’itération

Paramétrage géographique, temporel et spectral de la ressource recherchée

Téléchargement itératif des ressources images de niveau 2A (via Theia ou Sen2-Agri) :   
http://www.esa-sen2agri.org/

Concaténation des pixels valides dans le canevas de l’emprise générale

Temps nécessaire estimé : 15 H/Jeq pour le développement de la solution

1 H/Jeq pour la reprise par les agents concernés  
Certainement plusieurs heures de calculs pour une emprise géographique de l’ordre de grandeur d’un département

s

Temps nécessaire estimé : 1 H/Jeq pour la prise en main de l’environnement

s

1. Annexes

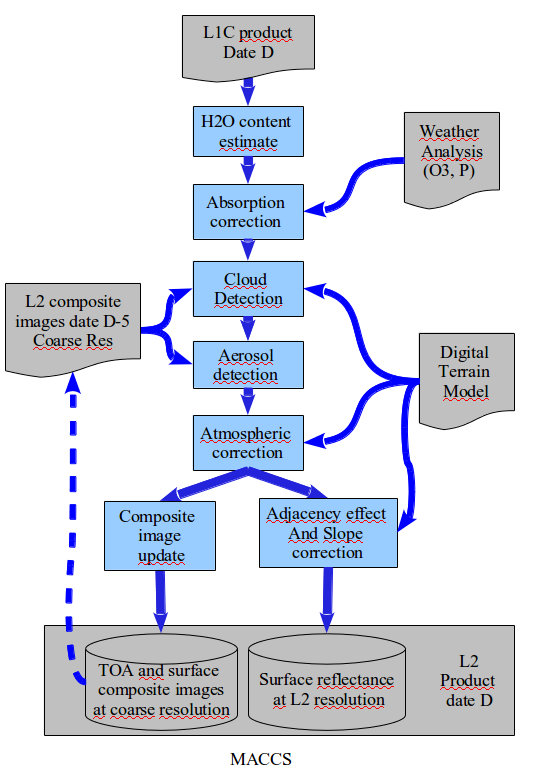


Figure : Chaîne de traitement MAJA pour données Sentinel-2 niveau 2A

Relance du bloc avec la couche RPG obtenue et la couche immédiatement plus ancienne

Push de la couche dans une table postgis dédiée

Affichage en sortie des paramètres d’accès à la ressource

MISE A DISPOSITION

Calcul des attributs dérivés (géométriques) et vote majoritaire sur les profils de culture antérieurs

ENRICHISSEMENT ATTRIBUTAIRE DE LA COUCHE OBTENUE

Statistiques zonales sur les profils de culture des années 2015-2019 depuis les couches RPG

Mise en forme de la table à 24 attributs et clé primaire

1. Cultures Intermédiaires Piège à Nitrate [↑](#footnote-ref-1)
2. Voir CESBIO : <https://labo.obs-mip.fr/multitemp/les-syntheses-mensuelles-de-niveau-3a-de-theia/> [↑](#footnote-ref-2)
3. Voir : <https://github.com/olivierhagolle/theia_download> [↑](#footnote-ref-3)
4. Voir : <https://theia.cnes.fr/atdistrib/rocket/#/home> [↑](#footnote-ref-4)
5. Voir : <https://pypi.org/project/eotile/> [↑](#footnote-ref-5)