

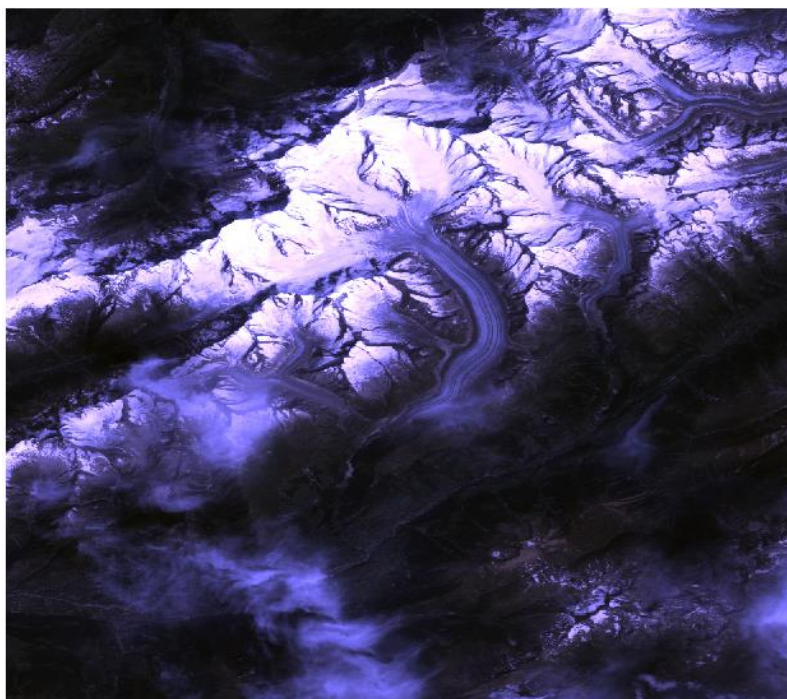
UNIVERSITÉ DE GENÈVE  
FACULTÉ DES SCIENCES DE LA SOCIÉTÉ  
DÉPARTEMENT DE GÉOGRAPHIE ET ENVIRONNEMENT  
CERTIFICAT COMPLÉMENTAIRE EN GÉOMATIQUE

---

Swiss Data Cube:  
Observations de la Terre pour le monitoring de l'environnement en  
Suisse dans le temps et dans l'espace  
Programme des Nations Unies pour l'Environnement  
GRID Genève

---

Laetitia Brossin  
Novembre 2017



Scène du glacier d'Aletsch provenant du satellite Landsat 5 datant du 18 août 1987, en vraies couleurs. Source: GRID-Genève, LB.

Superviseurs:  
M. Bruno Chatenoux  
Dr. Gregory Giuliani

## TABLE DES MATIÈRES

---

1. Résumé.....	5
2. Global Resource Information Database (GRID), Genève.....	6
3. Introduction.....	7
4. Le Cube de Données Suisse (Swiss Data Cube (SDC)).....	8
5. Objectifs et déroulement du stage.....	9
6. Méthodologie.....	10
6.1. Utilisation et évolution de la plateforme sdc.unepgrid.ch.....	10
6.1.1. <i>Custom Mosaic</i> .....	10
6.2. Suivi de la végétation.....	12
6.2.1. <i>Fractional Cover</i> .....	12
6.2.2. <i>NDVI Anomaly</i> .....	12
6.3. Tests des scripts établis à l'aide du notebook Jupyter.....	13
7. Résultats.....	14
7.1. Utilisation et évolution de la plateforme sdc.unepgrid.ch.....	14
7.1.1. <i>Custom Mosaic</i> .....	14
7.2. Suivi de la végétation.....	19
7.2.1. <i>Fractional Cover</i> .....	19
7.2.2. <i>NDVI Anomaly</i> .....	21
7.3. Tests des scripts établis à l'aide du notebook Jupyter.....	24
8. Discussion.....	32
8.1. Utilisation et évolution de la plateforme sdc.unepgrid.ch.....	32
8.1.1. <i>Custom Mosaic</i> .....	32
8.2. Suivi de la végétation.....	33
8.2.1. <i>Fractional Cover</i> .....	33
8.2.2. <i>NDVI Anomaly</i> .....	34
8.3. Avantages et améliorations de la plateforme sdc.unepgrid.ch.....	34
8.4. Tests des scripts établis à l'aide du notebook Jupyter.....	36
8.5. Avantages et améliorations des scripts.....	36
9. Conclusion.....	38
10. Déroulement du stage.....	39
11. Bibliographie.....	40
12. Annexes.....	43
A. Scripts principaux.....	43

A.1. <i>Data_Cube_API_Demo-CopyLB.py</i> .....	43
A.2. <i>test_scenes_video_LB.py</i> .....	45
A.3. <i>test_scenes_video_2_LB.py</i> .....	47
A.4. <i>test_kmean_clustering_LB.py</i> .....	49

Figure 1 - Le concept de Data Cubes. (CEOS, Data Cubes for Large Scale Data Analytics 2017).....	7
Figure 2 - Le Cube de Données : empilements des séries temporelles et scènes Landsat face aux tuiles de données. L'empreinte spatiale des scènes change à travers le temps, tandis que les données des tuiles conservent une empreinte constante. (Lewis, et al. 2015). .....	8
Figure 3 - Les tuiles recouvrant le territoire Suisse (Giuliani, Peduzzi, et al. 2017).....	8
Figure 4 - Interface du Swiss Data Cube.....	10
Figure 5 - Glacier d'Aletsch, mai à août 2013, Landsat 8. A gauche, en vraies couleurs (RGB). A droite, en fausses couleurs PIR, SWIR1, SWIR 2. Les <i>no data</i> apparaissent en noir.....	11
Figure 6 - (a) (b) Région des Portes du Soleil, hiver 2000-2001, Landsat 7. A gauche, en vraies couleurs. A droite, en fausses couleurs PIR, rouge, vert. (c) Couverture neigeuse, région au nord du lac de Neuchâtel, hiver 1999-2000, Landsat 7. Fausses couleurs PIR, SWIR 1, SWIR 2.	11
Figure 7 - Glaciers de Ferpècle et du Mont Miné (cercle rouge), Landsat 8. A gauche, problème du <i>cf_mask</i> , mai 2013 à décembre 2016, vraies couleurs. La neige et la glace apparaissent partiellement en bleu, les <i>no data</i> en noir. A droite, problème résolu, pour la même période. Les <i>no data</i> apparaissent en transparent ici, un fond noir a été ajouté pour faire ressortir la neige (images produites à des dates différentes).....	14
Figure 8 - Glaciers de Ferpècle et du Mont Miné, Landsat 7, mai à août 2002, vraies couleurs. A gauche (cercle rouge), la neige et la glace apparaissent partiellement en blanc, les <i>no data</i> en noir. A droite, zoom dans le cercle rouge. Les <i>no data</i> apparaissent en transparent ici, un fond noir a été ajouté pour faire ressortir la neige (images produites à des dates différentes dans le SDC). .....	15
Figure 9 - Glacier du Rhône (cercle rouge), Landsat 7, mai à août 2001, vraies couleurs.....	15
Figure 10 - Glacier d'Aletsch, Landsat 7, mai à août 2001, vraies couleurs. ....	15
Figure 11 - Evolution du couvert neigeux avec les bandes PIR, SWIR 1, SWIR 2. Région du Canton de Neuchâtel. A gauche, Landsat 5, hiver 1989-1990. A droit, Landsat 7, hiver 1999-2000. La neige apparaît en rouge, la végétation en beige, les <i>no data</i> en noir, les zones agricoles en orange et bleu, le tissu urbain en bleu-gris.....	16
Figure 12 - Evolution du couvert neigeux avec les bandes PIR, rouge, vert. Région des Portes du Soleil avec le tissu urbain de Villeneuve à Martigny (Valais). A gauche, Landsat 7, hiver 2000 – 2001. A droite, Landsat 8, hiver 2015 – 2016. ....	16
Figure 13 - Bâle, bandes PIR, rouge, vert. A gauche, Landsat 7, en 2000. A droite, Landsat 8, en 2016. ....	17
Figure 14 - Lausanne - Cossonay. <i>En haut</i> : vraies couleurs. La luminosité a été augmentée de 20%. A gauche, Landsat 7, mai à août 2000. A droite, Landsat 8, mai à août 2016. <i>En bas</i> : bandes PIR, rouge, verte, min. NDVI. A gauche, Landsat 7, mai à août 2000. A droite, Landsat 7, mai à août 2005.....	18
Figure 15 - Bulle (Fribourg). La luminosité a été augmentée de 20% pour les deux images. A gauche, Landsat 5, du 1 <sup>er</sup> janvier 1984 au 1 <sup>er</sup> janvier 1985. A droit, de mai à août 2013. ....	18
Figure 16 - Gamme de couleurs pour le Fractional Cover. PV: photosynthétique - NPV: Non photosynthétique - BS : sols nus.....	19
Figure 17 - Parc Jura vaudois et une partie de la zone agricole dans le Canton de Vaud, Landsat 8, mai à août 2016. On remarque une très forte proportion en sols nus (en bleu) alors que cette zone	

	est aussi composée de forêts. Les <i>no data</i> apparaissent en noir. La luminosité a été augmentée de 20%.....	19
Figure 18 -	Partie du Tessin. A gauche, de mai à août 2001. A droite, de novembre à décembre 2001. Landsat 7, méthode du pixel médian. La luminosité a été augmentée de 20% pour les deux images .....	20
Figure 19 -	Région des Portes du Soleil. A gauche, de mai à août 2010. A droite, de septembre à novembre 2010. Landsat 5, méthode du pixel médian. La luminosité a été augmentée de 20% pour les deux images. ....	20
Figure 20 -	Région des Portes du Soleil. A gauche, Landsat 5, de mai à août 1987. A droite, Landsat 8, de mai à août 2016. Méthode du pixel médian. La luminosité a été augmentée de 20% pour les deux images.....	20
Figure 21 -	Bassin genevois. A gauche, Landsat 7, mai à août 2000. A droite, Landsat 8, mai à août 2016. Méthode du pixel médian. La luminosité a été augmentée de 20% pour les deux images.....	21
Figure 22 -	Echelles des valeurs en couleurs pour (1) l'écart de NDVI (2) le pourcentage de changement du NDVI. Un pourcentage de changement de NDVI de +50% indique que la végétation s'est améliorée de +50% par rapport à la médiane. (3) Le NDVI.....	21
Figure 23 -	Extrait de l'onglet <i>Résultat</i> de la plateforme <i>sdc.unepgrid.ch</i> pour le NDVI Anomaly. A gauche, on peut voir la période de temps sélectionnée à partir de laquelle la scène unique sera choisie, les mois pour la <i>baseline</i> et les manières de représenter les résultats sur la carte (cases à cocher). A droite, la liste d'acquisition des scènes (pour la scène unique, ou le <i>Time Range</i> ).....	22
Figure 24 -	Gestionnaire des tâches, paramètres des tâches.....	22
Figure 25 -	Ecart de NDVI. Landsat 8. Zone de forêt dans la région du parc Jura vaudois et traversant la frontière française. La scène de référence est le mois d'août 2016, avec une scène unique datant du 31. La <i>baseline</i> est calculée avec les mois de juin, juillet et août. ....	23
Figure 26 -	Pourcentage de changement de NDVI. Idem que la Figure 24.....	23
Figure 27 -	<i>Baseline</i> pour le calcul de la médiane. Idem que la Figure 24. ....	23
Figure 28 -	Différence de rendu. Les deux scènes représentent la région des glaciers de Ferpècle et de Mont Miné, le 30 juin 2001 avec le Landsat 7. A gauche, la scène est téléchargée depuis le site <i>sdc.unepgrid.ch</i> . A droite, elle est produite via le script Python. ....	25
Figure 29 -	Exemples de rendu pour les scènes – glaciers de Ferpècle et du Mont Miné en 1987 et en 2010 (Landsat 5, bandes PIR, rouge, vert).....	26
Figure 30 -	Exemples de rendu pour les scènes – glacier du Rhône en 1990 et en 2010 (Landsat 5, bandes PIR, rouge, vert).....	26
Figure 31 -	Exemples de rendu pour les scènes uniques – glacier d'Aletsch en 1990, 2001 et 2016 (Landsat 5,7 et 8, vraies couleurs). Pour la scène de 1990, la luminosité a été augmentée de 20%. Pour celle de 2016, elle a été augmentée de 40%.....	27
Figure 32 -	Scènes uniques région des Portes du Soleil. L'image de droite a été produite avec le Landsat 7, celle de gauche avec le Landsat 8.....	28
Figure 33 -	Mosaïques pour la région des Portes du Soleil. À gauche, hiver 2001-2001, Landsat 7. A droite, hiver 2016-2017, Landsat 8.....	28
Figure 34 -	Bâle. Landsat 7 et 8, vraies couleurs. A gauche, en 2000. A droite, en 2016.....	29
Figure 35 -	Lausanne. Landsat 7 et 8, vraies couleurs. En haut, en 2000. En bas, en 2016.....	29
Figure 36 -	Zürich. Landsat 7 et 8, vraies couleurs. En haut, en 2000. En bas, en 2016.....	30
Figure 37 -	Genève. Landsat 7 et 8, vraies couleurs. A gauche, en 1999. A droite, en 2016.....	30
Figure 38 -	Bulle, bandes PIR, rouge, vert. A gauche, Landsat 5, du 1 <sup>er</sup> janvier 1984 au 31 janvier 1985. A droite, Landsat 8, de mai à août 2016. ....	31

## 1. RÉSUMÉ

---

Une interface internet a été créée par l'équipe du GRID-Genève (<http://sdc.unepgrid.ch/>) pour utiliser le Swiss Data Cube (SDC), nouvel environnement de travail open source permettant d'analyser les changements environnementaux grâce à des données satellitaires. Le but premier de ce stage a donc été de tester les fonctionnalités, possibilités et limites de l'environnement du SDC afin de montrer ses potentialités et de donner des conseils pour son évolution. L'algorithme employé pour ce premier objectif est la création de mosaïques (*Custom Mosaic*). Un second objectif a été de savoir s'il est possible d'utiliser deux autres algorithmes afin de mesurer les changements environnementaux et d'effectuer un suivi des écosystèmes à travers l'étude de la végétation (changements en qualité, en quantité et en type de végétation, comparaisons de milieux dans le temps et l'espace) : le *Fractional Cover* et le *NDVI Anomaly*. Pour terminer, des scripts permettant d'interagir avec le SDC ont été élaborés afin de reproduire de manière automatisée les images et vidéos établies lors des deux premiers objectifs.

Les options des filtres à disposition ont été testées sur l'interface. Les types de résultats ont été comparés entre eux pour déterminer quelles compositions de bandes faisaient le plus ressortir les phénomènes analysés. Concernant les méthodes de composition, des comparaisons entre le pixel le moins récent, le plus récent et le pixel médian ont été faites, ainsi que quelques tests avec les NDVI minimum et maximum. Les satellites Landsat 5, 7 et 8 ont été utilisés avec toutes les années à disposition (de 1984 à 2016) pour obtenir les meilleurs rendus et comparer les différents satellites. Pour le suivi de la végétation, des zones correspondant à différentes utilisations et occupations du sol ont été sélectionnées.

D'une manière générale, les tests de l'interface du SDC ont bien fonctionné. Selon les séries d'images produites et les objectifs de départ, il a été décelé que la pixellisation peut s'avérer trop grossière à certains endroits et qu'il y a beaucoup de *no data*. Mais dans la majorité des cas le rendu est relativement précis ; la qualité des images va du moyen à l'excellent. Il n'a toutefois pas été possible de répondre au deuxième objectif dû au ralentissement du server et à une mémoire parfois insuffisante rendant l'utilisation de l'interface impossible. Par conséquent, il n'y a pas eu une production suffisante d'images permettant d'observer des tendances sur l'évolution de la végétation. Concernant les scripts, ils ont permis un gain de qualité et de temps considérable.

A l'avenir, une étude approfondie et rapide des changements environnementaux devrait pouvoir être faite en vue des résultats encourageants obtenus. En effet, la qualité et précision des indices et de leur rendu graphique permettraient des analyses précises dans le temps et l'espace selon les objectifs des recherches.

## **2. GLOBAL RESOURCE INFORMATION DATABASE (GRID), GENÈVE**

---

Le GRID Genève est l'un des bureaux de la division scientifique de l'Organisation des Nations Unies (ONU). Les centres GRID ont plusieurs tâches (GRID Genève 2017) :

- ils facilitent l'accès à des données et à de l'information scientifique environnementales;
- grâce à ces données, ils participent à la prise de décision et la mise en place de politiques en évaluant les menaces et les changements environnementaux, notamment au travers d'alertes ou de pré-alertes à différentes échelles (du local au global);
- ils participent à la révision de l'état de l'environnement ainsi que de ses tendances.

Le GRID Genève est soutenu et collabore avec différentes institutions, dont l'Office Fédéral de l'Environnement (OFEV) et l'Université de Genève.

### 3. INTRODUCTION

---

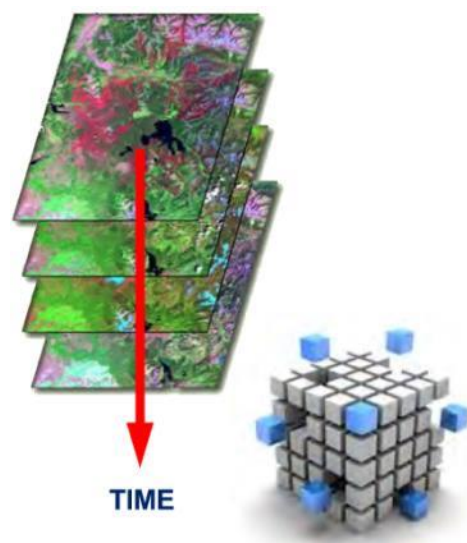
Le concept de Cube de Données (ou Data Cube, DC) part de la volonté d'améliorer l'utilisation des données d'observation de la Terre qui sont nombreuses, gratuites et en libre accès (CEOS and GRID Genève, The CEOS Open Data Cube Initiative 2016). En outre, elles diffèrent entre elles en termes de conditions d'observation (conditions atmosphériques, nuages, géométrie, etc.) et de capteurs individuels issus de multiples engins spatiaux (différentes résolutions spatiale, spectrale et radiométrique).

Les données satellitaires stockées dans le DC sont des Analysis Ready Data (ARD), autrement dit des « *données satellitaires qui ont été traitées selon un ensemble minimal d'exigences et organisées en une forme qui permet une analyse immédiate sans effort supplémentaire de l'utilisateur* » (Killough 2016). Les données des ARD doivent ainsi (i) décrire des métadonnées; (ii) avoir subi un calibrage radiométrique; (iii) avoir subi un calibrage géométrique; (iv) avoir subi des calibrages solaires et atmosphériques (pour les capteurs optiques) (Killough 2016).

C'est pourquoi le Data Cube s'est vu être une solution au défi d'intégration, d'analyse et d'application des données d'observation de la Terre. Comparé à des technologies antérieures, il offre la possibilité d'analyser, de traiter et de restructurer ces données spatiales **en un seul environnement de travail** (une seule infrastructure) de haute performance, accessible pour tout utilisateur. Il exploite ainsi des séries temporelles (allant jusqu'à l'échelle continentale) dont les données sont disponibles à travers plusieurs décennies.

L'image conceptuelle du Data Cube fait référence à un empilement multidimensionnel de séries temporelles de pixels alignés spatialement (Figure 1). Il possède donc les deux dimensions spatiales, x et y, ainsi qu'une dimension temporelle. Les pixels peuvent provenir de plusieurs jeux de données ou d'un seul jeu de données (CEOS, Data Cubes for Large Scale Data Analytics 2017, CEOS, Australian Government, CSIRO, USGS 2017).

Des algorithmes identifiant les thèmes des différentes couvertures et utilisations du sol (regroupés sous forme d'outils et d'indices) ont également été développés et implémentés dans les Data Cubes (Lewis, et al. 2015, Office fédéral de la statistique 2017).

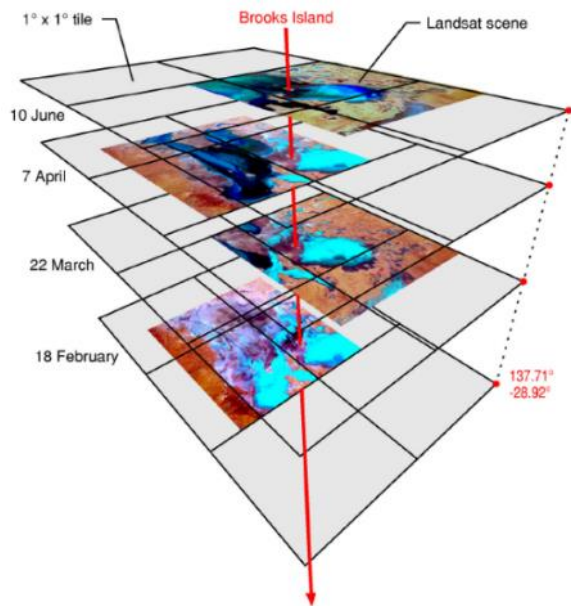


**Figure 1** - Le concept de Data Cubes. (CEOS, Data Cubes for Large Scale Data Analytics 2017)



## 4. LE CUBE DE DONNÉES SUISSE (SWISS DATA CUBE (SDC))

Les données des satellites Landsat ont été choisies pour être implémentées dans le SDC. En effet, elles offrent les plus longues séries temporelles de données multispectrales à résolution modérée et recouvrent la surface de la Terre (United States Geological Survey 2017). Elles permettent donc de créer des séries chronologiques, de comparer différentes périodes de temps et d'observer des tendances aux différentes échelles. Ajouté qu'en 2008, l'United States Geological Survey a rendu disponibles gratuitement toutes les données Landsat sur Internet (Giuliani, Chatenoux, et al. 2017).

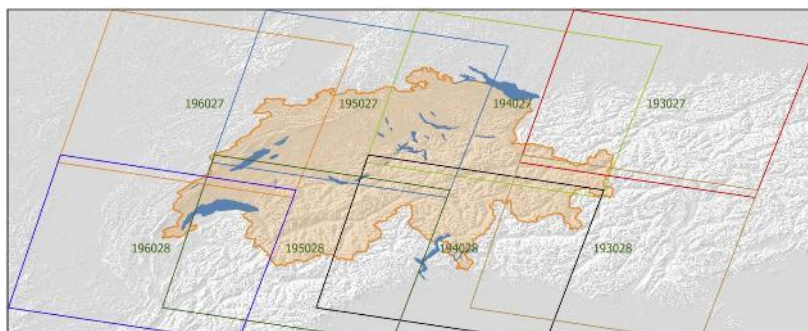


**Figure 2** - Le Cube de Données : empilements des séries temporelles et scènes Landsat face aux tuiles de données. L'empreinte spatiale des scènes change à travers le temps, tandis que les données des tuiles conservent une empreinte constante. (Lewis, et al. 2015).

Un critère essentiel à la mise en place du SDC a donc été la génération automatique des produits ARD Landsat. Ceci a permis de passer une autre étape clé, celle de la segmentation des données en "tuiles" (*tiles*) de même taille, ces tuiles formant une grille permettant une utilisation efficace des capacités de l'environnement du DC (Figure 2) (Lewis, et al. 2015, Giuliani, Chatenoux, et al. 2017).

Le territoire suisse est recouvert par huit tuiles (Figure 3). Le Swiss Data Cube comporte les scènes Landsat 5, 7 et 8 (de 1984 à aujourd'hui). La résolution des pixels est de 30m x 30m (Giuliani, Peduzzi, et al. 2017). Une interface graphique via internet a été créée pour interagir avec l'environnement du SDC : <http://sdc.unepgrid.ch/> (CEOS, GRID Genève 2017).

Les différents outils et indices implémentés dans ces interfaces comprennent la couverture nuageuse, le changement côtier, les mosaïques, la "couverture fractionnelle" (*Fractional Cover FC*), l'écart de NDVI (*NDVI Anomaly NDvA*), les risques de glissement de terrain, l'urbanisation et des analyses des eaux. Il est important de souligner que ce travail a été effectué au début des tests de fonctionnalités du SDC. Cela a



**Figure 3** - Les tuiles recouvrant le territoire Suisse (Giuliani, Peduzzi, et al. 2017).

impliqué des mises à jour du système et des modifications sur le serveur local durant le stage et donc des changements dans les outils à disposition, dans l'interface et de la qualité des images (CEOS, GRID Genève 2017).

## 5. OBJECTIFS ET DÉROULEMENT DU STAGE

---

Le stage s'est déroulé en trois parties. La première a tout d'abord consisté à tester les fonctionnalités, possibilités et limites de l'interface finalisée du SDC ([sdc.unepgrid.ch](http://sdc.unepgrid.ch)), afin de montrer ses potentialités et de donner des conseils pour son évolution. Pour ce faire, des animations et rendus graphiques sur des changements environnementaux connus ont été établis. Les observations se sont concentrées sur i) le retrait des glaciers ii) l'extension des zones urbaines et iii) l'évolution du couvert neigeux (Badoux and Rais 2015, WSL 2017, OFEV, Changements climatiques en Suisse 2007, OFEV, Office fédéral de météorologie et de climatologie 2017). L'outil (ou algorithme) employé pour ce premier objectif est la création de mosaïques (*Custom Mosaic, CM*) en vraies et fausses couleurs selon des séries temporelles définies.

Durant la deuxième partie, deux autres outils ont été testés. Le premier est le *Fractional Cover*, qui est un modèle en télédétection où la réflectance de pixels est considérée comme une combinaison linéaire d'aire proportionnelle de trois types de couverture : les pixels sont classifiés comme pourcentage de végétation photosynthétiquement active – photosynthétiquement inactive (comme la litière, les feuilles mortes ou les branches) et de sols nus (dont les roches). Le produit final est ainsi formé de trois bandes. Cet indice a pour but d'identifier les cycles et les tendances sur le long terme de la végétation et d'effectuer le suivi d'écosystèmes (p.ex. un pâturage) (Scarth, Röder and Schmidt 2011). Le second algorithme est l'écart de NDVI (Normalized Difference Vegetation Index). Le NDVI combine de manière arithmétique les canaux PIR et Rouge:  $NDVI = (PIR - ROUGE) / (PIR + ROUGE)$  (Allenbach and Dao 2017, Weier and Herring 2000). Les valeurs du NDVI varient de +1 à -1. Plus l'indice est positif, plus la végétation est en bonne santé ou dense. A l'inverse, des valeurs proches de 0 ou négatives indiquent une végétation clairsemée ou en mauvaise santé. L'écart de NDVI indique quant à lui la variation d'une scène unique par rapport aux enregistrements historiques (scènes à partir desquelles est calculée une médiane) (CEOS and GRID Genève, The CEOS Open Data Cube Initiative 2016). Une valeur positive indique une amélioration des conditions de végétation par rapport à la médiane, tandis qu'une valeur négative signale des conditions de végétation relativement mauvaises. Cet indice mesure donc un état relatif de la végétation en effectuant son suivi, relevant des zones où la végétation peut avoir subi un stress comme la détection d'éventuelles sécheresses (Organisation des Nations Unies pour l'alimentation et l'agriculture, FAO 2017).

Ces deux algorithmes permettent donc de mesurer les changements environnementaux et d'effectuer un suivi des écosystèmes (changements en qualité, en quantité et en type de végétation, comparaisons de milieux dans le temps et l'espace). L'objectif a été de savoir s'il est possible d'utiliser ces variables à ce dessein (Scarth, Röder and Schmidt 2011, Hamunyela, et al. 2016).

Pour terminer, un dernier objectif a été établi, celui de s'inspirer de scripts établis au travers de notebooks Jupyter (IPython 2017, CEOS, Australian Government, CSIRO, USGS 2017, NCI 2017) afin de produire des scripts pour interagir avec le SDC. Il s'agit de produire de manière automatisée les images et vidéos des différents outils vus précédemment.

## 6. MÉTHODOLOGIE

### 6.1. Utilisation et évolution de la plateforme sdc.unepgrid.ch

#### 6.1.1. Custom Mosaic

##### Evolution du couvert neigeux, des glaciers et des zones urbaines

Les mosaïques créent des images en vraies ou fausses couleurs de régions choisies en utilisant des données optiques provenant de plusieurs satellites Landsat. Des scènes sont ainsi assemblées pour former une mosaïque. L'utilité des mosaïques réside dans le fait de pouvoir gérer des informations à des résolutions diverses, c'est-à-dire spectrale, spatiale, temporelle et radiométrique (ESRI 2016, Hebert and Rougé 1984, CEOS, GRID Genève 2017).

La manipulation de l'interface est simple (Figure 4). La méthode pour effectuer les requêtes est la même pour les trois outils testés (CM, FC, NDvA). La description est faite ci-dessous avec les mosaïques.

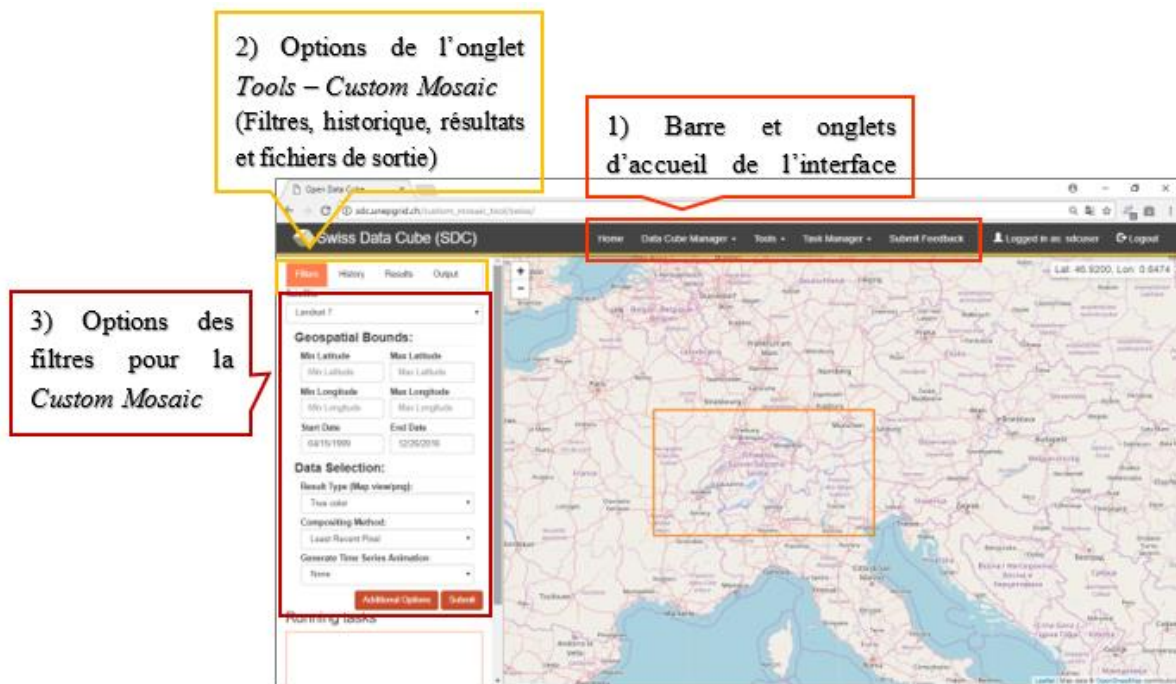


Figure 4 - Interface du Swiss Data Cube.

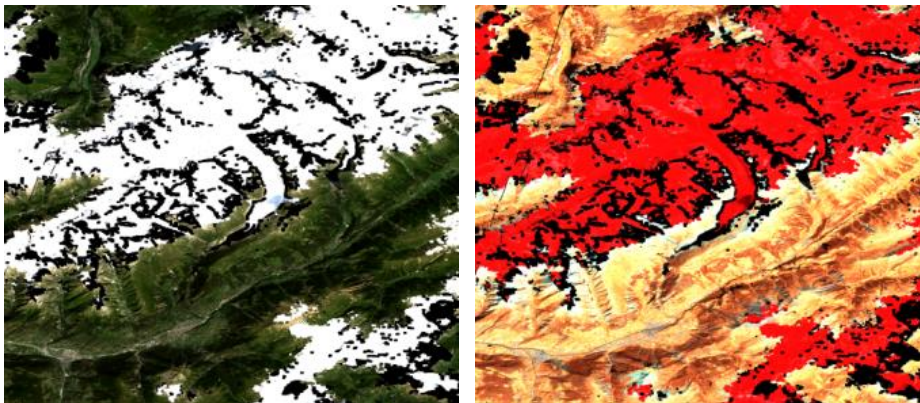
Une fois l'outil *Custom Mosaic* choisi sur l'interface (Figure 4 (1,2) ), il faut sélectionner les différents filtres (Figure 4 (3) ). Ceux-ci comprennent : (i) le choix du satellite (ii) les limites géospatiales (latitudes, longitudes et la date de début et de fin) (iii) la sélection des données (*type de résultats* (quelles bandes); *méthode de composition* (sélection des pixels dans les séries temporelles selon un calcul ou un ordre particulier comme le pixel le moins récent, le plus récent, le pixel médian, etc.) et (iv) la génération d'animations de séries temporelles.

Après qu'une mosaïque a été calculée, un historique permet d'aller rechercher les requêtes et de les re-télécharger sur la carte. L'onglet *résultats* affiche toutes les informations (date de début et de fin du calcul; méthodes de composition, liste d'acquisition des scènes, etc.). Pour terminer,

les données en sortie (*output*) peuvent être téléchargées en différents formats (\*.pgn, \*.tif, \*.netCDF, sous forme d'animations) (Figure 4 (2) ).

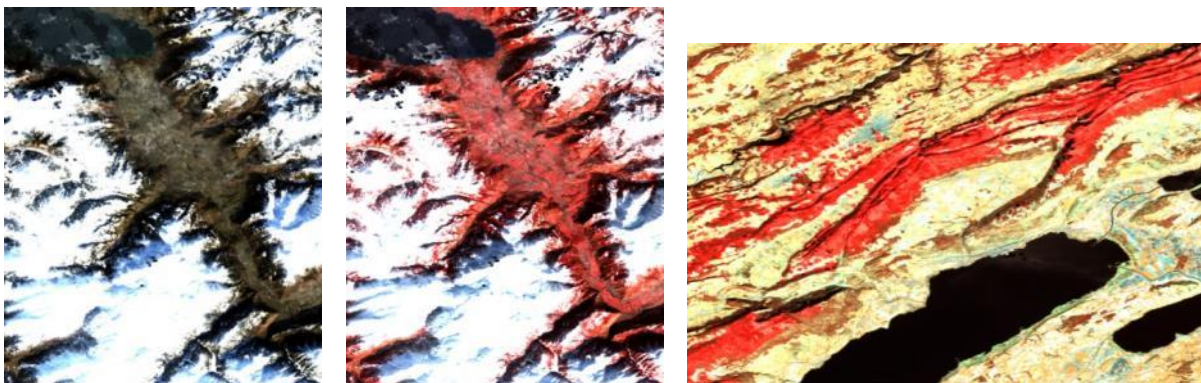
En ce qui concerne les zones d'étude, les glaciers de Ferpècle et du Mont Miné, du Rhône ainsi que celui d'Aletsch (Valais) ont été choisis. La région au nord du lac de Neuchâtel ainsi que la région des Portes du Soleil ont été sélectionnées pour l'évolution du couvert neigeux. Pour l'évolution des villes, Bâle, Bulle, La Chaux-de-Fonds, la région de Lausanne et de Cossonay (Vaud), Nyon ainsi que Genève ont été observées.

Des essais préliminaires ont été effectués avec les options des filtres à disposition. Les types de résultats ont été comparés entre eux pour voir quelles compositions de bandes faisaient le plus ressortir les phénomènes analysés. Pour les glaciers, les compositions en vraies couleurs (rouge-vert-bleu) et PIR, SWIR1, SWIR2 donnent les meilleurs rendus visuels (Figure 5).



**Figure 5** - Glacier d'Aletsch, mai à août 2013, Landsat 8. A gauche, en vraies couleurs (RGB). A droite, en fausses couleurs PIR, SWIR1, SWIR 2. Les *no data* apparaissent en noir.

Concernant la couverture neigeuse, les vraies couleurs, les bandes PIR, rouge, vert ou PIR, SWIR 1, SWIR 2 permettent de faire ressortir les différences entre les couvertures du sol (Figure 6).



**Figure 6** - (a) (b) Région des Portes du Soleil, hiver 2000-2001, Landsat 7. A gauche, en vraies couleurs. A droite, en fausses couleurs PIR, rouge, vert. (c) Couverture neigeuse, région au nord du lac de Neuchâtel, hiver 1999-2000, Landsat 7. Fausses couleurs PIR, SWIR 1, SWIR 2.

Pour les zones urbaines, les bandes PIR, rouge, vert ont été produites en plus des images en vraies couleurs.

Concernant les méthodes de composition, des comparaisons entre le pixel le moins récent, le plus récent et le pixel médian ont été faites, ainsi que quelques tests avec les NDVI minimum et maximum.

Les trois satellites Landsat ont été utilisés avec toutes les années à disposition (de 1984 à 2016) pour obtenir les meilleurs rendus et comparer les différents satellites.

Des GIF, vidéos et images ont été produits puis comparés entre les images en vraies couleurs et en fausses couleurs. Les GIF et les vidéos ont été élaborés à l'aide des sites <http://gifmaker.me/> et <https://ezgif.com/> (GIFMaker.Me 2017, Ezgif 2017). Le logiciel PhotoFiltre 7 (Bulot and Da Cruz 2005) a été employé pour ajouter les dates aux images.

## 6.2. Suivi de la végétation

### 6.2.1. Fractional Cover

Différents lieux ont été étudiés : le bassin genevois, le parc Jura Vaudois, la région du lac des Quatre Cantons, la région des Portes du Soleil (Valais) et des zones agricoles dans le canton de Vaud. Ces zones ont différentes utilisations et occupations du sol, comme des terrains agricoles, des forêts et des roches. La période de temps recouvre le printemps/l'été et l'automne afin de comparer la végétation vivante à la végétation en sénescence et à la végétation morte. Le nombre de mois entre chacune des requêtes n'a pas été le même à chaque fois. En raison de la quantité de données, il était parfois nécessaire de prendre des *time range* plus grands. La méthode de composition est celle du pixel médian.

### 6.2.2. NDVI Anomaly

La manière dont le SDC calcule le NDvA n'a pas pu être comprise de manière certaine. Comme dit précédemment, il calcule l'écart de NDVI en insérant une période de temps, par exemple du 1<sup>er</sup> au 30 juin 2014 (il faut inscrire une période et non une date précise sinon le calcul ne fonctionne pas) comme *scene selected* ou *single scene*. Le SDC semble prendre automatiquement la scène la plus récente à disposition (ici le 30 juin 2014). Une première composite (plusieurs scènes) calculant le NDVI semble être faite sur cette période de temps. Puis il faut choisir la *baseline*, qui est une seconde composite formée de scènes précédant la scène unique (Tableau 1). C'est le NDVI médian qui est calculé sur cette composite. Si nous sélectionnons les mois de juin, juillet et août avec le Landsat 8, le calcul se fera en prenant toutes les scènes de ces mois pour les années 2013 et 2014 (jusqu'au 30 juin 2014). L'écart de NDVI calcule l'évolution entre la *baseline* et la scène unique.

**Tableau 1** - Exemples de scènes uniques et des *baseline* sélectionnées comme filtre pour le calcul du NDVI Anomaly.

Single scene	Baseline
30 juin 2014	Avril, Mai
31 Août 2014	Juin, Juillet, Août
30 octobre 2014	Septembre, Octobre

Les régions d'étude sont : la région agricole du canton de Vaud (axe Saint-Prex – Aubonne – Bière jusqu'à la limite du parc Jura vaudois) et une zone forestière couvrant une partie du parc Jura Vaudois depuis Saint-Cergue et une partie de la France. Les périodes de temps ont été définies en fonction de la qualité des images disponibles et pour suivre l'exemple présenté dans le webinar établi par CEOS datant du 7 juillet 2017 (CEOS, Data Cubes for Large Scale Data Analytics 2017).

Les auteurs ont porté leur choix sur une scène en avril 2016 et ils l'ont comparée à un NDVI médian pour le même mois des années 2013 à 2016. La différence dans notre calcul est qu'il a été nécessaire de prendre plusieurs mois pour établir la *baseline*, le manque de données ne permettant pas toujours le calcul de l'écart de NDVI, ou un message d'erreur s'affichait, indiquant que le calcul n'était pas possible sur la période de temps inscrite.

### 6.3. Tests des scripts établis à l'aide du notebook Jupyter

Une fois les notebooks mis à jour pour le SDC, certains des codes ont été modifiés avec les données du SDC afin de voir s'il était possible de produire des images. Le langage des scripts est Python 3. Les premiers scripts ainsi testés sont les suivants :

*Data\_Cube\_API\_Demo-CopyLB.py* (Annexe A1) et le script via le lien [http://nbviewer.jupyter.org/github/data-cube/agdc-v2/blob/develop/examples/notebooks/Datacube\\_Summary.ipynb](http://nbviewer.jupyter.org/github/data-cube/agdc-v2/blob/develop/examples/notebooks/Datacube_Summary.ipynb)

Le script *Data\_Cube\_API\_Demo-CopyLB.py* est une démonstration pour l'utilisation de l'API (Application Programming Interface (Geoscience Australia 2017)), présentant les fonctions basiques du SDC, des productions de scènes simples (une seule date) et d'histogrammes des différentes bandes. Les métadonnées des scènes ainsi que la liste des acquisitions des scènes peuvent également être obtenues.

Le script *../Datacube\_Summary.ipynb*, établi par l'équipe de recherche australienne (CEOS, Australian Government, CSIRO, USGS 2017), présente les caractéristiques de mise en place de modules du DC afin de démontrer ses fonctionnalités. Il récupère toutes les informations d'un satellite Landsat et produit des graphes en fonction d'une série temporelle (par exemple une saison). Des fonctionnalités comme masquer les *no data*, les nuages ou encore le calcul du NDVI y sont présentées.

Une fois la prise en main de ces premiers scripts effectuée, d'autres scripts ont été testés. Par exemple le script *kmean\_clustering\_exemples.py* peut directement enregistrer sur le server des mosaïques en format TIF.

A partir de ces différents scripts, les scripts propres au SDC ont été élaborés. Le client utilisé pour accéder au server (machine distante) est FileZilla (Kosse 2017). Au travers d'une interface graphique, FileZilla permet de gérer les fichiers de la machine et de les éditer. Le client SSH PuTTY (Tatham 2017) a été employé quant à lui pour accéder, en lignes de commande, aux fichiers de la machine distante (exécution des fichiers édités par FileZilla).

Une dizaine d'essais de scripts ont été effectués afin d'expérimenter les différentes fonctions proposées par CEOS et les Australiens afin de tendre vers les meilleurs outputs finaux.

## 7. RÉSULTATS

---

### 7.1. Utilisation et évolution de la plateforme `sdc.unepgrid.ch`

En raison des mises à jour, de la gestion constante (nécessitant du temps et ralentissant les tests) et à l'instabilité de la plateforme, il a été difficile de maintenir des commentaires actuels sur son utilisation. Par exemple, un commentaire étant valable avant la mise à jour ne l'était plus après. De ce fait, la plupart des commentaires portant sur l'amélioration de la plateforme ne sont plus d'actualité. Les problèmes persistants sont abordés dans la partie discussion.

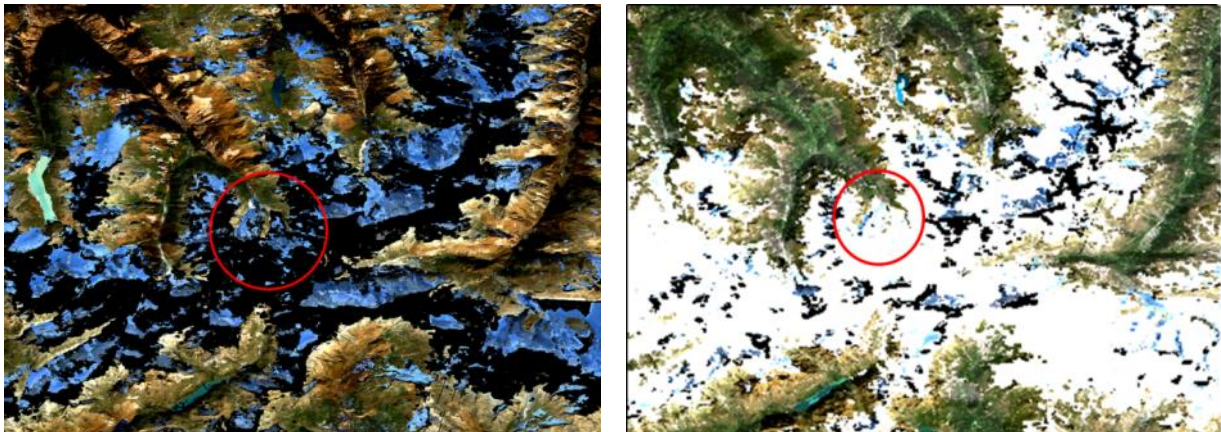
Toutes les images ainsi que les vidéos se trouvent sur l'owncloud du GRID.

#### 7.1.1. Custom Mosaic

##### Evolution des glaciers

Les périodes de temps conservées sont les mois durant lesquels la neige est le moins présente (de mai à septembre) afin d'éviter la couverture neigeuse qui cache de l'information (les glaciers ne peuvent apparaître correctement lorsqu'ils sont sous la neige, ainsi que les poussières recouvrant la neige qui peuvent poser des problèmes pour la réflectance).

Pour les glaciers de Ferpècle et du Mont Miné, des images ont été produites depuis 1987 jusqu'en 2014 (Landsat 5, 7, 8) avec comparaison entre vraies et fausses couleurs. Une vidéo a été effectuée sur les images de 2000 à 2005 avec la méthode du pixel le plus récent. La majorité des images a été sélectionnée en été mais une comparaison avec la période hivernale a été effectuée pour analyser la manière dont le SDC affiche la couverture neigeuse. Ces premiers tests ont décelé un problème avec le masque de la couverture de neige et de glace; ces deux surfaces n'apparaissaient pas ou très peu (Figure 7). Ce problème a pu être rapidement corrigé.

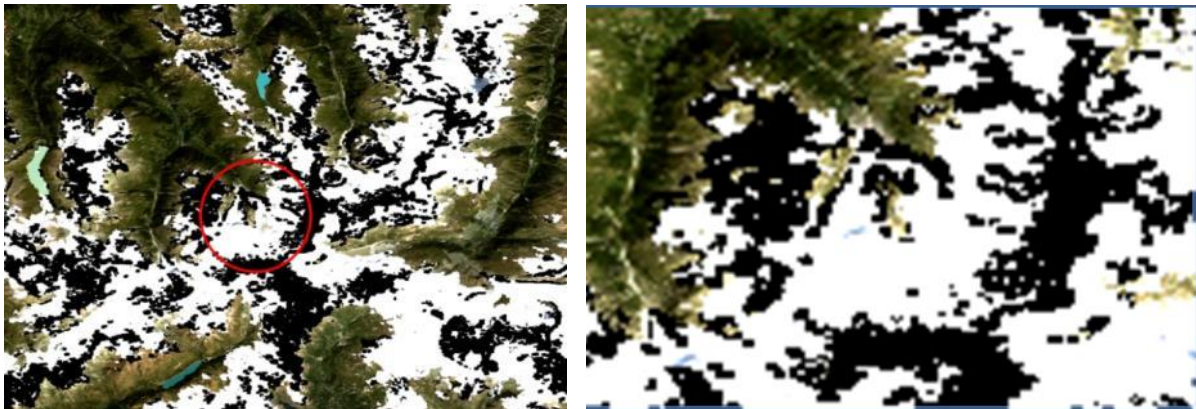


**Figure 7** - Glaciers de Ferpècle et du Mont Miné (cercle rouge), Landsat 8. A gauche, problème du `cf_mask`, mai 2013 à décembre 2016, vraies couleurs. La neige et la glace apparaissent partiellement en bleu, les `no data` en noir. A droite, problème résolu, pour la même période. Les `no data` apparaissent en transparent ici, un fond noir a été ajouté pour faire ressortir la neige (images produites à des dates différentes).

La période de temps sélectionnée recouvre plusieurs années, car lors des premiers tests il n'a pas suffi de conserver uniquement quelques mois sur une seule année pour faire apparaître la neige et la glace.

Malgré cette correction, on remarque que la pixellisation est encore grossière à certains endroits, qu'il y a beaucoup de `no data` et que le relief n'apparaît pas suffisamment clairement et ceci peu importe le Landsat utilisé. Mais on remarque surtout que le glacier i) n'apparaît pas correctement (même quand il n'y a pas de `no data`) ii) qu'il recouvre une surface

insuffisamment importante pour être détecté (un recadrage sur les glaciers ne permet évidemment pas d'obtenir un meilleur rendu étant donné la résolution du Landsat). (Figure 7, Figure 8).



**Figure 8** - Glaciers de Ferpèche et du Mont Miné, Landsat 7, mai à août 2002, vraies couleurs. A gauche (cercle rouge), la neige et la glace apparaissent partiellement en blanc, les *no data* en noir. A droite, zoom dans le cercle rouge. Les *no data* apparaissent en transparent ici, un fond noir a été ajouté pour faire ressortir la neige (images produites à des dates différentes dans le SDC).

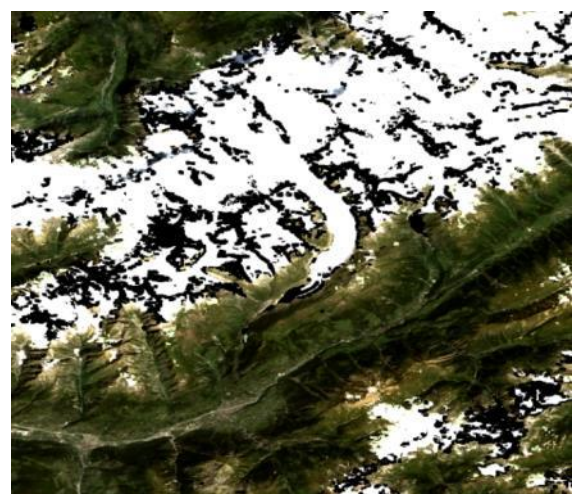


Pour le glacier du Rhône (Figure 9), les données recouvraient la période 1984 - 2016. La méthode était le pixel médian et en vraies couleurs. Des composites avec les Landsat 7 et 8 ainsi qu'une vidéo des mosaïques de 2010 à 2016 ont été produites. Les mêmes remarques que pour le glacier de Ferpèche sont faites quant à la qualité des images.

**Figure 9** - Glacier du Rhône (cercle rouge), Landsat 7, mai à août 2001, vraies couleurs.

Afin de comparer les différentes méthodes de composition, les tests pour le glacier d'Aletsch ont été effectués avec le pixel le plus récent, le moins récent et avec le pixel médian. Une vidéo recouvrant le période de 1984 à 2016 a été créée.

Plus de détails apparaissent (Figure 10) mais il persiste les mêmes problèmes que pour les deux autres glaciers. La différence est que la surface du glacier d'Aletsch est plus grande, donc plus visible que les deux autres. A noter également que seule une partie du glacier est représentée, la totalité des zones d'influences du glacier ne sont pas présentes (Wikipédia 2017, Schweiz Tourismus 2017).



**Figure 10** - Glacier d'Aletsch, Landsat 7, mai à août 2001, vraies couleurs.

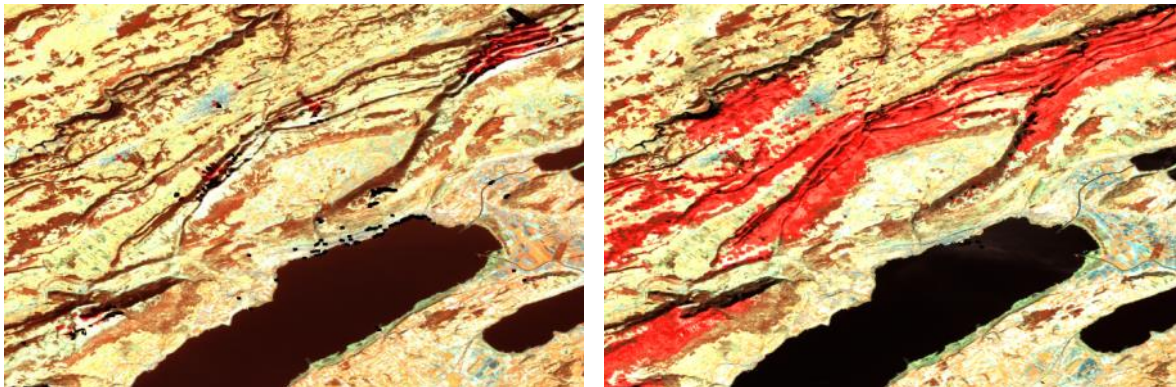


## Evolution du couvert neigeux

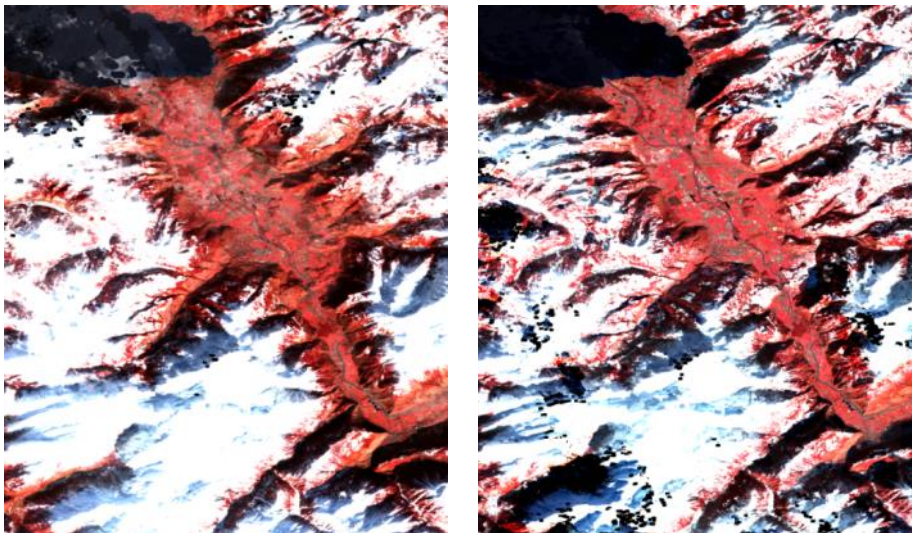
Pour Neuchâtel, des images ainsi qu'une vidéo de la période 1989 - 2015 avec les bandes PIR, SWIR 1, SWIR 2 ont été produites.

Les images de la région des Portes du Soleil ont été testées en vraies et fausses couleurs pour les années 1999 à 2016. La méthode du pixel médian a été conservée.

Sur la Figure 11, le couvert neigeux est bien visible avec les bandes sélectionnées. On peut observer des années avec beaucoup de neige face à celles où les précipitations sont moindres. On décèle aussi bien les différentes couvertures et utilisations du sol (villes, zones agricoles, eau, végétation) et, en conséquence, les zones où la neige persiste.



**Figure 11** - Evolution du couvert neigeux avec les bandes PIR, SWIR 1, SWIR 2. Région du Canton de Neuchâtel. A gauche, Landsat 5, hiver 1989-1990. A droite, Landsat 7, hiver 1999-2000. La neige apparaît en rouge, la végétation en beige, les *no data* en noir, les zones agricoles en orange et bleu, le tissu urbain en bleu-gris.



**Figure 12** - Evolution du couvert neigeux avec les bandes PIR, rouge, vert. Région des Portes du Soleil avec le tissu urbain de Villeneuve à Martigny (Valais). A gauche, Landsat 7, hiver 2000 – 2001. A droite, Landsat 8, hiver 2015 – 2016.

La Figure 12 montre une différence moins marquée entre les deux images pour la région des Portes du Soleil. La neige est moins présente en hiver 2015-2016. Il y a des différences de qualité des images dues à la différence de Landsat et à la position du soleil (ombres plus marquées sur l'image de droite).

Les vidéos de l'évolution du couvert neigeux pour la région de Neuchâtel et des Portes du Soleil montrent bien des différences de précipitations d'un hiver à l'autre. Par exemple, pour Neuchâtel, les hivers 1989-1990 et 2000-2001 ne montrent que très peu, voire aucune précipitation dans la majorité de la région. A l'inverse, les hivers 2002-2003, 2004-2005 et

2008-2009 montrent de fortes précipitations neigeuses. Le couvert neigeux en 2008-2009 semble avoir été particulièrement grand. Pour les Portes du Soleil, les fortes précipitations sont confirmées pour les hivers 2004-2005 et 2008-2009. Selon la date de prise de vue des scènes, le couvert neigeux se retrouve également dans les zones urbaines.

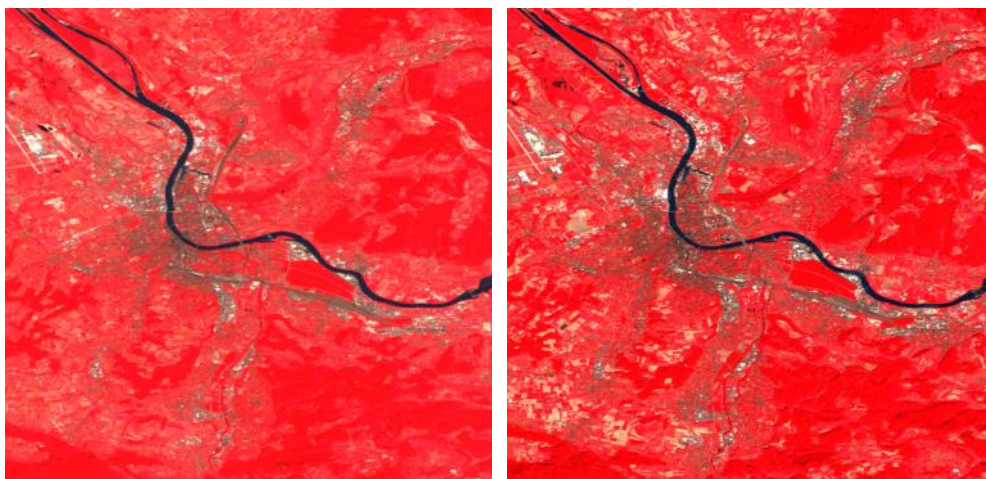
### **Zones urbaines**

Les périodes de temps sélectionnées concernent également les mois où la neige est la moins présente (généralement de mai à septembre, sauf pour certains endroits où il a été possible de faire des mosaïques sur une année sans que le résultat soit affecté par la couverture neigeuse).

Il faut préciser que les premiers tests ont été effectués avant la mise à jour du 14 août 2017. Après cette date, un outil "urbanisation" a été mis à disposition sur la plateforme finale, qui augmente l'exactitude et diminue le temps de traitement des requêtes. Le rendu est une mosaïque dont les couleurs sont définies par le NDBI (urbanisation), le NVI (végétation) et le NDWI (eau).

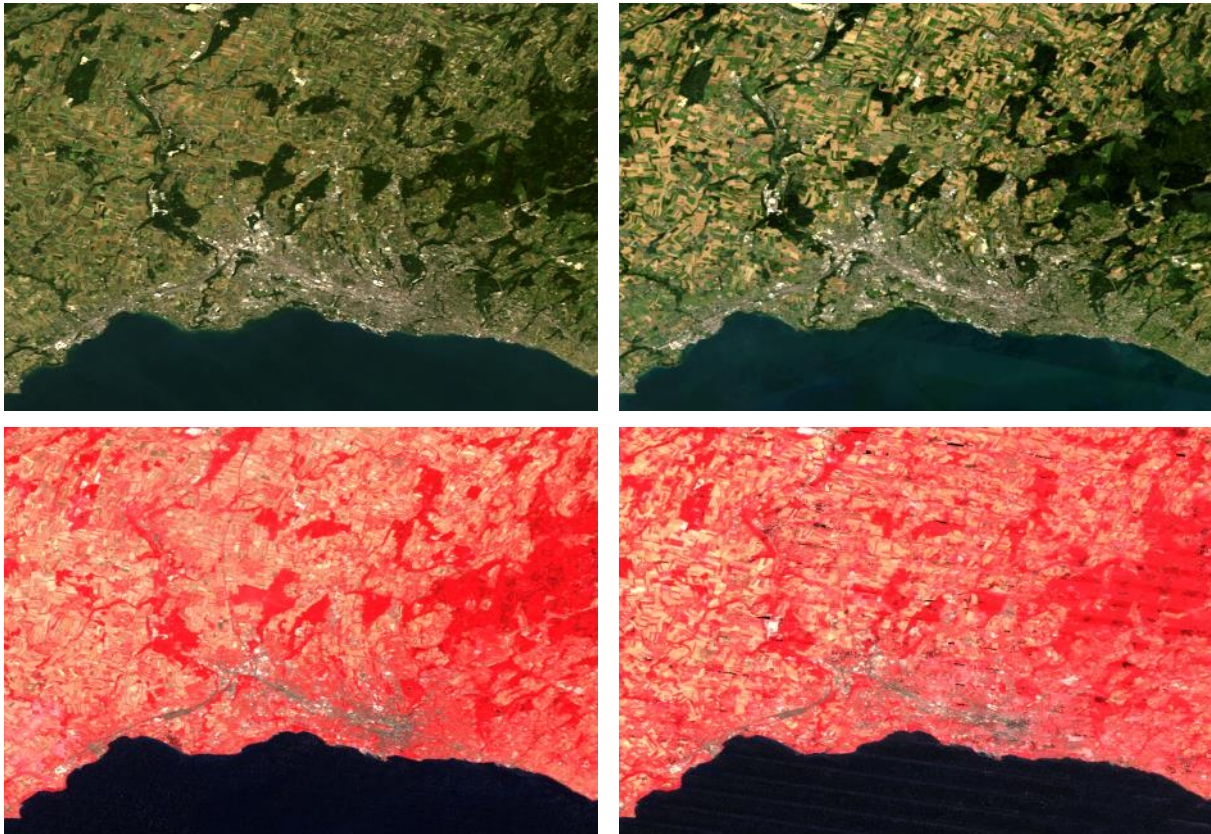
En raison d'un problème avec la plateforme (problème aperçu à partir du 09.10.2017), l'outil *Custom Mosaic* a rendu des images floues. Seules les images établies jusqu'alors ont pu être utilisées pour ce rapport, c'est pourquoi certaines dates pour analyser l'évolution ne coïncident pas forcément entre les différentes zones (p.ex. une analyse entre 2000 et 2016 pour certaines bandes et entre 2000 et 2005 pour d'autres bandes, car les tests de 2016 n'ont pas pu être effectués après coup). De surcroît, la taille des images lors du téléchargement en format PGN a également grandement diminué, rendant souvent impossible une visualisation suffisante.

Il s'est avéré difficile de bien distinguer le tissu urbain avec les vraies couleurs. Les meilleures bandes trouvées pour cette analyse sont les PIR, rouge et vert. La méthode du pixel médian est toujours utilisée, mais les tests avec les méthodes de NDVI maximum et minimum ont permis de faire ressortir les différences entre types de végétation et du même coup, de les différencier au mieux du tissu urbain.



**Figure 13** - Bâle, bandes PIR, rouge, vert. A gauche, Landsat 7, en 2000. A droite, Landsat 8, en 2016.

La ville de Bâle (Figure 13) montre une légère augmentation du tissu urbain. La végétation apparaît en rouge (plus dense comme la forêt) et orange (cultures), mais il est difficile de distinguer les différences sur ces images.



**Figure 14** - Lausanne - Cossonay. *En haut*: vraies couleurs. La luminosité a été augmentée de 20%. A gauche, Landsat 7, mai à août 2000. A droite, Landsat 8, mai à août 2016. *En bas*: bandes PIR, rouge, verte, min. NDVI. A gauche, Landsat 7, mai à août 2000. A droite, Landsat 7, mai à août 2005.

Dans la région de Lausanne (Figure 14), il est difficile de distinguer des changements avec cette résolution. On peut comparer les rendus entre les vraies et les fausses couleurs. Les différences de types de végétation sont plus nettes avec la méthode du NDVI minimum.



**Figure 15** - Bulle (Fribourg). La luminosité a été augmentée de 20% pour les deux images. A gauche, Landsat 5, du 1<sup>er</sup> janvier 1984 au 1<sup>er</sup> janvier 1985. A droite, de mai à août 2013.

La ville de Bulle (Figure 15) semble connaître un léger essor, mais à nouveau sur une surface trop faible pour être bien visible. Des images avec les bandes PIR, rouge, vert n'ont pas été possibles de par le problème cité précédemment.

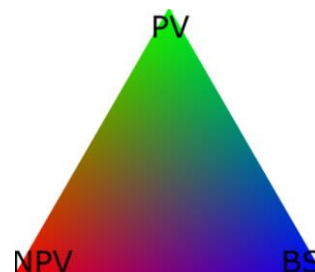
## 7.2. Suivi de la végétation

Pour les mêmes raisons énoncées à la page 13 sur l'instabilité du SDC, il a été décidé de se concentrer sur la réponse des deux algorithmes et de passer moins de temps sur leur lien possible avec les changements environnementaux ou d'utilisation du sol.

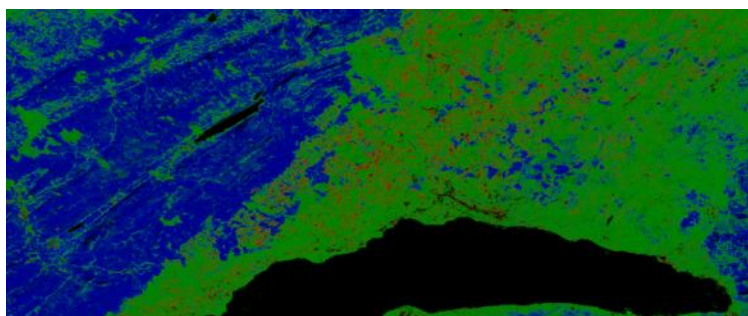
### 7.2.1. Fractional Cover

La Figure 16 illustre la gamme de couleurs des pixels classifiés utilisée pour visualiser le *Fractional Cover* (CEOS, GRID Genève 2017).

De manière générale le rendu est relativement précis. Il subsiste toutefois quelques doutes quant à certaines zones classées en sols nus (BS), comme pour la région du lac des Quatre Cantons ou le parc naturel Jura vaudois (Figure 17), où des zones de forêts sont classées comme telles (comparaison avec Google Earth (Google 2017)).



**Figure 16** - Gamme de couleurs pour le Fractional Cover. PV: photosynthétique  
- NPV: Non photosynthétique  
- BS : sols nus.



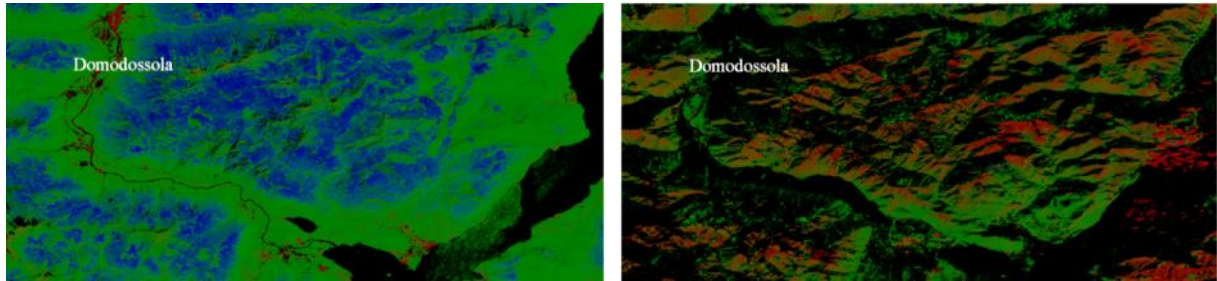
**Figure 17** - Parc Jura vaudois et une partie de la zone agricole dans le Canton de Vaud, Landsat 8, mai à août 2016. On remarque une très forte proportion en sols nus (en bleu) alors que cette zone est aussi composée de forêts. Les *no data* apparaissent en noir. La luminosité a été augmentée de 20%.

Certains *Fractional Cover* se sont modifiés entre 1984 et 2016 alors que d'autres ont peu, voire pas, subi de changement. Par comparaison, des tests effectués sur le Vietnam ont montré un grand nombre de différences flagrantes dues aux extrêmes climatiques (les moussons par exemple) et aux zones biogéographiques très diverses d'une saison ou d'une année à l'autre.

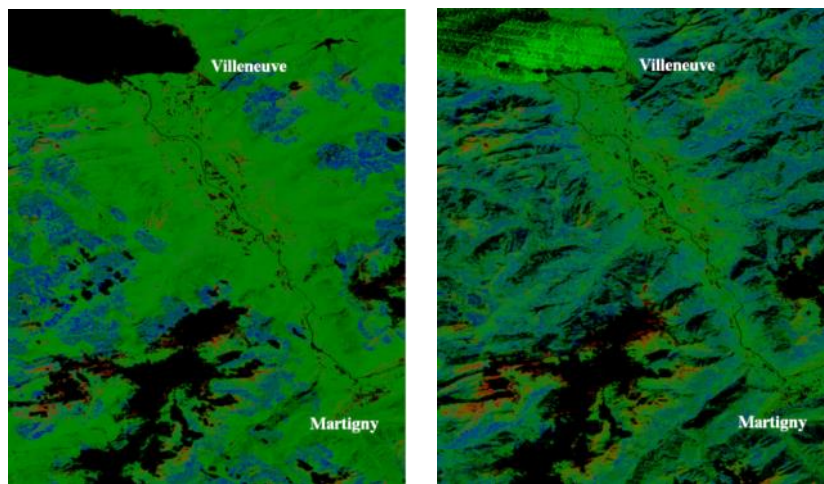
Les différences subsistent bien entendu en fonction des couvertures et de l'utilisation du sol, ainsi que pour les saisons. Ainsi, certaines parcelles agricoles sont classées en végétation non photosynthétique selon la période de récolte et la saison. Pour les forêts ou zones boisées, la différence entre printemps/été et automne/hiver est bien visible (Figure 18, Figure 19). Pour l'automne, on constate une plus forte proportion de sols nus (BS) et/ou de végétation non photosynthétique (NPV). Cela dépend aussi bien sûr du type de végétation (feuillue ou résineuse). Pour la partie tessinoise, une observation approfondie montre qu'en réalité la plupart des zones passant en NPV en automne sont des zones qui étaient classées en PV au printemps. Les zones où le sol nu passe en NPV pourraient s'expliquer par un déplacement des feuilles mortes par le vent ou éventuellement une végétation qui a poussé rapidement puis qui a séché.

Aux figures 19 et 20 on peut percevoir des zones classées PV et plus tard, pour la même période, classées en BS.

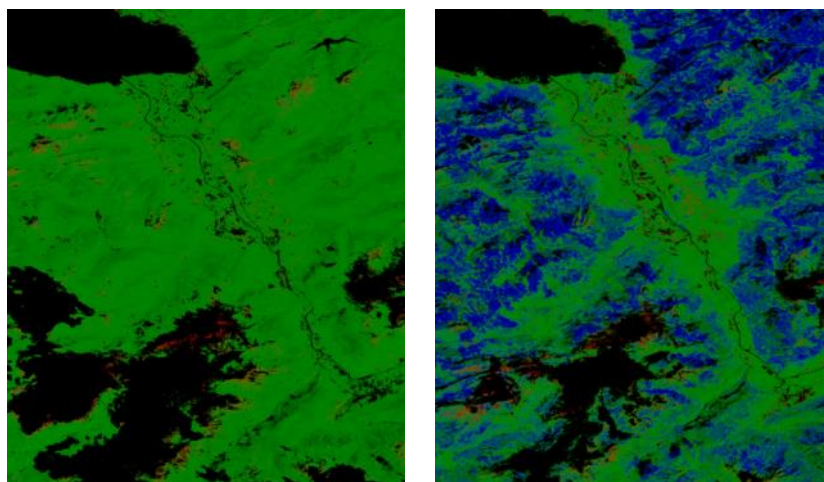
La couverture neigeuse ainsi que l'eau se transforment en *no data* (en noir) (Figure 17, Figure 18, Figure 19, Figure 20, Figure 21).



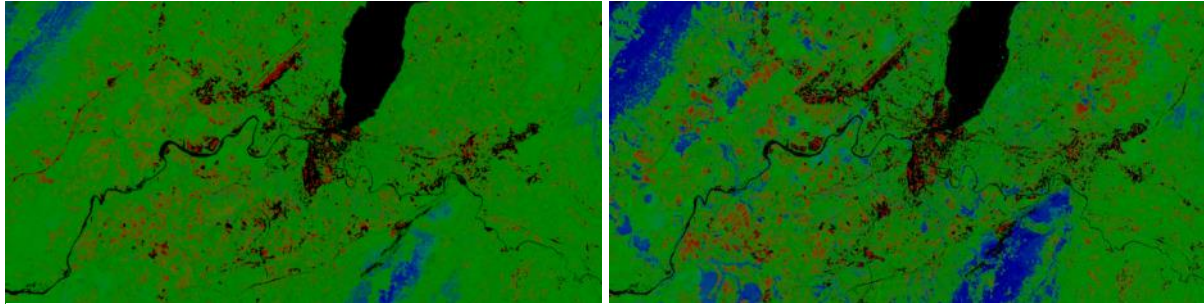
**Figure 18** - Partie du Tessin. A gauche, de mai à août 2001. A droite, de novembre à décembre 2001. Landsat 7, méthode du pixel médian. La luminosité a été augmentée de 20% pour les deux images



**Figure 19** - Région des Portes du Soleil. A gauche, de mai à août 2010. A droite, de septembre à novembre 2010. Landsat 5, méthode du pixel médian. La luminosité a été augmentée de 20% pour les deux images.



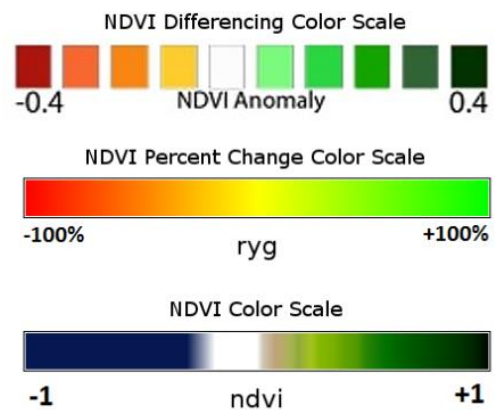
**Figure 20** - Région des Portes du Soleil. A gauche, Landsat 5, de mai à août 1987. A droite, Landsat 8, de mai à août 2016. Méthode du pixel médian. La luminosité a été augmentée de 20% pour les deux images.



**Figure 21** - Bassin genevois. A gauche, Landsat 7, mai à août 2000. A droite, Landsat 8, mai à août 2016. Méthode du pixel médian. La luminosité a été augmentée de 20% pour les deux images.

### 7.2.2. NDVI Anomaly

Le résultat visuel se trouve sous quatre formes : la première est l'écart (ou différence) de NDVI, la seconde est le changement de pourcentage de NDVI, la troisième la scène unique et la dernière la composite servant de référence (*baseline*). Les échelles de couleurs associées à ces différentes formes sont mises à disposition sur la plateforme (Figure 22). Il est ainsi possible, depuis l'onglet *Résultats*, de choisir quelle image montrer sur la plateforme et de voir les scènes à disposition (Figure 23). L'onglet *Output* permet, en plus du choix du format, de choisir laquelle de ces images peut être téléchargée. Par contre, ni dans l'onglet *Résultat* ni dans l'onglet *Output* il n'est écrit que le calcul se base sur une scène unique. Seul le *Task Manager* (gestionnaire des tâches), dans les *Task Parameters*, indique quelle scène unique a été sélectionnée (Figure 24).



**Figure 22** - Echelles des valeurs en couleurs pour (1) l'écart de NDVI (2) le pourcentage de changement du NDVI. Un pourcentage de changement de NDVI de +50% indique que la végétation s'est améliorée de +50% par rapport à la médiane. (3) Le NDVI.

### Completed Tasks

Latitude Range:	(46.4632, 46.5483)
Longitude Range:	(6.2687, 6.4824)
Time Range:	(06/01/2014, 06/30/2014)
Baseline Months:	April, May
Show NDVI Difference:	<input checked="" type="checkbox"/>
Show NDVI Percentage Change:	<input type="checkbox"/>

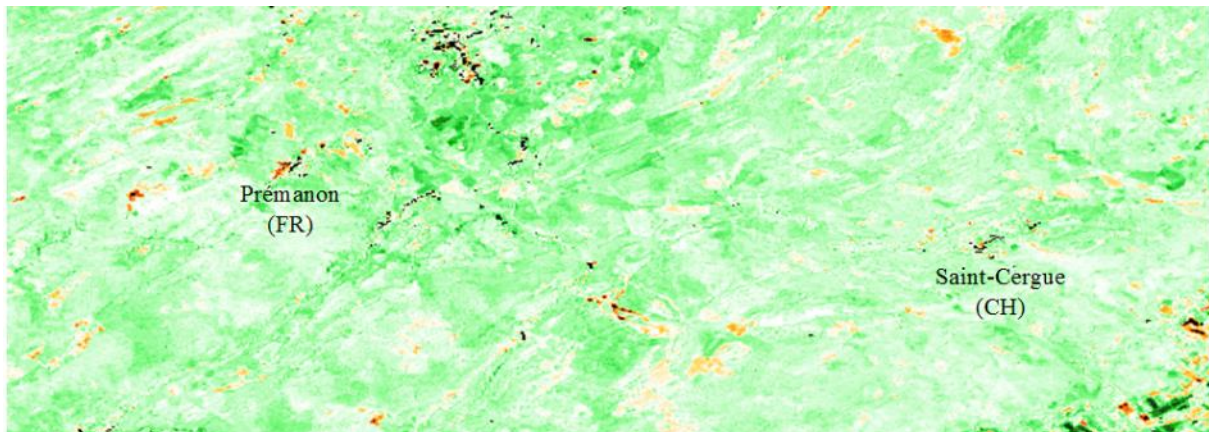
### Acquisitions List

<b>06/24/2014</b>	<b>Load this scene</b>
Clean pixels:	0
Total Pixels:	248220
Clean Pixel Percentage:	0.00%
<b>06/24/2014</b>	<b>Load this scene</b>
Clean pixels:	0
Total Pixels:	248220
Clean Pixel Percentage:	0.00%
<b>06/15/2014</b>	<b>Load this scene</b>
Clean pixels:	248872

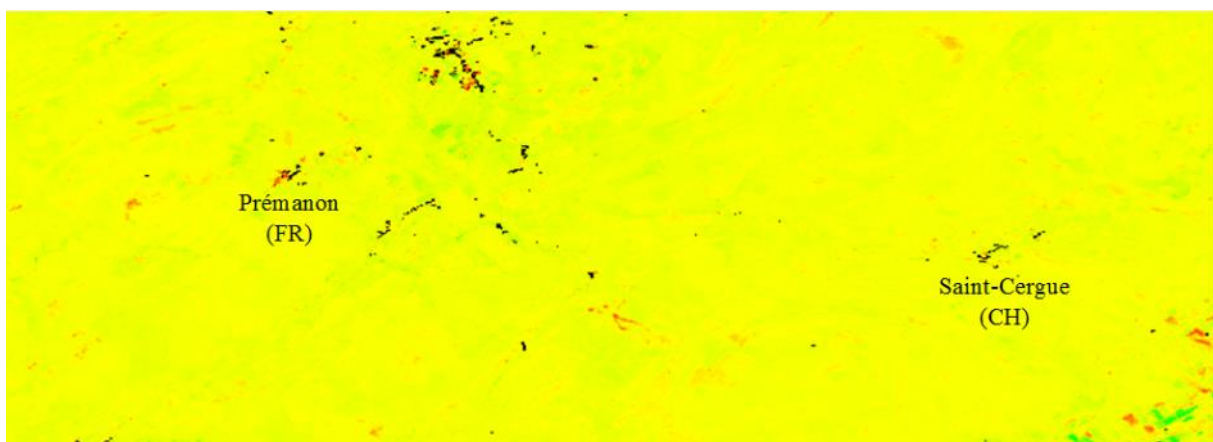
**Figure 23** - Extrait de l'onglet *Résultat* de la plateforme *sdc.unepgrid.ch* pour le NDVI Anomaly. A gauche, on peut voir la période de temps sélectionnée à partir de laquelle la scène unique sera choisie, les mois pour la *baseline* et les manières de représenter les résultats sur la carte (cases à cocher). A droite, la liste d'acquisition des scènes (pour la scène unique, ou le *Time Range*).

Task Details	
<b>Title</b>	NDVI Anomaly Query
<b>Description</b>	None
<b>Status</b>	Complete
<b>Start Time</b>	10/02/2017 03:58
<b>End Time</b>	10/02/2017 03:58
Task Parameters	
(Lat, Lon) Min	(46.463200, 6.268700)
(Lat, Lon) Max	(46.548300, 6.482400)
<b>Scenes Selected</b>	June 30, 2014
<b>Baseline Months</b>	April, May

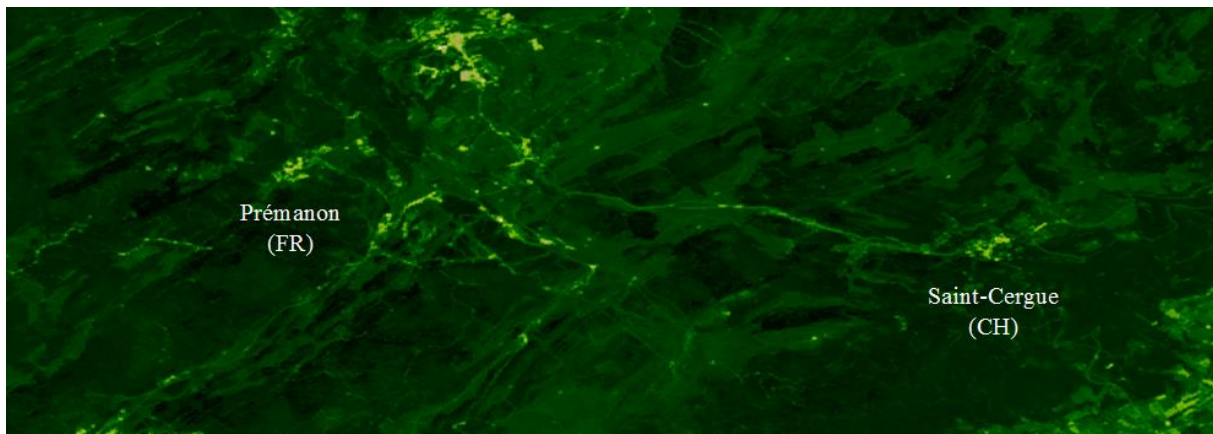
**Figure 24** - Gestionnaire des tâches, paramètres des tâches.



**Figure 25** - Ecart de NDVI. Landsat 8. Zone de forêt dans la région du parc Jura vaudois et traversant la frontière française. La scène de référence est le mois d'août 2016, avec une scène unique datant du 31. La *baseline* est calculée avec les mois de juin, juillet et août.



**Figure 26** - Pourcentage de changement de NDVI. Idem que la Figure 24.



**Figure 27** - *Baseline* pour le calcul de la médiane. Idem que la Figure 24.

La Figure 25 illustre un écart de NDVI majoritairement positif (+ 0.08). Certaines zones montrent un écart négatif (entre -0.08 et -0.16). Dans l'ensemble, la végétation est devenue plus dense et/ou en meilleure santé au 31 août 2016 comparé à la médiane des mois de juin, juillet, août entre 2013 et 2016 (le Landsat 8 commençant en 2013).

En pourcentage, le changement de NDVI oscille entre 0% et 50% (Figure 26). L'analyse du pourcentage de changement de NDVI n'est pas aisée, le choix des couleurs (jaune-vert proche



du fluorescent) n'étant pas des plus adéquats pour l'œil humain. La Figure 27 quant à elle indique un NDVI médian entre 0.6 et 1 pour la végétation la plus dense (forêt) et autour des 0.4 pour les zones plus clairsemées (bords des routes, villes, villages, ...).

### 7.3. Tests des scripts établis à l'aide du notebook Jupyter

Trois scripts ont été finalisés :

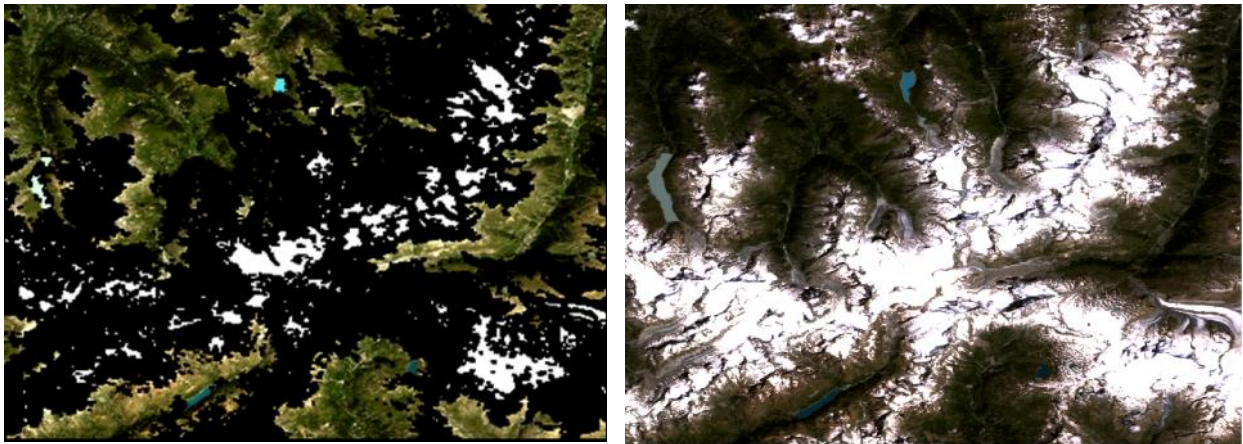
Les scripts *test\_scenes\_video\_LB.py* et *test\_scenes\_video\_2\_LB.py* (Annexe A2 et A3), inspirés d'une partie du notebook "*Datacube\_Summary*" des Australiens (cf lien de la page 11), produisent des scènes uniques qui sont ensuite mises ensemble pour produire des vidéos sur l'évolution spatio-temporelle des phénomènes étudiés (en vraies et fausses couleurs). L'unique différence entre les deux scripts est l'insertion ou non d'images blanches au début et à la fin des vidéos, les images blanches permettant d'amorcer et de terminer les transitions entre les images. L'automatisation de la production des scènes comprend (i) l'importation de modules, dont des spécifiques au DC (p.ex. *datacube.Datacube()* est une interface qui cherche, décrit et écrit un datacube), (ii) la définition d'arguments (permettant d'aller rechercher les informations des scènes Landsat en ligne de commande dans PuTTY) (module *argparse*) (iii) le chargement des données Landsat du Datacube (*dc.load*, comprenant le produit, les coordonnées géographiques, la période de temps, les mesures, etc.) (iv) la production des scènes Landsat (*clipped\_visible2* engendre un tableau de N dimensions comprenant des coordonnées et des dimensions labellisées. *Scaled2* recalcule les dimensions du tableau). La seconde partie du script (à partir d'*os.system ("rm \*.png")*) crée une chaîne (instruction *for a in*) afin de produire les scènes sous format PNG dans un même dossier. L'instruction alternative *if size* sélectionne les scènes contenant le plus d'informations (taille des images en bytes) afin d'éviter au maximum les zones de nuages ou sans données. La sélection de la taille se fait de manière manuelle, en ouvrant les fichiers PNG et notant la taille des images. Une vidéo rassemblant ces scènes (ligne de commande *ffmpeg*) est produite avec la dernière ligne du code. Dans l'exemple des Annexes A2 et A3, les scènes sont produites sur le glacier d'Aletsch en vraies couleurs (RGB).

Différents problèmes sont survenus aux différentes étapes de la prise en main et de l'automatisation. Il a fallu par exemple récupérer des informations dans les variables (p.ex. l'information *time* dans *clipped\_visible2* qui recherche la date de la prise de vue des scènes). Ainsi, des titres et légendes ont pu être rajoutés à chacune des images.

Pour le script *test\_kmean\_clustering\_LB.py* (Annexe A4) (créé à partir du notebook *kmean\_clustering.ipynb*), il a également fallu importer des fonctions et modules propres au DC, comme les fonctions utilitaires dans les modules d'espace de stockage via *utils*. La fonction *write\_geotiff\_from\_xr* enregistre la mosaïque en format TIF. La fonction *create\_geographic\_chunks* divise les données géographiques (latitude et longitude ici) du tableau en petits "morceaux" ("*chunks*") afin de pouvoir faire tenir les données dans la mémoire (CEOS, Australian Government, CSIRO, USGS 2017). Ces images sont ensuite ouvertes avec QGIS (OSGeo 2017). Dans l'exemple de l'Annexe A4, la mosaïque est faite sur le bassin genevois, avec le Landsat 7 et allant du 1<sup>e</sup> janvier 1999 au 31 janvier 2000.

Avec les scripts *test\_scenes\_video\_LB.py* et *test\_scenes\_video\_2\_LB.py*, des vidéos sur les glaciers, les zones urbaines et l'évolution de la couverture neigeuse ainsi que des scènes correspondantes ont été utilisées comme produits finaux pour ce stage.

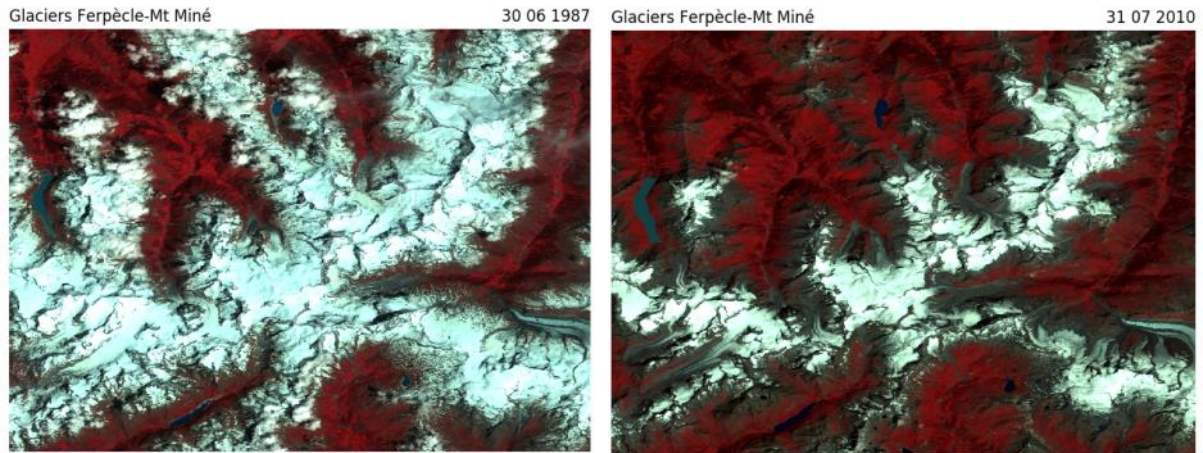
Il faut noter une remarque importante ici. En comparant les scènes ou les mosaïques obtenues à partir de la plateforme à celles produites avec le script, on observe une grande différence en termes de qualité des images. En date du 04.10.17, la Figure 28 présente les tests effectués et donne un exemple des différences.



**Figure 28** - Différence de rendu. Les deux scènes représentent la région des glaciers de Ferpèche et de Mont Miné, le 30 juin 2001 avec le Landsat 7. A gauche, la scène est téléchargée depuis le site [sdc.unepgrid.ch](http://sdc.unepgrid.ch). A droite, elle est produite via le script Python.

## Glaciers - scènes uniques en vraies et fausses couleurs

Seules des scènes uniques ont été testées pour les glaciers.



**Figure 29** - Exemples de rendu pour les scènes – glaciers de Ferpèche et du Mont Miné en 1987 et en 2010 (Landsat 5, bandes PIR, rouge, vert).



**Figure 30** - Exemples de rendu pour les scènes – glacier du Rhône en 1990 et en 2010 (Landsat 5, bandes PIR, rouge, vert).

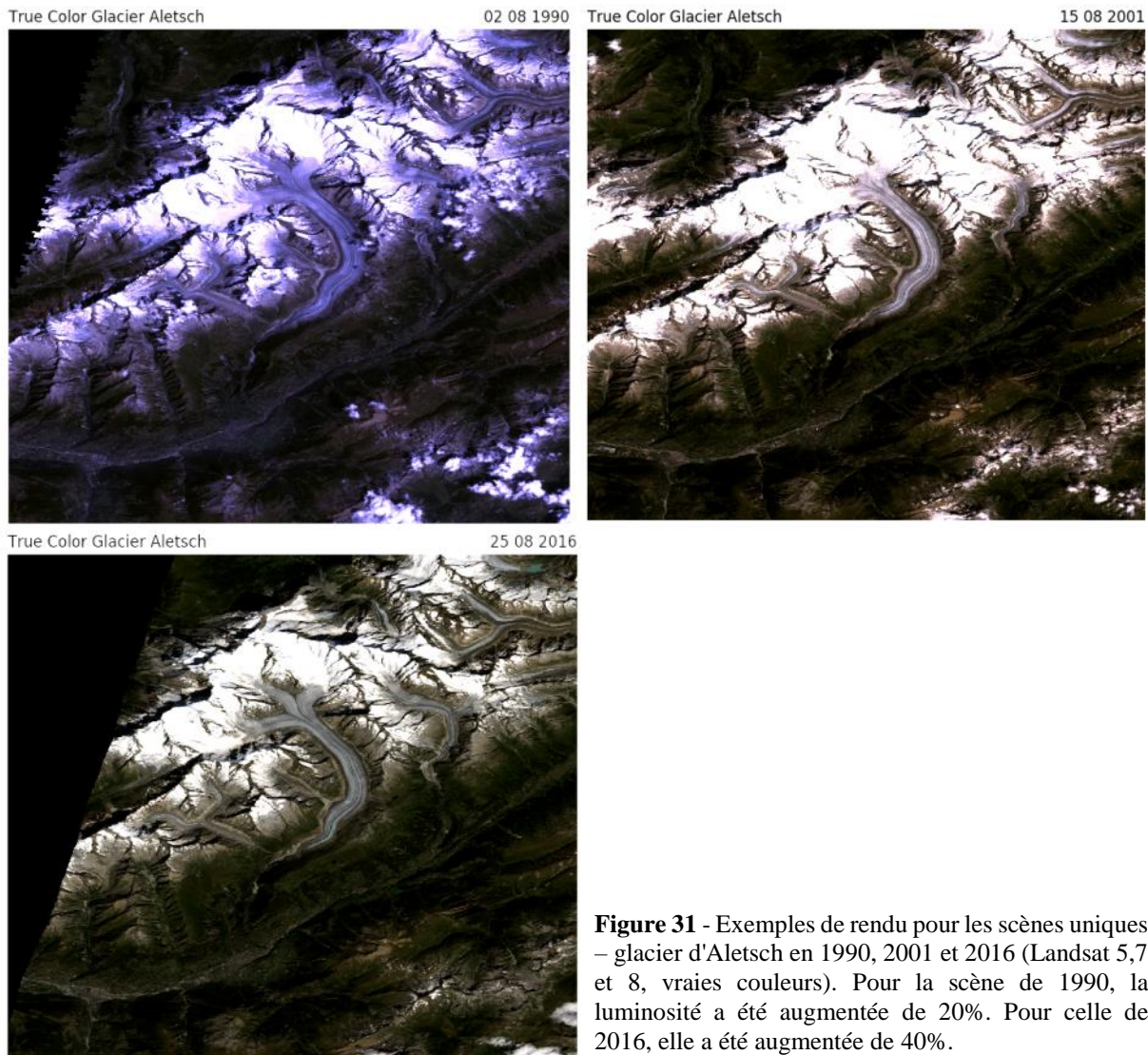
Les figures 28, 29 et 30 montrent l'excellente qualité des scènes obtenues au travers des scripts.

A la Figure 29, le problème de la résolution face à la taille des glaciers reste présent. On observe quand même un couvert neigeux plus important dans la scène de 1987 comparé à celui de 2010. Les détails des glaciers sont donc encore moins visibles. Mais l'indication sur les quantités de précipitations est bien évidemment primordiale pour l'analyse de l'évolution des glaciers.

C'est l'inverse à la Figure 30, où les précipitations neigeuses sont légèrement plus importantes à la date ultérieure. En 1990, le relief du glacier est donc mieux visible et la qualité du détail toujours impressionnante.

La Figure 31 illustre le glacier d'Aletsch. La teinte des images varie en fonction des Landsat (capteurs différents). Les scènes en été permettent de voir en détail la morphologie du glacier, mais il manque l'information sur les quantités de précipitations durant la période hivernale. Au mois d'août 1990, la surface du glacier semble légèrement plus importante qu'en 2001 et 2016. La morphologie et le relief sont également plus visibles en 2016, ce qui pourrait être la traduction d'une diminution en volume (hauteur) de la glace.

Lors de l'observation des glaciers, la longueur ainsi que les parties latérales sont à prendre en compte. Ces observations sont bien sûr plus faciles avec le glacier d'Aletsch.



**Figure 31** - Exemples de rendu pour les scènes uniques – glacier d'Aletsch en 1990, 2001 et 2016 (Landsat 5,7 et 8, vraies couleurs). Pour la scène de 1990, la luminosité a été augmentée de 20%. Pour celle de 2016, elle a été augmentée de 40%.

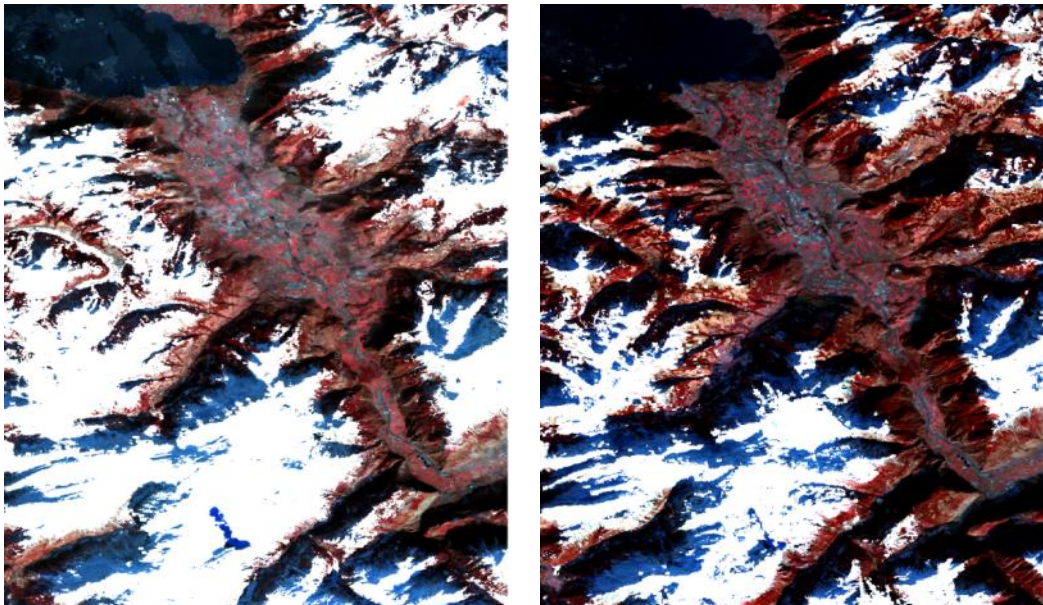
## Couvert neigeux – scènes uniques et mosaïques



**Figure 32** - Scènes uniques région des Portes du Soleil. L'image de droite a été produite avec le Landsat 7, celle de gauche avec le Landsat 8.

La Figure 32 est un exemple de scènes tirées des mosaïques. Elles permettent de voir la qualité des données à partir desquelles sont calculés les pixels médians pour faire les mosaïques.

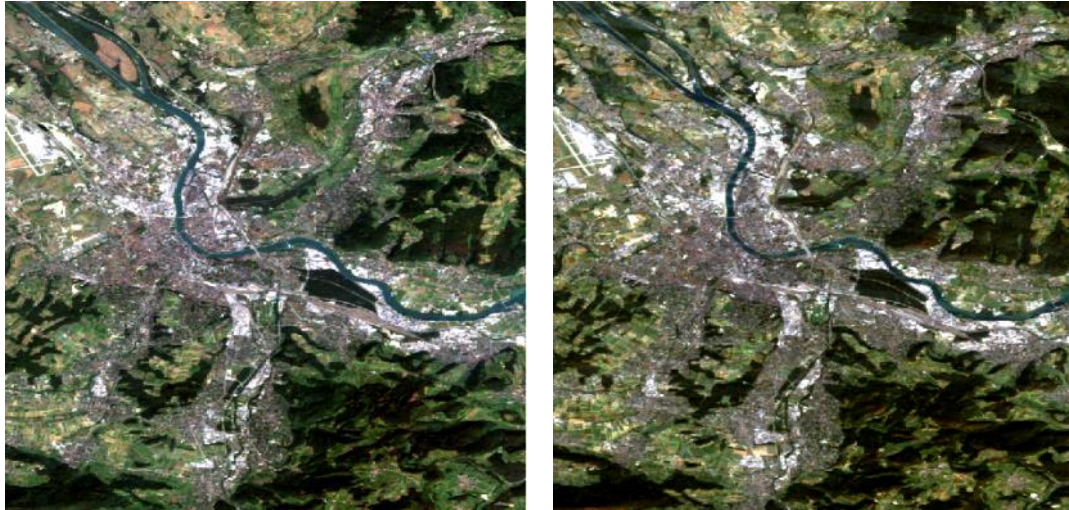
Concernant les mosaïques, les images TIF nommées "*MOS-lsdtX\_hiver20XX-20XX\_PDS\_NRV\_MedPix*" placées sur l'owncloud donne un exemple d'évolution du couvert nuageux pour la région des Portes du Soleil. La Figure 33 montre un exemple des rendus des mosaïques. Les données sont quasiment les mêmes qu'à la Figure 12 à une année d'intervalle.



**Figure 33** - Mosaïques pour la région des Portes du Soleil. À gauche, hiver 2001-2001, Landsat 7. A droite, hiver 2016-2017, Landsat 8.

### Zones urbaines - scènes uniques et mosaïques

Des composites avec la méthode du pixel médian ont été produites pour Zürich, Genève, Bâle, Lausanne et Bulle. Les images TIF nommées " *lsX\_median\_mosaic\_période\_nomdeVille\_...*" placées sur l'owncloud en sont le résultat. Les figures 33 à 37 sont des exemples des rendus des mosaïques.



**Figure 34** - Bâle. Landsat 7 et 8, vraies couleurs. A gauche, en 2000. A droite, en 2016.



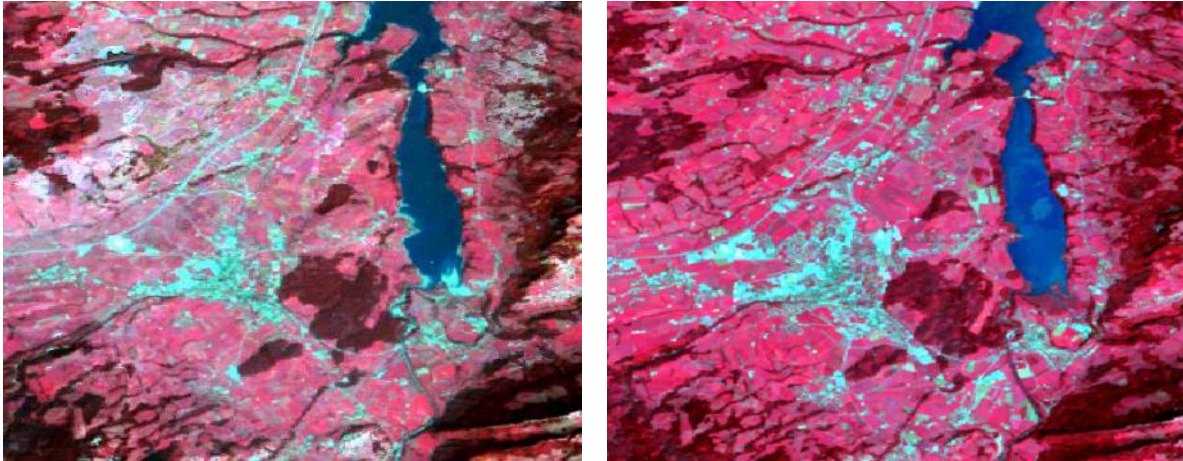
**Figure 35** - Lausanne. Landsat 7 et 8, vraies couleurs. En haut, en 2000. En bas, en 2016.



**Figure 36** - Zürich. Landsat 7 et 8, vraies couleurs. En haut, en 2000. En bas, en 2016.



**Figure 37** - Genève. Landsat 7 et 8, vraies couleurs. A gauche, en 1999. A droite, en 2016.



**Figure 38** - Bulle, bandes PIR, rouge, vert. A gauche, Landsat 5, du 1<sup>er</sup> janvier 1984 au 31 janvier 1985. A droite, Landsat 8, de mai à août 2016.

Comme pour les *Custom Mosaic*, il est difficile de déceler des changements importants entre les deux scènes pour chaque ville. Les vraies couleurs ont été présentées ici par pur choix subjectif. Pour Genève, du fait du nombre d'utilisation et d'occupation du sol différentes au sein d'un même territoire, il a été difficile de choisir des bandes permettant de bien faire ressortir le tissu urbain. Les images pour les quatre plus grandes villes avec les bandes PIR, rouge, vert se trouvent sur l'owncloud.

Les mosaïques en fausses couleurs ont été choisies pour présenter la ville de Bulle (Figure 38), étant la seule ville sélectionnée qui a montré un étalement du construit plus évident.



## 8. DISCUSSION

---

### 8.1. Utilisation et évolution de la plateforme `sdc.unepgrid.ch`

Pour commencer, un conseil global pour envisager de futures recherches est qu'il faudrait pouvoir se concentrer sur des zones géographiques précises. En ne sachant pas vraiment où chercher, un temps considérable a été perdu pour trouver, tester et comparer les zones géographiques. De plus, pour répondre aux objectifs sur les changements environnementaux considérés, il faut au préalable recouvrir une bibliographie suffisante pour interpréter les résultats des observations satellites étudiées (retrait des glaciers en termes de volume ? De surface ? A quelle vitesse ? Couvert neigeux : augmentation, diminution des précipitations ? Importance de la fréquence ? etc.). Or, dans le temps imparti, les recherches dans la littérature se sont concentrées sur la compréhension du SDC et des différents outils.

Concernant la visualisation des résultats, comme dans toute analyse en télédétection, il faut prêter attention lors de la comparaison des différents satellites Landsat (différences dans les capteurs, les bandes, la résolution).

#### 8.1.1. Custom Mosaic

##### Evolution des glaciers

Même si des différences entre années (essentiellement pour Aletsch) sont visibles, on ne peut tirer de tendances sur l'évolution glaciaire uniquement avec les images produites. On peut citer plusieurs hypothèses quant à la difficulté de voir le retrait des glaciers (Badoux and Rais 2015, IPCC 2017) : (i) le nombre d'images produit n'est pas assez élevé, notamment l'unique période de temps sélectionnée (l'été) pourrait être insuffisante, il faudrait prendre l'évolution des glaciers sur toute l'année puis comparer les années entre elles. (ii) La résolution des données peut être un frein selon l'étendue de la zone. (iii) La qualité, la quantité et le traitement des données, comme les premières images recueillies par le Landsat 5, les stries du Landsat 7 (à partir de mai 2003) ou encore la couverture `cf_mask` pour la glace. (iv) Les archives Landsat commençant en 1984, il n'est alors pas possible de comparer avec des données antérieures, comme le permettent les photos (les premières photographies des glaciers datent du 19<sup>ème</sup> siècle) et les mesures d'archives recouvrent une période de temps beaucoup plus longue. Elles peuvent ainsi montrer des différences plus flagrantes avec une proximité et des angles de vue qu'on ne retrouve pas avec les satellites. (Badoux and Rais 2015, Wikipédia 2017, IPCC 2017, Pronatura-Aletsch 2017). (v) Les méthodes de mesure sur le terrain permettent des analyses plus approfondies (mesure de la profondeur de la glace/du volume de la glace p.ex.), apportant des informations complémentaires dans la littérature sur le retrait des glaciers.

Par conséquent, aucune des vidéos produites n'a permis de tirer des observations catégoriques. Il est surtout possible de percevoir des différences dans les précipitations d'une année à l'autre.

Il faudra prêter attention car selon les mises à jour, il faudra à nouveau implémenter ou retravailler le `cf_mask`, autrement la glace n'apparaît pas (Delos Santos, ceos-seo/data\_cube\_utilities 2017).

## **Evolution du couvert neigeux**

Les remarques sont similaires que pour celles des glaciers. Les différences entre années au niveau des précipitations sont décelables, mais il n'est pas possible de relever des tendances quant à cette évolution avec les zones testées. Il faudrait tout d'abord pouvoir produire des images de qualité et de manière continue de 1984 à 2017 (notamment s'assurer que la neige apparaisse correctement), étendre l'analyse à d'autres régions puis les comparer à d'éventuels autres moyens de mesure et à des statistiques (durée et accumulation des précipitations, couverture neigeuse mesurées au sol, etc.) (OFEV, Office fédéral de météorologie et de climatologie 2017).

### **Zones urbaines**

Pour le tissu urbain, même si les bandes PIR, rouge et vert et PIR, SWIR1, SWIR2 sont des couleurs moins agréables pour l'œil humain, elles permettent généralement de mieux différencier le bâti du reste des couvertures du sol. Ainsi, en Suisse, il ne semble pas y avoir eu de grande évolution entre 1984 et 2016 pour les villes sélectionnées. Même si la surface bâtie tend à l'augmentation en Suisse (OFEV, Urbanisation et environnement 2017), le manque de résultat concernant l'évolution du tissu urbain pourrait provenir des résolutions où il est impossible de voir s'il y a bel et bien eu une évolution du tissu urbain ou par le fait que le territoire suisse recouvre une petite superficie; l'aménagement du territoire ne peut se faire sur de grands espaces. On peut imaginer que les nouvelles constructions ne recouvrent pas des aires suffisamment grandes pour être correctement décelées au travers des tests effectués. De plus, l'analyse possible uniquement en surface et non en volume de tissu urbain va influencer la perception du bâti. En télédétection, l'étude se fait principalement en surface (la position des satellites ne permet pas de distinguer la hauteur totale des constructions).

## **8.2. Suivi de la végétation**

### **8.2.1. Fractional Cover**

Malgré que le lien entre cet algorithme et l'étude des changements environnementaux et du suivi des écosystèmes n'a pas pu être investigué de manière approfondie, il devrait cependant être tout à fait possible d'utiliser cet algorithme à cette fin vu la qualité et la précision des trois bandes produites. Les résultats ont très bien montré des changements entre deux mois ou deux saisons.

Concernant l'interprétation des proportions de surfaces classées comme sols nus (BS) lorsqu'il semble y avoir plutôt de la forêt (comparaison avec Google Earth), cela pourrait être expliqué par le fait que l'alternance entre sols nus et surfaces boisées soit proche et que les bandes se superposent à certains endroits (cette observation avait déjà été faite pour le couvert neigeux, où le masque semblait parfois imprécis).

Pour les zones classées à une date en PV puis en BS pour la même période, on peut supposer des coupes dans la végétation ou lorsque cette dernière avait des conditions différentes qui ne lui permettaient pas de persister ultérieurement.

Il faut garder en mémoire que des comparaisons seulement entre deux périodes de temps très espacées dans le temps n'est pas pertinente pour évaluer l'évolution de la végétation grâce aux

trois bandes fournies par le FC. Il faudrait à nouveau faire des observations pour chaque année aux mêmes périodes durant toutes les années à disposition.

### **8.2.2. NDVI Anomaly**

Tout d'abord, le nombre de requêtes effectuées est faible, dû (i) à la compréhension du calcul effectué (ii) aux *bugs* du server et (iii) à la mémoire (RAM) saturée.

Ensuite, il a été compliqué de tester le NDVI Anomaly sans avoir une question de recherche plus précise ni des zones géographiques. Selon la visée de l'étude, la scène unique et la *baseline* vont changer (en matière de dates et de durées (scènes uniques et *baseline* sur un, deux, X mois?). Les tests effectués se sont inspirés des études déjà faites afin de tester l'algorithme, mais l'objectif général de voir les liens entre les changements d'écart de NDVI et un lien possible avec le suivi des écosystèmes et les changements environnementaux est trop vaste.

Il n'est à nouveau pas aisé de tirer des conclusions hâtives avec les tests effectués. On aperçoit des améliorations comme des diminutions en quantité (ou en qualité) de végétation, mais le grand nombre de facteurs de causalités rendent l'analyse de l'évolution de la végétation complexe (saisons, couverture du sol (forêt vs agricole p.ex), type de végétation (feuillue vs résineuse), météo, écart entre les *baselines* et les scènes uniques, quantité de données pour chaque calcul, etc.), ajouté au fait que le calcul de l'écart de NDVI n'est pas clair. Si le calcul se base bel et bien sur une seule scène, la quantité et la qualité de données (nombre de *clean pixels* p.ex.) pourrait être insuffisante. Concernant la *baseline*, il est difficile de prendre seulement un mois en compte, car de nombreuses fois il n'y a pas assez d'informations pour faire les calculs.

Malgré tout, la possibilité de visualiser quatre images avec des légendes différentes offre de multiples possibilités. On peut observer les changements de NDVI dans les scènes uniques ou le NDVI médian dans les *baselines*, qui donnent une première information sur l'état de la végétation. Les informations sur l'écart de NDVI apportent une dimension ainsi qu'une comparaison temporelle plus importantes dans les analyses. Une possibilité intéressante serait de pouvoir sélectionner un seul mois comme *baseline*, afin de comparer la scène unique avec des données d'archives correspondant au même mois. Par la suite, changer la scène unique de mois en mois et modifier la *baseline* en conséquence. Tout dépend de l'objectif de l'étude (savoir tout d'abord si l'on compare une même zone dans le temps ou des zones différentes, quelles sont les hypothèses quant à l'évolution de ces zones, quel est le but de calculer un écart de NDVI dans ces cas, etc.).

### **8.3. Avantages et améliorations de la plateforme sdc.unepgrid.ch**

Pour terminer cette section, les tests des trois outils ont été rassemblés afin de lister les avantages et améliorations possibles sur l'utilisation de la plateforme sdc.unepgrid.ch.

#### **Avantages**

- Le fond de carte, qui a été amélioré depuis le début du stage, avec la localisation des glaciers, les noms des principales villes, des parcs naturels, etc.
- La possibilité de sélectionner une zone d'intérêt manuellement ou d'inscrire les longitudes et latitudes dans les filtres.

- Les filtres offrent de nombreuses possibilités, avec un historique, les résultats obtenus et les outputs. Les résultats de chaque scène sont détaillés (nombre de clean pixels, nombre total de pixels, etc. par scène, ce qui permet de connaître celles qui contiennent le plus d'information et ainsi de savoir si la quantité d'information sur une certaine période est suffisante ou non pour être utilisable). On obtient également le nombre de scènes par mosaïque. Dans les *résultats* et les *outputs*, il est possible de choisir et d'afficher uniquement une scène composant la mosaïque. Dans les *outputs*, il est possible d'obtenir une synthèse de ce que l'image contient comme information et de télécharger l'image sous différents formats.
- Les fonctionnalités des outils proposés sont multiples (mosaïque, fractional cover, etc.) (onglet *tools*). Ils sont pertinents pour analyser l'environnement sous différents angles (analyse de la végétation, des glissements de terrain (et par extension des zones à risque), de l'eau, etc.). Différentes combinaisons de bandes sont mises à disposition afin d'élargir les types de rendus (vraies et fausses couleurs faisant ressortir les différentes occupations ou utilisations du sol (neige, glace, eau, tissu urbain, zone agricole, forêt, etc.)).
- Grand avantage de la plateforme : elle offre une utilisation "use-friendly" composée d'une seule interface pour générer des séries temporelles. Le résultat est un gain de temps, une visualisation aisée avec des comparaisons avec des données d'archives.

### **Pistes pour améliorer la plateforme**

- Même si le fond de carte s'est grandement amélioré, il serait utile soit d'ajouter encore des informations sur la carte déjà établie (noms des lieux plus précis, tracé des routes, ...) soit d'en avoir plusieurs regroupées sous forme de cartes thématiques par exemple.
- Ajouter une barre de recherche afin de sélectionner un lieu.
- Pouvoir aller rechercher une scène dans le *task manager* et la télécharger à nouveau sur la carte interactive. Actuellement, il existe la possibilité d'aller dans l'historique ou d'aller re-sélectionner à la main les minimum et maximum de la longitude et de la latitude. Mais les scènes ne se retrouvent pas toujours dans l'historique et il existe un risque de les perdre.
- Un problème survenu plus tard durant le stage est la taille des images téléchargées au format PNG. Lors du téléchargement, elle est grandement diminuée et selon l'étendue de la zone d'étude au départ il devient impossible de voir quoi que ce soit.
- La possibilité de générer des animations en série temporelle n'est pas encore au point (que ce soit dans l'onglet filtre ou dans les possibilités de formats dans les outputs).
- La résolution et la qualité des images pourraient ne pas être adéquates selon l'objectif de l'étude voulue.
- Même si elle s'est améliorée depuis le début du stage, la durée de traitement des requêtes reste "longue" (relativement normal vu la quantité d'informations demandée en même temps) alors que les zones d'étude restent petites (échelle des cantons, voire des communes). Ainsi, des mosaïques sur l'entier de la Suisse demandent encore un certain temps.
- Les mises à jour engendrent souvent de nouveaux problèmes; il faut prêter attention à la qualité des images, à la mémoire disponible, à la gestion des outils (cf. les images

floues *Custom Mosaic*), à la gestion du *cf\_mask* notamment pour les couvertures de neige et de glace.

#### 8.4. Tests des scripts établis à l'aide du notebook Jupyter

La remarque concernant les problèmes dus aux mises à jour est maintenue. Il n'a pas été toujours aisé d'obtenir des images avec une résolution adéquate, les tests ont été ralentis durant les manipulations à cause des bugs (souvent dus à la saturation de la mémoire) et des mises à jour des notebooks. Un certain temps a également été perdu en raison d'un problème d'utilisation de certains modules. Par exemple, les commandes `"import dc_utilities as utilities/dc = DataAccessApi()" / "from utils.data_access_api import DataAccessApi"` ne fonctionnaient pas sous mon répertoire "Laetitia" mais ont fonctionné sous Datacube – data\_cube\_notebooks. La localisation des scripts est donc importante dans FilZilla. Ces différents problèmes ajoutés à une maîtrise du langage Python insuffisante ont fait qu'il n'a pas été possible de transcrire la production de scènes, de mosaïques et de vidéos avec le *Fractional Cover* et le *NDVI Anomaly* sous forme de script permettant de générer leurs résultats automatiquement.

Concernant la différence de qualité entre les images élaborées avec la plateforme et les images produites avec les scripts, elle n'a pas pu être expliquée.

Les observations faites sur l'évolution des glaciers, du couvert neigeux et du tissu urbain pour les images produites avec la plateforme sont valables aussi ici. Même avec l'augmentation de la qualité, il n'est pas possible de tirer des tendances claires sur ces évolutions avec les images obtenues durant ce stage. Toutefois, avec le gain de temps permis grâce aux scripts, il sera plus aisé de produire les images (avec une qualité adéquate) et de manière continue entre 1984 et 2017. Malgré que les vidéos ou ensembles de scènes produites n'ont pas permis des analyses approfondies (la mise en place des scripts et le tri de chaque scène manuellement prend énormément de temps), on peut déjà observer de manière certaine des différences dans le couvert neigeux et par conséquent (dans certains cas) des différences dans les zones d'accumulation ou les fronts des glaciers (sous l'influence des dates et de la météo au moment de la prise de vue).

#### 8.5. Avantages et améliorations des scripts

Pour les scripts, les avantages et les améliorations possibles sont les suivants :

##### Avantages

- Grâce aux scripts du notebook, il était déjà possible de produire des images et d'obtenir les informations sur les données de manière beaucoup plus rapide. Avec l'automatisation, le gain de temps est encore augmenté. Ainsi, il est possible de produire un grand nombre d'images en peu de temps, rendant les comparaisons et la sélection des meilleures images beaucoup plus aisées.
- L'augmentation de la qualité des images.
- La création de vidéos à partir des scènes.
- La production de mosaïques en format TIF enregistrées directement sur le server. La lecture des mosaïques se fait ensuite avec un autre logiciel (comme QGIS p.ex.)

## **Pistes pour améliorer les scripts**

Malgré tous les tests effectués, les scripts comportent encore des lacunes. Un grand nombre d'améliorations et d'ajouts de fonctions restent à faire, notamment :

- L'automatisation n'est pas encore complète. Il faudrait par exemple pouvoir sélectionner uniquement les "meilleures scènes", à savoir celles qui contiennent le plus de données et les données les plus complètes (autrement dit l'abandon des scènes contenant des nuages et le maintien de celles n'en contenant pas). Pour ce faire, il faut obtenir (i) un *cf\_mask* cachant les nuages et les ombres; ces manipulations sont possibles dans les notebooks des Australiens (fonctions implémentées dans l'Australian Geoscience DC), mais il nous manque un paramètre pour le faire fonctionner dans le SDC (ii) enlever les *no data*.
- Combiner les trois Landsat au sein du même script.
- Sélectionner uniquement des périodes (comme les saisons) sur plusieurs années.
- Avoir du code supplémentaire permettant d'éclaircir les images en cas de besoin.
- Faire des vidéos à partir des images en format TIF (les mosaïques/composites) ou pouvoir les établir dans un format (idéalement en PNG) permettant de les assembler sous forme de vidéo.
- Étendre les scripts pour tous les outils.

## 9. CONCLUSION

---

Au moment du stage, le SDC venait d'être terminé. Il a été très intéressant de se situer au début d'un tel projet et de voir en quatre mois les améliorations. Mais les tests des différentes fonctionnalités et des plateformes ont aussi été très frustrants. L'instabilité et même les impossibilités d'utiliser le SDC à certains moments ont grandement ralenti les essais et les objectifs du début ont dû être modifiés. A la fin de ce stage, les conclusions sur les fonctionnalités des outils utilisés n'apportent rien de nouveau; tous les collaborateurs ont pu faire leur propre évaluation de la plateforme et tester les différents outils.

Malgré que l'analyse du lien entre les deux algorithmes (FC et NDvA) et les changements environnementaux n'ait pas été faite, une étude approfondie devrait pouvoir être réalisée à l'avenir compte tenu des résultats encourageant obtenus. En effet, la qualité et précision des indices et de leur rendu graphique permettraient des analyses précises dans le temps et l'espace selon les objectifs des recherches.

Ainsi, le Suisse Data Cube offre une nouvelle technologie du traitement de l'information spatiale et temporelle avec un fort potentiel de développement, capable de mesurer différemment les changements environnementaux.

## **10. DÉROULEMENT DU STAGE**

---

Le stage au sein du GRID Genève a été une bonne expérience. Au départ, les objectifs du stage ne correspondaient pas exactement à mes attentes (manque d'utilisation des compétences en géomatique acquises), mais après discussion avec mes superviseurs, une redirection ainsi qu'un ajout d'objectif a permis d'y remédier en partie.

Malgré des conditions qui n'étaient pas toujours faciles (uniquement dues à des problèmes techniques liés à la mise en place de la plateforme `sdc.unepgrid.ch`), l'environnement social de travail a été excellent grâce au très bon accueil des collaborateurs du GRID et au suivi de mes superviseurs.

Le seul regret est peut-être de ne pas avoir appliqué davantage de connaissances géomatiques pouvant être plus utiles auprès d'un futur employeur (p.ex. une utilisation plus avancée d'ArcGIS ou de QGIS).

Dans le cadre d'un stage au sein du GRID Genève, je conseille au futur collaborateur de se renseigner sur le sujet proposé au préalable et de ne pas hésiter à demander des modifications des objectifs afin qu'ils correspondent à ses attentes.



## 11. BIBLIOGRAPHIE

---

- Allenbach, Karin, and Hy Dao. "Cours Geotools-RS : télédétection et traitements d'images." Université de Genève, 2017.
- Badoux, Sophie, and Cécile Rais. *Le dramatique recul des glaciers suisses*. 2015. <https://www.rts.ch/info/suisse/7151867-le-dramatique-recul-des-glaciers-suisses.html#story-anchor-En+savoir+plus> (accessed 2017).
- Bulot, Jean-Claude, and Antonio Da Cruz. "PhotoFiltre 7." *PhotoFiltre Studio*. PhotoFiltre Studio, 2005.
- CEOS. *Data Cubes for Large Scale Data Analytics*. July 7, 2017.
- CEOS, and GRID Genève. *The CEOS Open Data Cube Initiative*. Genève, 2016.
- CEOS, Australian Government, CSIRO, USGS. *Open Data Cube*. 2017. <https://www.opendatacube.org/contacts> (accessed 2017).
- . *Open Data Cube - xarray*. 2017. <http://xarray.pydata.org/en/stable/dask.html> (accessed 2017).
- CEOS, GRID Genève. *Swiss Data Cube*. 2017. <http://sdc.unepgrid.ch/> (accessed 2017).
- Chatenoux, Bruno. "GIS Officer." Geneva: GRID Genève, 2017.
- Delos Santos, Alfredo. *ceos-seo/data\_cube\_utilities*. Juin 2017. [https://github.com/ceos-seo/data\\_cube\\_utilities/blob/03e6bd41d620ec44150649945bed8f1f9fd5e8ee/dc\\_utilities.py#L34](https://github.com/ceos-seo/data_cube_utilities/blob/03e6bd41d620ec44150649945bed8f1f9fd5e8ee/dc_utilities.py#L34) (accessed 2017).
- . "Software Developer." Hampton: Analytical Mechanics Associates, 2017.
- ESRI, ArcGIS. *ArcGIS for Desktop - Qu'est-ce qu'une mosaïque*. 2016. <http://desktop.arcgis.com/fr/arcmap/10.3/manage-data/raster-and-images/what-is-a-mosaic-dataset.htm> (accessed 2017).
- Ezgif. *Ezgif.com*. 2017. <https://ezgif.com/> (accessed 2017).
- Geoscience Australia, CSIRO, NCI. *Open Data Cube*. 2017. <https://datacube-core.readthedocs.io/en/stable/dev/api.html> (accessed 2017).
- GIFMaker.Me. *GIFMaker.Me*. 2017. <http://gifmaker.me/> (accessed 2017).
- Giuliani, Gregory, et al. "Building an Earth Observations Data Cube: lessons learned from the Swiss Data Cube on generating Analysis ready Data." *Big Earth Data*, 2017.
- Giuliani, Gregory, Pascal Peduzzi, Bruno Chatenoux, Karin Allenbach, and Jean-Philippe Richard. *The Swiss Data Cube*. Geneva: GRID Geneva, 2017.
- Google. *Google Earth*. 2017.
- GRID Genève. *GRID-Geneva*. 2017. <http://www.grid.unep.ch/index.php?lang=fr> (accessed 2017).

- Hamunyela, Eliakim, Jan Verbesselt, Sytze de Bruin, and Martin Herold. "Monitoring Deforestation at Sub-Annual Scales as Extreme Events in Landsat Data Cubes." *Remote Sensing*, 2016: 16.
- Hebert, Philippe, and Bernard Rougé. "Mosaïques numériques d'images." *XVth ISPRS Congress*. Toulouse: Centre National d'Etudes Spatiales, 1984. 104-113.
- IPCC. 2017. <https://www.ipcc.ch/> (accessed 2017).
- IPython. *Jupyter*. September 20, 2017. <http://jupyter.org/> (accessed 2017).
- Killough, Brian. "CEOS Analysis Ready Data for Land." 2016.
- Kosse, Tim. "FileZilla." FileZilla, 2017.
- Lewis, Adam, et al. "Rapid, high-resolution detection of environmental change over continental scales from satellite data - the Earth Observation Data Cube." *International Journal of Digital Earth*, October 20, 2015: 6.
- NCI, National Computational Infrastructure. *NCI*. 2017. <http://nci.org.au/> (accessed 2017).
- OFEV. *Changements climatiques en Suisse*. scientifique, Berne: Confédération suisse, 2007.
- . *Office fédéral de météorologie et de climatologie*. 2017. <http://www.meteosuisse.admin.ch/home.html?tab=overview> (accessed 2017).
- . *Urbanisation et environnement*. 2017. <https://www.bafu.admin.ch/bafu/fr/home/themes/urbanisation/en-bref.html> (accessed 2017).
- Office fédéral de la statistique . *Utilisation et couverture du sol*. 2017. <https://www.bfs.admin.ch/bfs/fr/home/statistiques/espace-environnement/utilisation-couverture-sol.html> (accessed 2017).
- Organisation des Nations Unies pour l'alimentation et l'agriculture, FAO . *Observation de la Terre - indicateurs mondiaux*. 2017. [http://www.fao.org/giews/earthobservation/asis/index\\_2.jsp?lang=fr](http://www.fao.org/giews/earthobservation/asis/index_2.jsp?lang=fr) (accessed 2017).
- OSGeo, Open Source Geospatial Foundation. "QGIS." *QGIS*. OSGeo, 2017.
- Pronatura-Aletsch. *Centre Pronatura d'Aletsch*. 2017. <https://www.pronatura-aletsch.ch/rechauffement-climatique> (accessed 2017).
- Richard, Jean-Philippe. "IT Officer." Geneva: GRID Genève, August 14, 2017.
- Scarth, Peter, Achim Röder, and Michael Schmidt. *Tracking Grazing Pressure and Climate Interaction - The Role of Landsat Fractional Cover in Time Series Analysis*. Indooroopilly, March 28, 2011.
- Schweiz Tourismus. *Aletsch, la route du grand glacier*. 2017. <https://www.myswitzerland.com/fr-ch/aletsch-panoramaweg-la-route-du-grand-glacier.html> (accessed 2017).
- Tatham, Simon. "PuTTY." *putty*. Windows, 2017.

- United States Geological Survey . *About Landsat*. 2017. <https://landsat.usgs.gov/about-landsat> (accessed 2017).
- Weier, John, and David Herring. *Measuring Vegetation*. 2000. [https://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring\\_vegetation\\_1.php](https://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_1.php) (accessed 2017).
- Wikipédia. *Glacier d'Aletsch*. 2017. [https://fr.wikipedia.org/wiki/Glacier\\_d%27Aletsch](https://fr.wikipedia.org/wiki/Glacier_d%27Aletsch) (accessed 2017).
- . *Histoire de la photographie*. 2017. [https://fr.wikipedia.org/wiki/Histoire\\_de\\_la\\_photographie](https://fr.wikipedia.org/wiki/Histoire_de_la_photographie) (accessed 2017).
- WSL, Institut fédéral de recherches sur la forêt, la neige et le paysage. *Neige et glace*. 2017. <https://www.wsl.ch/fr/neige-et-glace.html> (accès le 2017).

## 12. ANNEXES

---

### A. Scripts principaux

#### A.1. Data\_Cube\_API\_Demo-CopyLB.py

```
# coding: utf-8

import datacube
from datetime import datetime
import matplotlib.pyplot as plt

dc = datacube.Datacube()

# **Low level Data Cube functions**
# List off all of the products in the datacube
# refers to the product definitions (datacube -v product add <prod>), and the output type in #
the ingestion configuration
dc.list_products()
# Lists all the measurement information for all products, including ones added by the user
# and ones added during ingestion
dc.list_measurements()

#http://agdc-
#v2.readthedocs.io/en/stable/dev/generate/datacube.Datacube.load.html#datacube.Datacube.lo
#ad

data = dc.load(product="ls7_ledaps_swiss", platform="LANDSAT_7",
               measurements=['red', 'green', 'blue', 'cf_mask'],
               latitude=(46.25, 46.62), longitude=(7.82,8.24),
               time=(datetime(2000,5,1), datetime(2000,7,30)))

#hist = data.red.plot()

data_plot = data.isel(time=28).red.plot(cmap='BuPu',figsize=(10, 8))
hist = data.red.plot()
data_plot = data.isel(time=28).cf_mask.plot(cmap='BuPu',figsize=(10, 8))

# it is also possible to resample data on load using the output_crs, resolution, and resampling
#options.
#data = dc.load(product="ls7_ledaps_swiss", platform="LANDSAT_7",
#               #measurements=['red', 'green', 'blue', 'cf_mask'],
#               #latitude=(45.9, 46.2), longitude=(7.4,7.8),
#               #time=(datetime(2001,1,1), datetime(2002,1,1)),
#               #output_crs='EPSG:4326',
#               #resolution=(-30, 30),
#               #resampling='cubic')
```

```

#hist = data.red.plot()
#data_plot = data.isel(time=1).red.plot(cmap='BuPu')

# **Our Data Access API**

from utils.data_access_api import DataAccessApi
dc = DataAccessApi()

# get metadata for an entire product
dc.get_datacube_metadata('LANDSAT_7', 'ls7_ledaps_swiss')

# Get metadata for specific acquisitions - includes storage unit paths, extents, etc.
#dc.get_scene_metadata(product="ls7_ledaps_swiss", platform="LANDSAT_7",
#    latitude=(46.25, 46.62), longitude=(7.82,8.24),
#    time=(datetime(2000,5,1), datetime(2000,7,30)))

dc.list_acquisition_dates(product="ls7_ledaps_swiss", platform="LANDSAT_7",
    latitude=(45.9, 46.2), longitude=(7.4,7.8),
    time=(datetime(2000,5,1), datetime(2000,7,30)))

# Load data just like with the Data Cube API, just wrapped in a container for consistency
#data = dc.get_dataset_by_extent(product="ls7_ledaps_swiss", platform="LANDSAT_7",
#    latitude=(46.25, 46.62), longitude=(7.82,8.24),
#    time=(datetime(2000,5,1), datetime(2000,7,30)))
#print(data)
#hist = data.red.plot()
#data_plot = data.isel(time=1).red.plot(cmap='BuPu')

```

## A.2. test\_scenes\_video\_LB.py

```
import os
import argparse
import datacube
from datetime import datetime
import matplotlib
matplotlib.use('Agg')
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd

dc = datacube.Datacube()

parser = argparse.ArgumentParser(description='Plotting multi-band images of Aletsch Glacier')
parser.add_argument('product', help='Product name e.g. ls7_ledaps_swiss')
parser.add_argument('platform', help='Product name e.g. LANDSAT_7')
parser.add_argument('lat_min', help='Product value e.g 46.25')
parser.add_argument('lat_max', help='Product value e.g 46.62')
parser.add_argument('long_min', help='Product value e.g 7.82')
parser.add_argument('long_max', help='Product value e.g 8.24')
parser.add_argument('start_time', help='Time value e.g 2001-7-1')
parser.add_argument('end_time', help='Time value e.g 2001-8-31')
args = parser.parse_args()

rgb2 = dc.load(product=args.product, platform=args.platform,
               latitude=(args.lat_min, args.lat_max), longitude=(args.long_min, args.long_max),
               time=(args.start_time, args.end_time),
               measurements=['red', 'green', 'blue'],
               stack='color').transpose('time', 'latitude', 'longitude', 'color')

fake_saturation2 = 3000
clipped_visible2 = rgb2.where(rgb2 < fake_saturation2).fillna(fake_saturation2)
max_val2 = clipped_visible2.max(['longitude', 'latitude'])
scaled2 = (clipped_visible2 / max_val2)

os.system("rm *.png")
max_time = (len(clipped_visible2['time']))
cnt = 1
for i in range(0, max_time):
    plt.figure(figsize=(8, 8))
    sel = scaled2.isel(time=i)
    plt.imshow(sel)
    plt.axis('off')
    t = pd.to_datetime(str(sel['time'].values))
```

```

plt.title(t.strftime('%d %m %Y'),loc='right')
plt.title('RGB Glacier Aletsch',loc='left')
plt.savefig('tmp.png')
# plt.savefig('myfig_%i.png' % (i)) # cette ligne de code est à faire au lieu de la ligne
#du dessus pour visualiser les images et conserver les meilleures (et voir leur taille)
size = os.stat('tmp.png').st_size
if size >= 582000:
    cnt += 1
    os.rename('tmp.png', 'myfig_%03d.png' % (cnt))

#----avec images blanches au début et à la fin de la vidéo----#

os.system('convert myfig_002.png -threshold -1 -alpha off myfig_001.png')
os.system('convert myfig_002.png -threshold -1 -alpha off myfig_%03d.png' % (cnt+1))
os.system('convert *.png -morph 22 %04d.morph.png')
os.system('ffmpeg -framerate 8 -i %04d.morph.png -vcodec mpeg4 -b 1000k movieN.avi')

```

### A.3. test\_scenes\_video\_2\_LB.py

```
import os
import argparse
import datacube
from datetime import datetime
import matplotlib
matplotlib.use('Agg') #bug
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd

dc = datacube.Datacube()

parser = argparse.ArgumentParser(description='Plotting multi-band images of Aletsch Glacier')
parser.add_argument('product', help='Product name e.g. ls7_ledaps_swiss')
parser.add_argument('platform', help='Product name e.g. LANDSAT_7')
parser.add_argument('lat_min', help='Product value e.g 46.25')
parser.add_argument('lat_max', help='Product value e.g 46.62')
parser.add_argument('long_min', help='Product value e.g 7.82')
parser.add_argument('long_max', help='Product value e.g 8.24')
parser.add_argument('start_time', help='Time value e.g 2001-7-1')
parser.add_argument('end_time', help='Time value e.g 2001-8-31')
args = parser.parse_args()

rgb2 = dc.load(product=args.product, platform=args.platform,
               latitude=(args.lat_min, args.lat_max), longitude=(args.long_min, args.long_max),
               time=(args.start_time, args.end_time),
               measurements=['red', 'green', 'blue'],
               stack='color').transpose('time', 'latitude', 'longitude', 'color')

fake_saturation2 = 3000
clipped_visible2 = rgb2.where(rgb2 < fake_saturation2).fillna(fake_saturation2)
max_val2 = clipped_visible2.max(['longitude', 'latitude'])
scaled2 = (clipped_visible2 / max_val2)

os.system("rm *.png")
max_time = (len(clipped_visible2['time']))
cnt = 0
for i in range(0, max_time):
    plt.figure(figsize=(8, 8))
    sel = scaled2.isel(time=i)
    plt.imshow(sel)
    plt.axis('off')
    t = pd.to_datetime(str(sel['time'].values))
```



```

plt.title(t.strftime('%d %m %Y'),loc='right')
plt.title('True Color Glacier Aletsch',loc='left')
plt.savefig('tmp2.png')
# plt.savefig('myfig_%i.png' % (i)) # cette ligne de code est à faire au lieu de la ligne
#du dessus pour visualiser les images et conserver les meilleures (et voir leur taille)
size = os.stat('tmp2.png').st_size
if size >= 582500:
    cnt +=1
    os.rename('tmp2.png','myfig_%02d.png' % (cnt +1))

#---sans images blanches au début et à la fin de la vidéo---#
os.system('convert *.png -morph 20 %04d.morph.png')
os.system('ffmpeg -framerate 7 -i %04d.morph.png -vcodec mpeg4 -b 1000k movie8.avi')

```

#### A.4. test\_kmean\_clustering\_LB.py

```
from utils.data_access_api import DataAccessApi
from utils.dc_utilities import write_geotiff_from_xr, create_cfmask_clean_mask
from utils.dc_chunker import create_geographic_chunks, combine_geographic_chunks
from utils.dc_mosaic import create_median_mosaic

# from utils.dc_mosaic import create_mosaic_iterative, create_max_ndvi_mosaic,
# create_min_ndvi_mosaic

from datetime import datetime
import xarray as xr
import gdal
import numpy as np
import matplotlib
import matplotlib.pyplot as plt

dc = DataAccessApi()

lat_range = (46.10001, 46.4232)
lon_range = (5.9494, 6.3716)
time_range = (datetime(1999,1,1), datetime(2000,1,1))

product = "ls7_ledaps_swiss"
platform = "LANDSAT_7"

product_details = dc.dc.list_products()[dc.dc.list_products().name == product]
acquisitions = dc.list_acquisition_dates("LANDSAT_7", product, time=time_range,
                                         longitude=lon_range, latitude=lat_range)
geographic_chunks = create_geographic_chunks(latitude=lat_range, longitude=lon_range,
                                             geographic_chunk_size=.05)

measurements = ['nir', 'swir1', 'swir2', 'cf_mask']
# measurements = ['blue', 'green', 'red', 'nir', 'swir1', 'swir2', 'cf_mask']

product_chunks = []

for index, chunk in enumerate(geographic_chunks):
    data = dc.get_dataset_by_extent(product, time=time_range,
                                   longitude=chunk['longitude'], latitude=chunk['latitude'],
                                   measurements=measurements)
    clean_mask = create_cfmask_clean_mask(data.cf_mask)
    product_chunks.append(create_median_mosaic(data, clean_mask=clean_mask))
# create_median_mosaic, create_max_ndvi_mosaic, create_min_ndvi_mosaic

final_mosaic = combine_geographic_chunks(product_chunks)

write_geotiff_from_xr("ls7_median_mosaic_01-1999-2000_Geneve_NS1S2.tif",
                    final_mosaic, measurements, nodata=-9999)
```