

Analyse des séries temporelles d'image Sentinel-2 pour la détection des épicéas scolytés.

ULG - Gembloux Agro-Bio Tech
September 29, 2021

Contents

1	Introduction	2
1.1	Sentinel-2	2
2	Méthodologie	3
2.1	Suivi d'un indice spectral pour la détection précoce de stress .	3
2.2	Ensemble de règles pour attribuer un état sanitaire	6
2.3	Illustrations	8
3	Le logiciel s2_timeSerie	8
3.1	Organisation du code source	15
3.2	La classe <code>catalogue</code>	15
3.3	La classe <code>tuileS2OneDate</code>	17
3.4	La classe <code>TS1Pos</code>	17
4	Les post-traitements avec s2_postProcess	18
5	Validation	19
5.1	Introduction	19
5.2	Validation terrain	20
5.2.1	Résultats	21

1 Introduction

Dans le cadre du projet InterReg RegioWood 2 et de l'Accord cadre de recherche et vulgarisation forestière, la crise du Typographe qui touche les pessières wallonnes est étudiée au moyen de la télédétection. L'objectif est de dresser des cartes d'état sanitaire pour chaque année étudiée sur lesquelles on puisse distinguer les arbres sains des arbres dépérissants. L'utilisation de série temporelle d'images satellites dispose d'un potentiel très intéressant pour le suivi de la phénologie des arbres.

Le présent document est un document interne de l'unité de Gestion des ressources forestières (Ulg). Pour toutes questions ou remarques, contactez Lisein jonathan par courriel (liseinjon@hotmail.com).

1.1 Sentinel-2

L'imagerie multispectrale des deux satellites Sentinel-2 (A et B, mis en orbite en juin 2015 et mars 2017) du programme Copernicus est à la base de la méthodologie présenté dans ce document. La résolution des bandes spectrales de Sentinel-2 est de 10 mètres dans le meilleurs des cas (certaines fréquences utilisées pour le suivi sanitaire sont captées à 20 mètres). La fréquence de revisite à l'équateur est de 5 jours, mais on dispose en Belgique de l'ordre d'une 12aine de prises de vue par ans pour lesquelles la couverture nuageuse est suffisamment faible. Les prises de vue font l'objet de traitement et sont redécoupées selon un carroyage de tuile (figure 1). Le shapefile des tuiles et des centroïdes sont disponible sous <https://github.com/justinelliotmeyers/Sentinel-2-Shapefile-Index>.

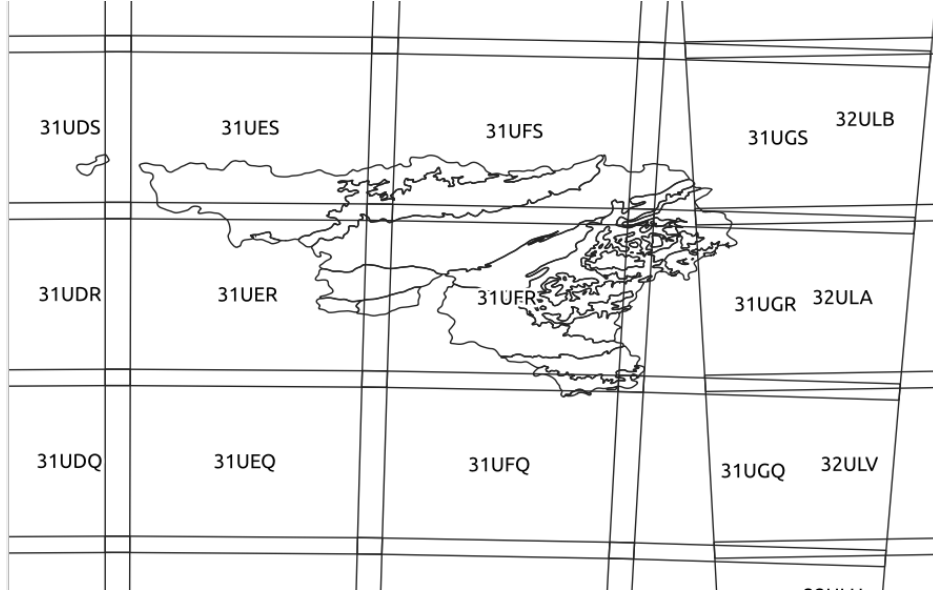


Figure 1: La Wallonie est couverte par un total de 8 tuiles Sentinel-2, mais la tuile 31UFR couvre à elle seule une majorité de l'Ardenne Belge.

2 Méthodologie

2.1 Suivi d'un indice spectral pour la détection précoce de stress

L'idée générale est de se calquer sur la méthodologie de l'INRAE (INRAE, UMR TETIS, R. Dutrieux, K. Ose, J.-B. Féret) développée depuis 2017.

L'indice spectral permettant la détection précoce (stade vert du dépérissement) des dépérissements provoqués par l'attaque de Typographes est le CRSWIR :

$$CR_{SWIR1} = \frac{SWIR1}{(NIRa + (\lambda_{SWIR1} - \lambda_{NIRa}) * (\frac{SWIR2 - NIRa}{\lambda_{SWIR2} - \lambda_{NIRa}}))}$$

avec

$$\lambda_{NIRa} = 865$$

$$\lambda_{SWIR1} = 1610$$

$$\lambda_{SWIR2} = 2190$$

A noter que CR est l'acronyme de Complex Ratio. Les valeurs d'un peuplement soumis à un stress physiologique voient leur CRSWIR augmenter de manière précoce.

$$f(t) = a_1 + b_1 \sin\left(\frac{2\pi}{T}t\right) + b_2 \cos\left(\frac{2\pi}{T}t\right) + b_3 \sin\left(\frac{2\pi}{T}2t\right) + b_4 \cos\left(\frac{2\pi}{T}2t\right) \quad (1)$$

L'équation 1 permet de modéliser les variations saisonnières de CRSWIR pour un peuplement d'épicéa sain. La constante T est égal à 365.25. Cette équation ajustée sur 300 pessières saines est illustrée pour 3 années sur la figure 2. Pour un pixel donnée, l'étude de la série temporelle de valeurs de CRSWIR, correspondant donc à chaque date pour lesquelles une image Sentinel-2 peu ennuagée est disponible, permet de déterminer si le peuplement est soumis à un stress. Plusieurs situations existent en dehors de celle d'une pessière saine et de celle d'une pessières scolytées. En effet, il est également nécessaire de déterminer si le peuplement est coupé (détection de sol nu), ou si le peuplement présente un stress temporaire dû probablement à un déficit hydrique temporaire. Les pessières sur des sols à Régime Hydrique Alternatif sont fortement représentatif de cette dernière situation, avec typiquement un déficit hydrique estival entraînant un stress passager qui se manifestera par une augmentation de CRSWIR, mais qui se différencie d'une attaque de scolyte par un retour à un état sanitaire normal (diminution du CRSWIR) peu de temps après.

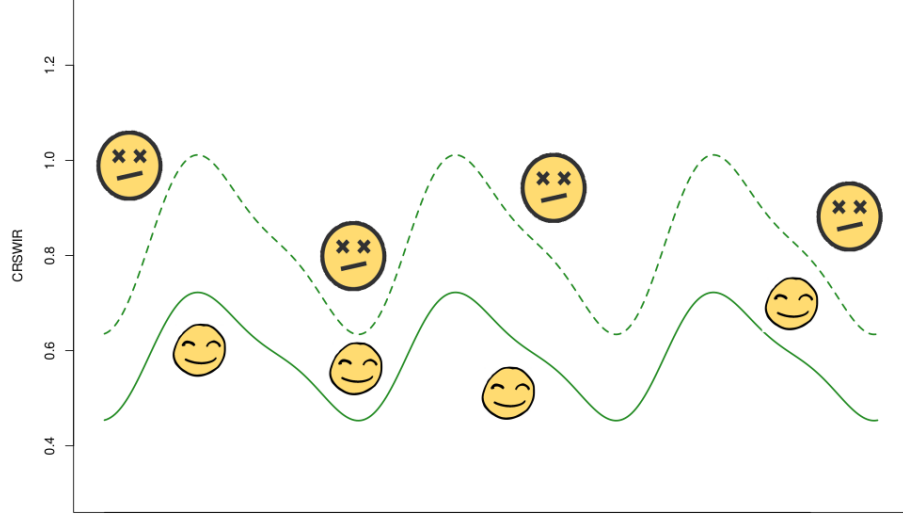


Figure 2: L'indice CRSWIR pour un peuplement sain varie durant l'année selon une fonction harmonique (trait plein en vert sur la figure). Un arbre dépérissant va présenter une valeur de CRSWIR qui sera supérieure à la normale. On décide arbitrairement que les valeurs qui dépassent la courbe en pointillé sont considérées comme représentative d'un stress.

Afin de faciliter la comparaison d'une valeur de CRSWIR pour une date donnée avec la valeur de référence, qui est la valeur de CRSWIR d'une pessière saine représenté par l'équation 1, nous utilisons un ratio tels que défini ci-dessous :

$$CR_{SWIR_{norm}}(t) = \frac{CRSWIR_{observe}}{CRSWIR_{theorique}(t)} \quad (2)$$

le $CRSWIR_{theorique}(t)$ représente bien entendu la valeur calculée au moyen de l'équation 2. Le CRSWIR normalisé est donc un ratio qui, lorsque sa valeur est de 1, représente l'état typique d'une pessière saine, et lorsqu'il dépasse un certain seuil (fixé entre 1.5 et 1.7), est considéré comme représentatif d'un stress végétatif. Une fois le CRSWIR normalisé calculé pour chaque date de prise de vue, la suite des traitements peut s'effectuer sans à avoir à considérer la variation saisonnière inhérente aux changements phénologiques de la végétation.

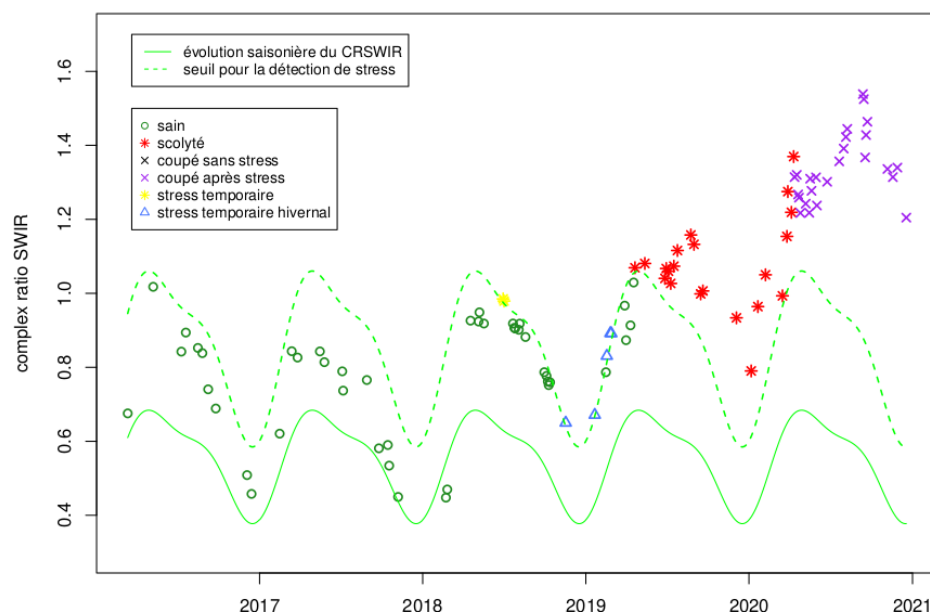




Figure 3: Exemple d'une série temporelle d'observation de CRSWIR pour un pixel donné. Un état sanitaire présumé est attribué à chaque observation en fonction de la valeur de CRSWIR, de la présence ou non de sol nu et à son état sanitaire pour les dates antérieures et postérieures. Par exemple, en 2020, l'état sanitaire attribué est celui "coupé après stress" (croix mauve sur le graphique) car d'une part le pixel est détecté comme étant en sol nu, d'autre part les observations antérieures (été 2019 et hiver 2020) sont considérées comme étant "scolytées" (étoile rouge sur le graphique)

2.2 Ensemble de règles pour attribuer un état sanitaire

1. ; sain
2. ; dépérissement
3. ; coupé mais sans dépérissement détecté avant la coupe
4. ; coupé après avoir été dépérissant

5.  ; stress passager, donc à priori plus un stress lié à un déficit hydrique estival qu'à une attaque de scolyte.
6.  ; pixel qui est un mélange, détecté car présente un stress passager en hiver (feuillus sans photosynthèse) suivi d'un retour à la normale en été (résineux + feuillus en photosynthèse)

Les filtres et les règles de décisions permettant d'attribuer un état sanitaire à une observation d'une série temporelle sont brièvement présentés ici. On peut considérer que la série temporelle pour un pixel donné est synthétisée sous forme d'un tableau à 2 colonnes pour lequel chaque ligne constitue une date pour laquelle on dispose d'une information (prise de vue non ennuagée pour cette position). La première colonne est celle contenant l'information de la date de l'observation. La deuxième colonne contient un état sanitaire présumé, qui peut prendre la valeur de 1, 2 ou 3 avant de subir l'analyse présentée ici. La valeur de 1 signifie que la valeur de $CRSWIR$ permet de supposer que le peuplement est sain. La valeur de 2 indique par contre que le peuplement semble en situation de stress, avec un $CRSWIR_{normalise}$ supérieur au seuil fixé (ex. 1.4, voir ligne en pointillé de la figure 2).

L'analyse commence par un premier filtre qui vise à retirer les valeurs aberrantes de la série temporelle. On va pour ce faire contrôler que aucune valeur de 2 (stress) ou 3 (coupé) n'apparaît de manière isolée. L'hypothèse est simplement qu'un pixel détecté comme stressé à une date mais considéré comme sain à la date précédente à la date suivante correspond à une observation aberrantes qui sera donc retiré ici.

Ensuite, un deuxième filtre vise à classer les coupes (code 3) et les coupes sanitaires (code 4). On va pour se faire identifier toutes les occurrences d'au moins 3 détection de sol nu consécutifs, ou encore de 2 détections de sol nu mais dont les 2 dates sont séparées d'au moins 40 jours (pour considérer les situations de sols nus au dessus d'une zone fortement ennuagé et donc avec une densité temporelle d'observations assez faible). Lorsqu'une situation similaire est détectée, l'état de toutes les observations suivantes sont fixées à "coupé" ou à "coupe sanitaire" si un stress a été détecté à la date antérieure à la détection du sol nu. L'idée est bien d'éviter qu'une position détectée comme sol nu ne soit classée comme étant une pessière saine les mois et années suivant la coupe. Nous attirons l'attention que la méthodologie employée ici sur des images de 10m de résolution n'as pas pour objectif d'obtenir de bon résultats sur des jeunes plantations, pour lesquels la réponse spectrale

est un mélange résultant de celle des petits houppiers et de la végétation environnante.

La troisième partie de l'analyse vise à la détection des situations de peuplements scolytés. Si un stress est observé plusieurs fois d'affilé, on va modifier les états sanitaires postérieurs pour s'assurer que ceux-ci correspondent soit à un stress, soit à une coupe sanitaire. Néanmoins, un *retour à une situation normale* est accepté dans le cas où un état sanitaire sain est constaté plus de 3 fois consécutivement, pour une durée excédant 30 jours et seulement dans les cas où le dépérissement observé précédemment n'a pas duré plus longtemps qu'un certain nombre de jours (par défaut 90 jours). Lorsqu'un **retour à la normale** est ainsi détecté, l'état sanitaire durant la période de stress est considéré comme étant un stress temporaire (code 5) et non pas une situation de présence de scolyte.

Enfin, un dernier filtre permet d'appréhender la situation assez complexe mais marginale d'un pixel qui couvrirait une surface hétérogène du point de vue de la présence de l'épicéa, tels qu'un mélange épicéa-feuillus ou un mélange épicéa-sol nu. La détection de mélange vise à réduire le taux de faux positif pour les zones scolytées.

2.3 Illustrations

3 Le logiciel s2_timeSerie

Bien que conceptuellement la méthode ne soit pas extrêmement compliquée, le volume de donnée est considérable et nécessite une organisation soignée ainsi qu'une méthode de calcul optimisée. C'est dans ce but qu'est développé l'application **s2_timeSerie**, sous C++ et dans un environnement Linux.

L'application peut s'utiliser de la manière suivante :

```
./s2_timeSerie --help  
./s2_timeSerie --catalogue 1 --tuile T32ULU  
./s2_timeSerie --catalogue 2 --tuile T31UGR --srCR 1.7
```

l'argument **tuile** renseigne le nom de une ou plusieurs tuiles Sentinel-2. l'argument **catalogue** peut recevoir les valeurs de 1 ou de 2. Deux modes de création d'un catalogue, c'est à dire d'une série temporelle d'images pour une tuile donnée, sont disponibles.

Avec le mode 1, **s2_timeSerie** effectue le téléchargement des données depuis la plateforme theia. Les informations nécessaires sont le nom de la tuile Sentinel-2 et une plage de date durant laquelle on souhaite effectuer l'analyse.

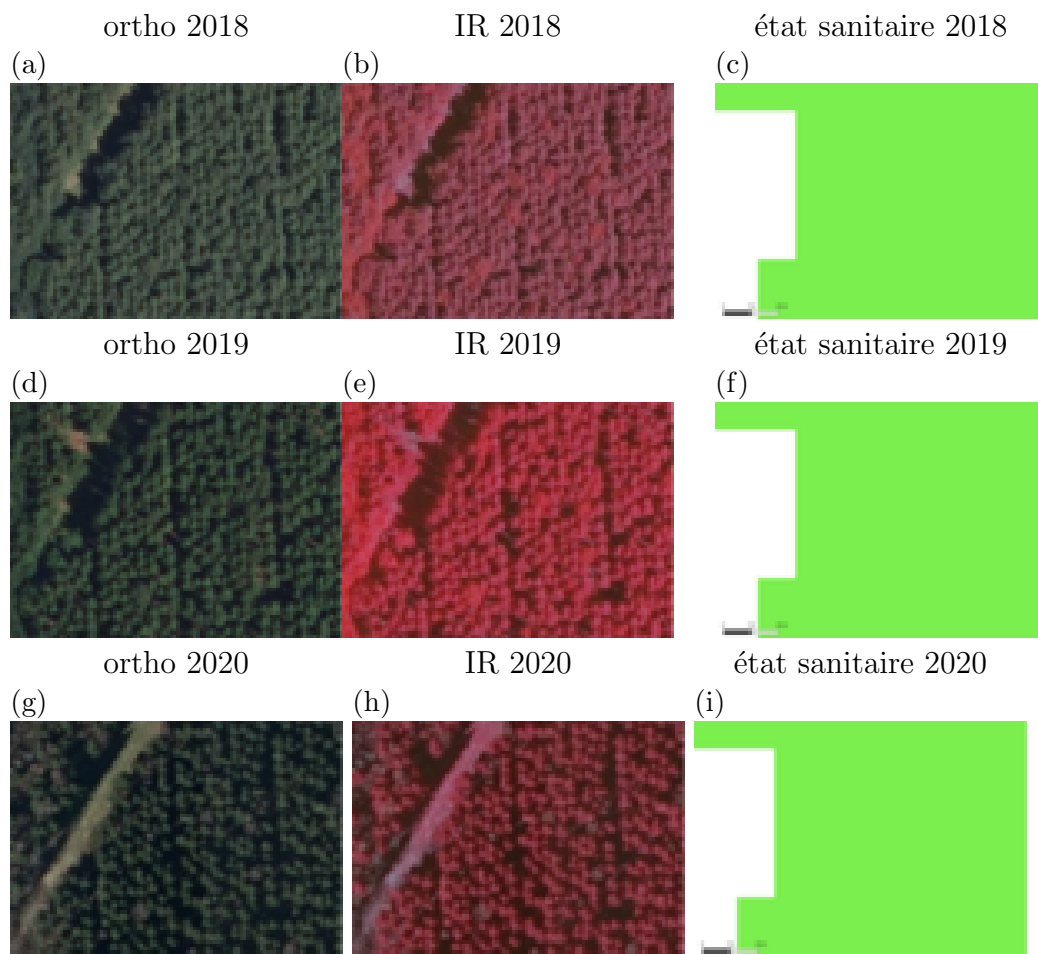


Figure 4: Illustration pour une pessière saine



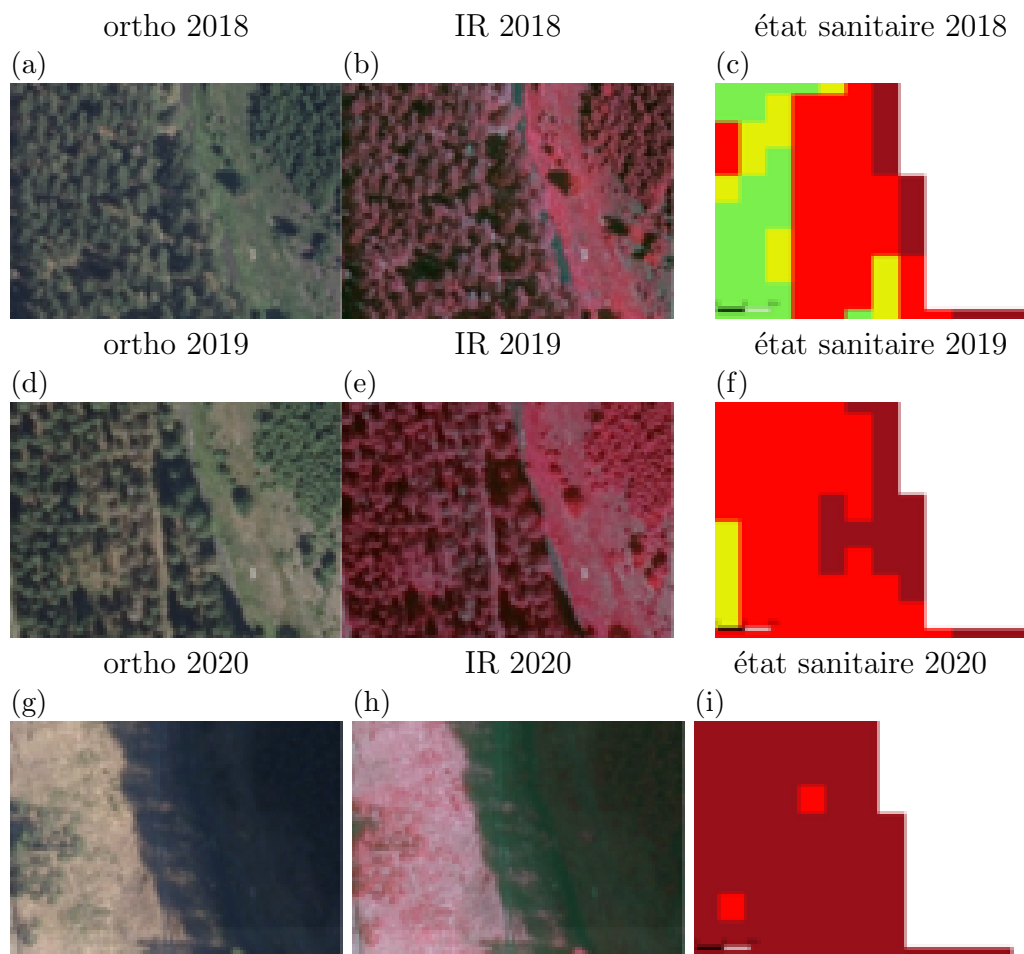
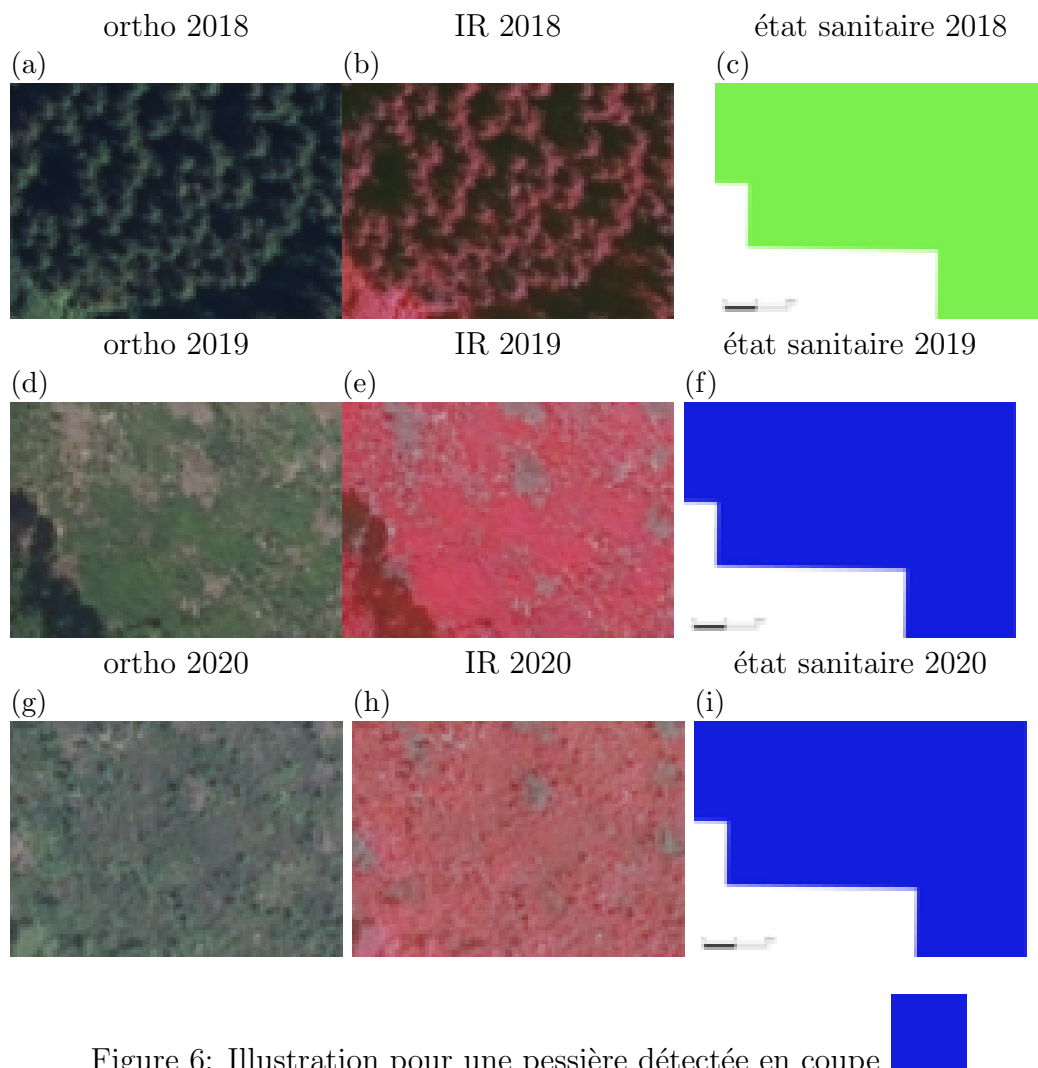


Figure 5: Illustration pour une pessière scolytée





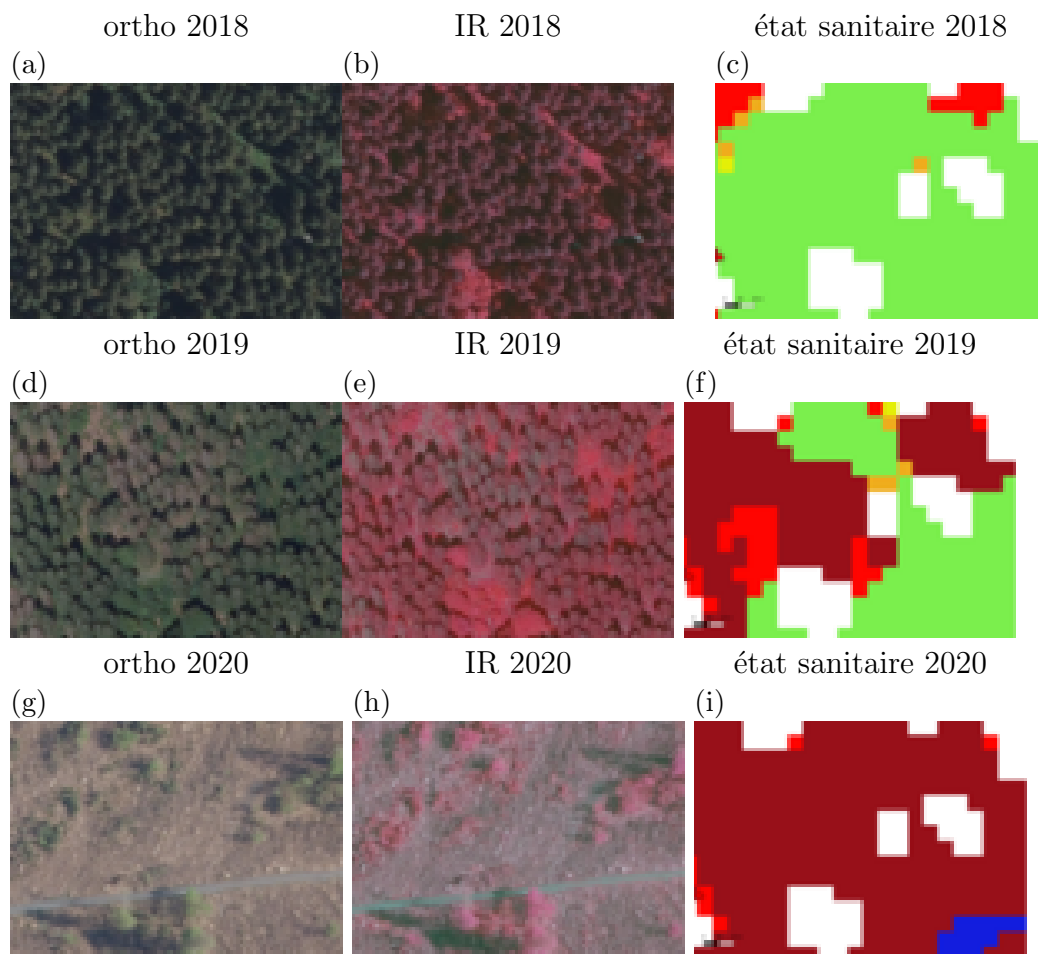


Figure 7: Illustration pour une pessière détectée en coupe sanitaire

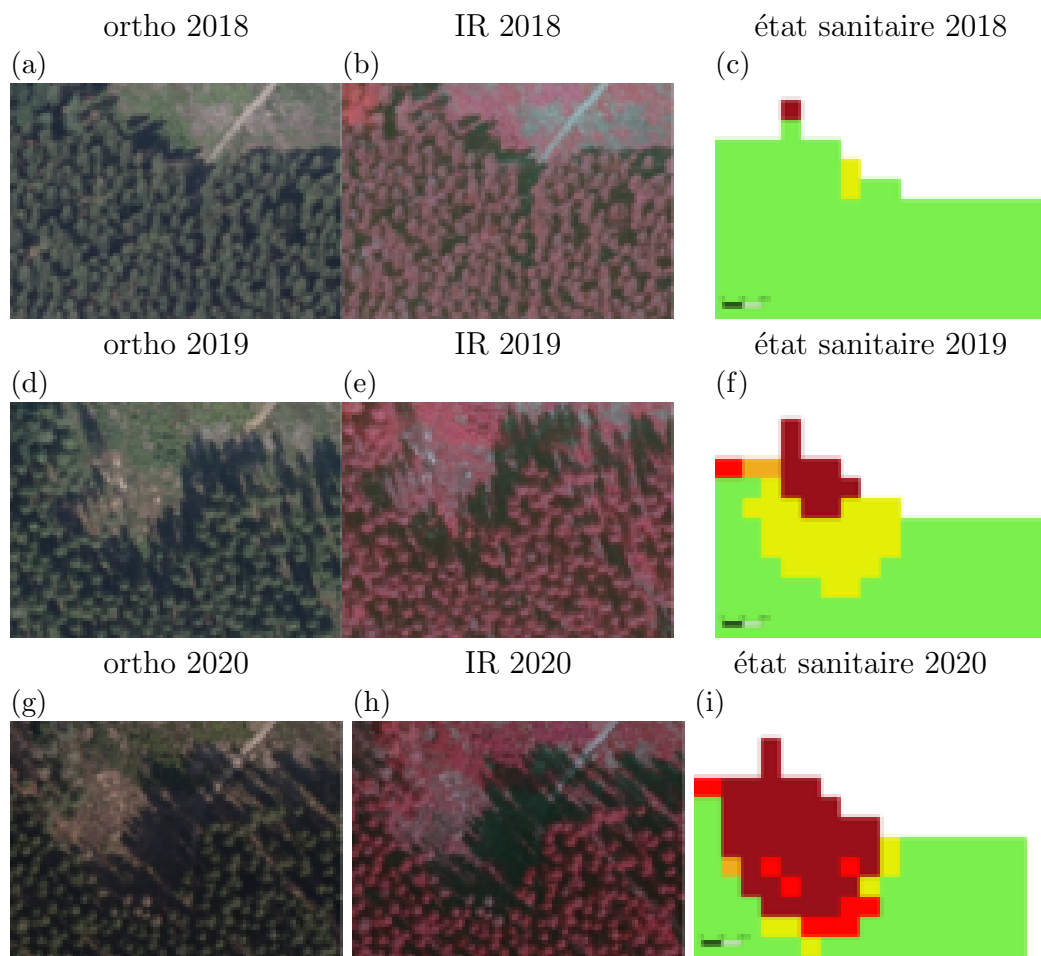


Figure 8: Illustration pour une pessière soumis à un stress temporaire



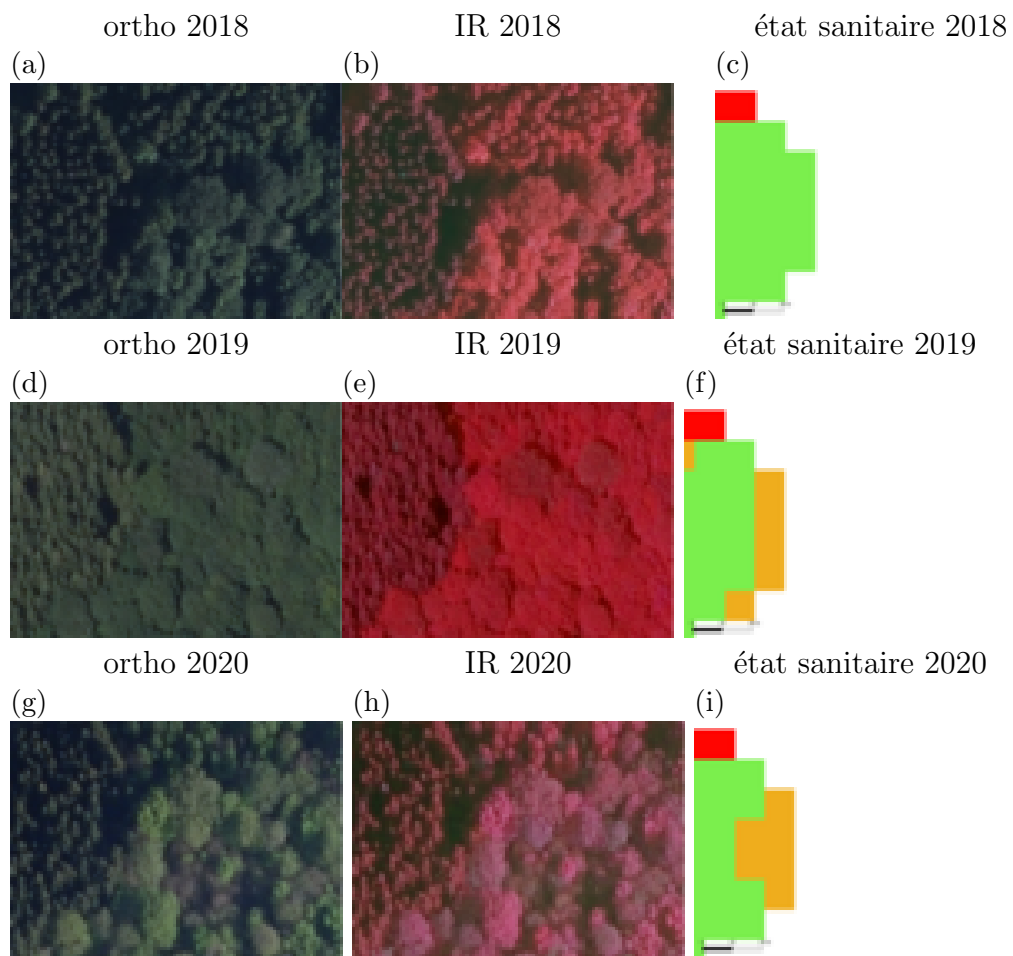


Figure 9: Illustration pour des pixels détectés comme étant en mélange



Cette plage de date peut être définie avec l'argument `-dates` . par exemple : `-dates 2016-01-01 2017-01-01`.

Avec le mode 2 (`-catalogue 2`), le catalogue est créé sur base des dossiers téléchargés précédemment. Chaque dossier correspond à une date de prise de vue. Seul le nom de la tuile est donc requis.

L'appli utilise OTB comme calculatrice raster, pour tout ce qui concerne les éléments ci-dessous :

- calcul de carte de masques (no data et nuage)
- rééchantillonnage des bandes spectrales de Sentinel-2 de 20m vers 10m de résolution
- calcul de carte pour la détection du sol nu
- le calcul des cartes de CRSWIR
- le calcul des cartes de CRSWIR normalisé

L'ensemble de ces opérations sont effectuées date par date de manière indépendante, et restent donc de l'ordre du monotemporel. Leur relative simplicité font que ces opérations sont considérées comme des prétraitements. Ces prétraitements sont évidemment préliminaires à l'analyse de la série temporelle à proprement parler.

3.1 Organisation du code source

Trois classes d'objets sont définies dans le projet C++ et permettent d'effectuer toutes les opérations nécessaires. Ces classes sont présentées brièvement ci-dessous. Ces informations sont une aide destinée à l'utilisateur qui souhaite lire le code source pour le comprendre et/ou le modifier. La figure ?? schématise le fonctionnement des 3 classes d'objets les unes par rapports aux autres.

3.2 La classe catalogue

C'est la classe qui contient toutes la série temporelle sur l'emprise d'une tuile. La série temporelle est constituée par l'ensemble des prises de vue sur cette tuile. Les prises de vue retenues sont celles qui présentent un taux d'enneigement inférieur à un seuil fixé à 35%. C'est au niveau de cette classe que sont défini les dates de début et de fin de la série temporelle, mais

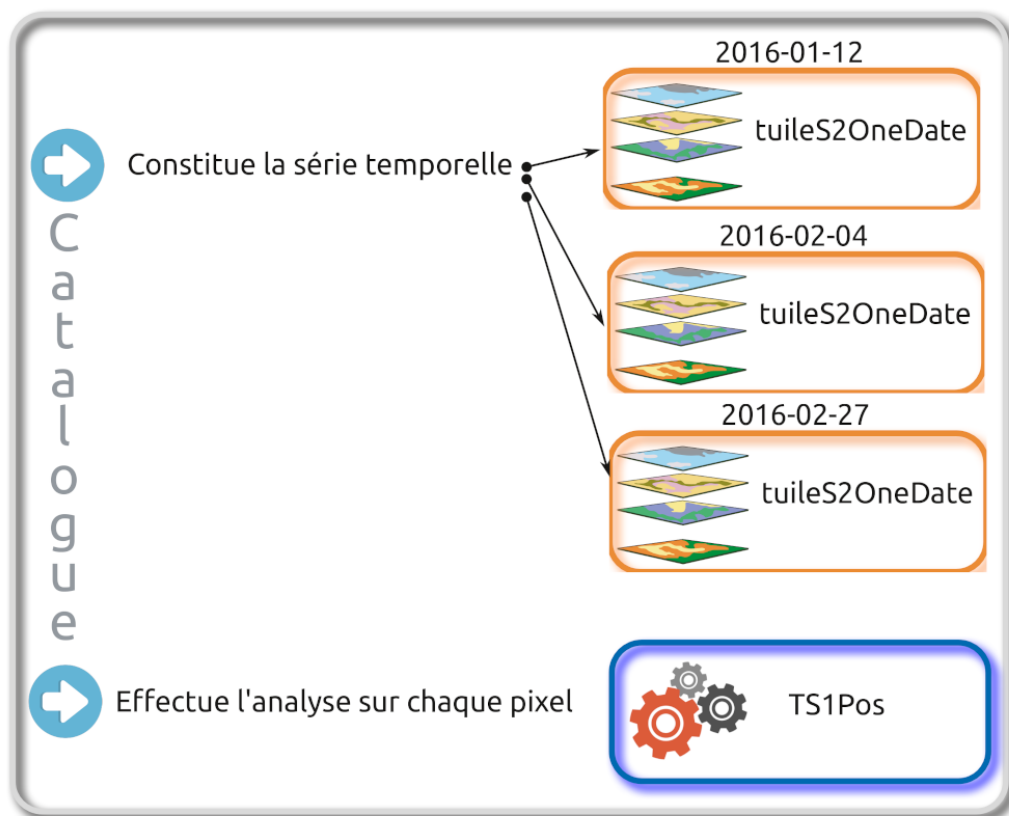


Figure 10: L'application est développée avec l'approche de programmation Orientée Objet, dont les 3 classes principales sont les classes **catalogue**, **tuileS2OneDate** et **TS1Pos** représentés sur ce graphique.

aussi que la zone d'intérêt pour laquelle on souhaite effectuer les calculs est définie. En effet, l'analyse n'est pas effectuée sur l'ensemble du territoire mais uniquement sur les zones considérées comme étant des pessières. Le masque épicéa est donc utilisé comme donnée d'entrée et ce masque, qui couvre toute la Région Wallonne, est découpée par la classe catalogue afin d'épouser l'emprise de la tuile.

3.3 La classe **tuileS2OneDate**

C'est la classe qui contient toutes les informations relatives à une tuile pour une date donnée ; archive compressée, décompressée, métadonnées (date, nuage, nom du produit, répertoire où j'ai mis les images, etc.). Cette classe effectue la lecture de nombreuses informations depuis le fichier XML de métadonnées du produit Sentinel-2A pour cette date. Plus d'information sont disponibles sous le lien <https://labo.obs-mip.fr/multitemp/sentinel-2/theias-sentinel-2-l2a-product-format/> à propos du format de données des produits 2A de Sentinel-2.

Les objets **tuileS2OneDate** regroupent l'ensemble des images nécessaires à l'analyse temporelle pour la détection de stress provoqué par les scolytes. En plus des images de chaque bande spectrale du capteur de Sentinel-2, les images calculées par OTB listées précédemment dans ce document sont gérées par la classe **tuileS2OneDate**.

3.4 La classe **TS1Pos**

C'est la classe qui effectue l'analyse de la série temporelle (**TS** pour Time Serie) pour une position donnée (**Pos** pour Position, qui est représenté par le centre de chaque pixels de l'image). C'est avec cette classe que l'on fait l'analyse de la succession de l'état sanitaire pour déterminer si la pessière en ce point est scolytée, coupée, saine, etc. La classe **TS1Pos** n'effectue aucune opération cartographique, c'est strictement l'analyse du CRSWIR normalisé au fil du temps, combiné avec la détection ou non de sol nu, qui sert d'information. Outre l'analyse de la série temporelle pour une position, cette classe a une méthode pour exporter un résultat synthétique pour cette position, en l'occurrence un état sanitaire *synthétique* pour chaque année. En effet, disposer d'un état sanitaire pour chaque année est un compromis entre trop de détails et trop de synthèse.

4 Les post-traitements avec `s2_postProcess`

Les traitements illustrés jusqu'à présent sont des calculs qui s'effectuent sur base de la série temporelle, mais sans prendre en compte les effets de voisinage. On peut dire que la dimension spatiale n'est pas prise en compte. Pour assurer une certaine cohérence dans les classes d'état sanitaires assignées à chaque pixel dans les cartes annuelles, un certain nombre de post traitements sont effectués et développés dans une seconde application. Cette application est assez polyvalente, avec un nombre d'outils assez élevés. Les outils permettent les actions suivantes ;

0. masquer les cartes d'états sanitaires avec la carte de probabilité de présence de l'épicéa, en utilisant un seuil de présence spécifié. Pour rappel, l'analyse de la série temporelle est effectuée sur une zone définie sur base d'une probabilité de présence de l'épicéa de plus de 50%. On peut donc avec cet outil se montrer plus restrictif en augmentant ce seuil, par exemple avec un seuil de 70%.
1. nettoyage des cartes d'états sanitaires. Pour le moment, il s'agit uniquement de localiser les zones de coupe non-sanitaire de petites dimensions qui sont entourées d'une coupe sanitaire. On met l'hypothèse que ces épicéas ont également été attaqué par le scolyte avant d'être abattu, comme leur voisins direct, et on change la classe pour les mettre en coupe sanitaire.
2. calcul des cartes d'évolutions de l'états sanitaires. Chaque carte d'état sanitaire est comparée avec celle de l'année précédente afin de classer les zones de scolyte en nouveau scolyte ou ancien scolyte, et les zones de coupe sanitaire en nouvelle ou ancienne coupe sanitaire.
3. compression des cartes geotiff
4. calcul de statistique (surface) pour chaque classe d'état sanitaire.
5. projection des cartes vers le SRC 31370 Belge Lambert 72.
6. Extraction de valeurs d'état sanitaire pour les points d'un shapefile, dans un objectif de validation de la méthode. Diverses indices de voisinage sont également calculés, comme la distance la plus proche à un foyer de scolyte.

Voici les codes pour l'évolution de l'état sanitaire ainsi que la couleur associée:

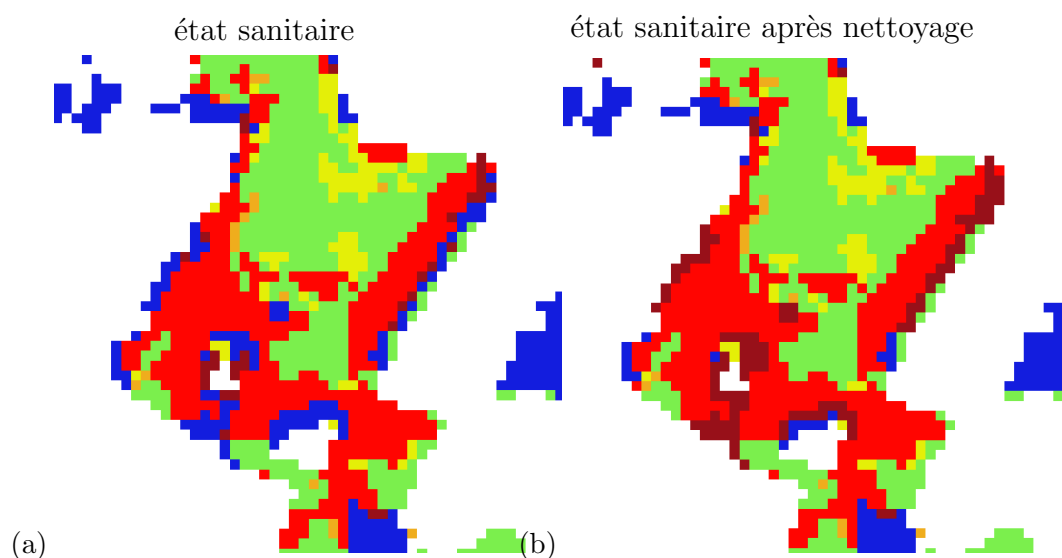




Figure 11: Un des post-traitement vise à améliorer les cartes d'état sanitaire, sur cette illustration on voit que certaines coupe "normale" sont reclassée en coupe sanitaire sur base de relation de voisinage.

- 21.  ; ancien scolyte
- 22.  ; nouveau scolyte
- 41.  ; ancienne coupe sanitaire
- 42.  ; nouvelle coupe sanitaire

5 Validation

5.1 Introduction

La méthode de détection des scolytes sera validée selon plusieurs méthodes, en fonction des données dont nous disposons.

5.2 Validation terrain

L'équipe du CRPF Grand-est a effectué des mesures sur le terrain en 2021 (projet RégioWood II) pour un total de 127 placettes en pessière. L'échantillonnage est stratifié pour disposer de relevés en pessières scolytées et non-scolytées représentant de manière équilibrée les 3 classes d'altitudes choisies : 0-600 mètres, 600-800 mètres et 800-1000 mètres et plus. Le protocole de mesure vise à décrire le peuplement en place au niveau dendrométrique (diamètre, statut social, essence, hauteur) et sanitaire (évaluation selon protocole DEPERISS). La station est décrite au moyen d'un sondage pédologique et via la position topographique. L'exposition, la pente et l'apport en eau sont également évalués.

Nous utilisons ces données pour valider la méthode de détection des scolytes par analyse de série temporelle d'image sentinel-2, bien que la méthode aie été adaptée pour un fonctionnement optimal sur les pessières Wallonnes. En effet, les fluctuations normales de l'indice spectral utilisé ont été calibrées sur des pessières Wallonnes (voir 2.1).

Les 127 placettes du CRPF sont passées en revue de manière minutieuse pour appréhender les données et corriger certaines incohérences. En effet, il est dans certains cas difficile de distinguer une situation de dépérissement dû à des conditions stationnelle et climatique contraignante d'un dépérissement dû à une attaque de scolyte. Par exemple, un examen attentif de la platte numéro 92, pour laquelle le commentaire relevé sur le terrain indique "non scolyte mais stressé", nous a permis de confirmer que cette placette ne présente pas de symptôme du au scolyte. Par ailleurs, une placette d'épicéa sain peut-être installée juste à coté d'un peuplement scolyté (ex placette numéro 43). C'est une situation ambiguë, étant donné que la position du centre de placette est relevée sur le terrain au moyen d'un GPS doté d'une précision de l'ordre de 10 à 20 mètres. La probabilité de localiser la placette saine numéro 43 sur un pixel scolyté étant grande, cette placette est retirée du jeu de donnée. Seul les placettes situées dans la tuile T32ULU sont utilisées pour la validation, ce qui nous amène à retirer 6 placettes. Enfin, les pessières sont localisées au moyen de la classe "sapin-épicéa" de la BD forêt de l'inventaire forestier Français (IGN France). Certaines placettes sont en dehors de cette classe "sapin-épicéa" et sont donc également mise de coté.

Un total de 109 placette sont utilisées pour la validation. Elles comprennent 58 placettes de pessières saines et 54 pessières scolytées. Les pessières scolytées ayant subies une coupe sanitaire se voient attribuer la classe numéro 4, alors que les autres sont classées avec le code 2 (voir 2.2). Ces classes étant dérivées de l'observation de terrain, nous les renseignons sous la terminologie

état sanitaire terrain. L'état sanitaire en 2020 est extrait de la carte d'état sanitaire : il s'agit à contrario de *l'état sanitaire SIG*. De plus, la distance au foyer de scolyte le plus proche est également calculé, ainsi que les proportion de surface scolyté, coupée et en coupe sanitaire sur une vignette de 7 pixels de large (soit 70 mètres de coté, donc 35 mètres de distance du centre de placette dans toutes les directions). La distance au foyer le plus proche est utilisé pour changer l'état sanitaire de certaines placettes ; en effet, une placette située à une distance de maximum 20 mètres d'un foyer de scolytes se voit attribuer une classe d'état sanitaire 2 ou 4, en fonction de la majorité en surface des pixels du voisinage.

5.2.1 Résultats

La matrice de confusion présentée figure 12 montre que la concordance entre les observations terrains et la détection des pessières scolytées par télédétection est satisfaisante. Bien que la précision globale de classification n'exède pas 69.6%, la confusion la plus courante concerne les classes 2 et 4. En effet, tantôt des pessières scolytées sont détectées par télédétection comme étant des coupes sanitaires (n=7), tantôt des coupes sanitaires sont détectées par télédétection comme étant des pessières scolytées (n=13). Il va de soit qu'une globalisation de ces deux classes en une seule classe *pessières scolytées ou coupe sanitaire* ne porte pas préjudice aux objectifs poursuivis par cette recherche, c'est-à-dire la détection des dégâts provoqués par l'attaque du typographe.

On notera que le nombre de faux positifs, des placettes détectées comme étant scolytées alors qu'il s'agit de peuplements sains, est ici de 0. Il est évident que le taux de faux positif est supérieur à 0%, mais cette validation illustre le coté conservateur de la méthode : elle privilégie un nombre d'omissions élevés (n=5 dans le tableau) et un nombre de faux positif faible (nb=0), afin de garantir au mieux que les pixels détectés comme scolytés le soient le plus souvent possible.

Enfin, la matrice de confusion met en lumière le fait qu'une proportion considérable de coupes sanitaires sont détectées comme étant des coupes normales (n=11). La méthode de détection des scolytes nécessite en effet de disposer de minimum 2 images satellites sur lesquelles on détecte un stress préalablement à la coupe sanitaire. Si la coupe sanitaire est effectuée très rapidement après l'attaque de scolyte, ou si des conditions d'enneigement prolongés ne permettent pas d'observer la pessière au moment du stade vert du dépérissement, la méthode échoue.

		Terrain		
		1	2	4
SIG	1	56	5	1
	2	0	9	4
	3	2	3	11
	4	0	7	13
	6	0	1	0

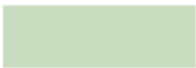
 classification correcte

Figure 12: Matrice de confusion illustrant la confrontation des données terrains de présence/absence de scolyte avec les résultats de la méthode de détection par analyse de série temporelle d'images satellites. Les observations par télédétection montrent une bonne concordance avec les données terrains.