

# Quelles utilisations de l'imagerie satellitaire pour le suivi de la végétation dans les emprises SNCF ?

Anne-Sophie Onody<sup>1</sup> Fatoumata Bah<sup>2</sup> Grégory Duval<sup>2</sup> Hervé Nicolas<sup>1</sup> Pauline Dusseux<sup>1</sup>

<sup>1</sup> Agrocampus Ouest

<sup>2</sup> SNCF Réseau, Direction Ingénierie et Projets (I&P LVE ENV)

anne-sophie.onody@agrocampus-ouest.fr

## 1 Introduction

SNCF réseau est le gestionnaire du réseau ferré français. A ce titre, il entretient et modernise l'infrastructure ferroviaire tout en assurant la circulation des trains. Dans la cadre des missions d'entretien du réseau, la maîtrise de la végétation représente l'un des principaux postes de dépense. Cette maîtrise de la végétation est une nécessité aussi bien pour la conservation des ouvrages que pour la sécurité des circulations ferroviaires et des cheminements le long des voies [9].

Les 95 000 ha de propriété foncière de SNCF réseau nécessitent des outils industriels pour programmer les interventions de maintenance de la végétation sur les voies et les abords du réseau. Dans ce cadre, l'objectif de ces recherches est d'évaluer l'utilisation de l'imagerie satellitaire pour classer les végétations présentes dans les emprises des voies ferrées selon leur type de formation (herbacée, arbustive et arborée) en vue d'améliorer la programmation et la réalisation des travaux.

## 2 Méthodologie

### 2.1 La zone d'étude

L'intérêt de l'étude est d'obtenir un pourcentage de couverture de chaque strate de végétation (absence, herbacée, arbustive et arborée) dans trois zones différentes constituant les emprises SNCF: voie et piste, bande de proximité et abords. Cette étude est menée sur la LGV Nord et a pour objectif d'être reproductible sur d'autres lignes grande vitesse dans un premier temps. La reproductibilité sur les lignes classiques sera testée ultérieurement.

### 2.2 Les données de télédétection utilisées

Les images satellitaires utilisées sont des images Très Haute Résolution Spatiale (THRS) stéréo Pléiades de résolutions spatiales de 0.5m et 1.5m respectivement dans le panchromatique et le multi-spectral et composées de 4 bandes spectrales (Rouge, Vert, Bleu et Proche Infra Rouge). Deux acquisitions de la région lilloise sont utilisées ici, une première datant du 17 février 2016 et une seconde datant du 14 février 2017.

### 2.3 Les traitements effectués

Les images sont reprojctées en Lambert 93 et les comptes numériques transformés en réflectance exo-atmosphérique avant d'effectuer une fusion entre l'image panchromatique et l'image multi spectrale afin d'augmenter la résolution spatiale de l'image multi-spectrale selon la méthode du pan-sharpening.

En plus des quatre bandes spectrales, un ensemble d'indices issus de la bibliographie et identifiés comme d'intérêt pour la discrimination de la végétation sont calculés pour les deux images ; des indices spectraux par calcul sur les différentes bandes spectrales (C3, ExG [7], NDVI, CVI, DVI, OSAVI [10], indice de brillance [2]) ainsi que des indices contextuels de texture selon les méthodes Haralick [4] et SFS (Structural Feature Set) [5] et de morphologie [8] ; ces deux derniers étant calculés à l'aide de la librairie Orfeo Tool Box sur l'image panchromatique.

Un modèle numérique de surface (MNS) est calculé à l'aide du logiciel S2P, un logiciel libre et gratuit permettant une reconstruction 3D automatique à partir d'images stéréo à haute résolution spatiale [3]. Un modèle numérique d'élévation (DEM) en est ensuite dérivé par soustraction entre le MNS et le MNT (Modèle Numérique de Terrain) produit par l'IGN à une résolution de 5m pour l'image de février 2017 et de 25m pour l'image de février 2016. La résolution spatiale du MNT à 5m, inférieure à la largeur des emprises ferroviaires, est jugée comme acceptable pour répondre à la question.

Le calcul du NDVI (Normalized Difference Vegetation Index) et son seuillage empirique permet de distinguer les zones couvertes par de la végétation des zones non-couvertes par de la végétation. [6]. Par ailleurs, un seuillage sur la bande panchromatique permet d'éliminer les zones couvertes par des ombres [1], limitant ainsi les confusions de classification par la suite.

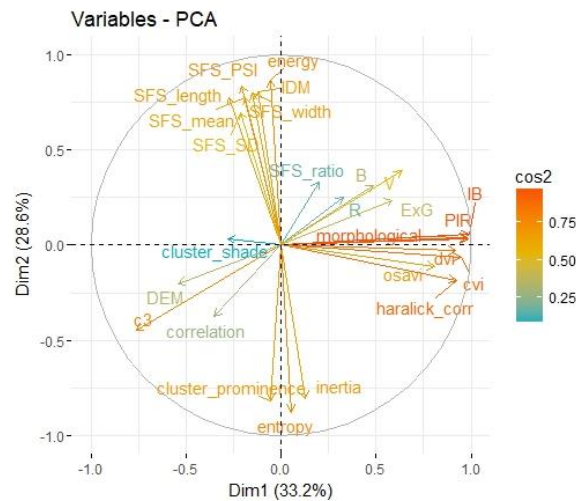
Dans un second temps, un lot d'échantillons est créé par photo-interprétation sur l'image Pléiades de février 2017. Ce lot est constitué de 139 petites zones (surface de 78m<sup>2</sup> chacune) appartenant à quatre types de végétation distincts : arbre, culture, pelouse et mélange arbustes / herbacées hautes repéré principalement sur les talus en bordure de voie ferrée. La pelouse étant difficilement distinguable des cultures du fait de leur stade phénologique

peu avancé au moment de l'acquisition de l'image, les classes « pelouse » et « culture » sont regroupées sous une même classe appelée « herbacée ». Des statistiques zonales sont ensuite calculées sur l'ensemble des variables décrites précédemment avant d'effectuer une ACP (Analyse en Composantes Principales) permettant d'identifier les paramètres qui discriminent le mieux les différentes classes et ainsi de les conserver pour la suite de l'étude. Cette sélection de variables en amont permet de limiter les temps de calcul de la classification dans un souci de reproductibilité à grande échelle. Les variables retenues sont rassemblées afin de former une image multi-bande. Un modèle de classification supervisée selon la méthode Random Forest est entraîné en utilisant le jeu de données séparé au préalable en échantillons d'entraînement et de validation (75% - 25%) afin d'évaluer la classification. Le modèle précédent est utilisé pour prédire et classer l'image de février 2017 et l'image de février 2016. Un deuxième lot de 81 échantillons est créé par photo-interprétation de l'image de février 2016 et des matrices de corrélation sont ensuite calculées pour estimer la validité et la robustesse de l'application du modèle de 2017 sur une image de 2016.

### 3 Résultats

#### 3.1 Les variables retenues

D'après l'ACP réalisée sur l'ensemble des informations extraites de l'image de février 2017 et présentée en figure 1, les variables retenues pour entraîner le modèle de classification sont sélectionnées selon leur poids respectifs sur chacun des axes et sont détaillées dans le Tableau 1.



**Figure 1** - Analyse en Composantes Principales réalisée sur le jeu d'échantillons de l'image de février 2017 et l'ensemble des variables extraites

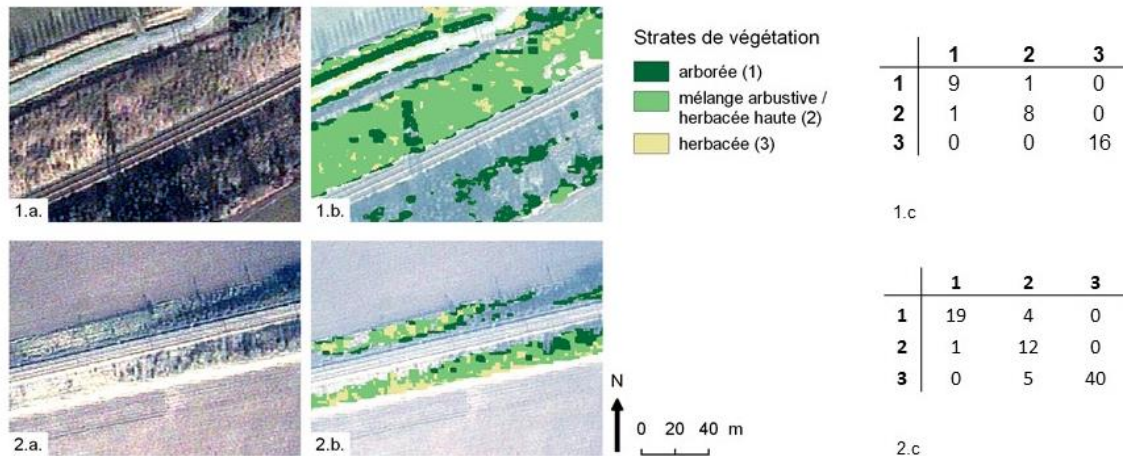
**Tableau 1** - Variables sélectionnées pour la réalisation des classifications de la suite de l'étude

	Indices spectraux	Indice de texture	Autres
<b>Axe 1</b>	C3 IB	Haralick correlation	DEM morphological
<b>Axe 2</b>		Energy Entropy SFS mean SFS PSI	
<b>Axe 3</b>	ExG		

#### 3.2 Résultats de classification

Les premiers résultats de classification par Random Forest montrent une bonne corrélation (accuracy : 94%, kappa : 0.91) entre les échantillons de calibration et de validation de l'image de 2017 ainsi que de bons résultats lors de l'application du modèle à l'image de février 2016 (accuracy : 87% et kappa : 0.79). Les confusions se font

principalement entre la strate arbustive et les deux autres strates (arborée et herbacée) comme présenté dans la figure 2.



**Figure 2** - Extraits des images Pléiades fusionnées de février 2017 et février 2016 (1.a., 2.a. respectivement), du résultat de la classification Random Forest réalisée sur ces images (1.b. et 2.b.) et les matrices de corrélation respectives (1.c. et 2.c.)

### Conclusion et perspectives

Les premiers résultats obtenus montrent que la fine résolution spatiale et l'information fournie par la stéréoscopie des images Pléiades permettent de bien distinguer trois strates végétales différentes par classification supervisée sur un ensemble de variables à une date donnée. L'extrapolation du modèle à une autre image de la même saison fournit également des résultats encourageants. Les résultats de la classification sont ensuite utilisés pour déterminer le pourcentage de couverture de chaque classe dans les polygones d'emprises SNCF.

Les images utilisées dans ce début de recherche sont des images hivernales, qui ne sont pas optimales pour la détection de la végétation, d'une part à cause de la faible activité chlorophyllienne des plantes et d'autre part par la forte présence d'ombres portées, due à une élévation faible du soleil, qui diminuent le nombre de pixels clairs. L'application de cette méthodologie sur des images estivales devrait permettre une meilleure distinction des différentes formations végétales en palliant les inconvénients cités précédemment. Enfin, la récolte de vérités terrain et de données de topographie in situ devraient permettre d'améliorer la précision de la classification et d'évaluer la précision du MNS calculé.

### Références bibliographiques

- [1] K. R. M. Adeline, M. Chen, X. Briottet, S. K. Pang, et N. Paparoditis, Shadow detection in very high spatial resolution aerial images: A comparative study, *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 80, p. 21-38, juin 2013.
- [2] Drone Imaging, Indice de brillance.
- [3] C. de Franchis, E. Meinhardt-Llopis, J. Michel, J.-M. Morel, et G. Facciolo, An automatic and modular stereo pipeline for pushbroom images, *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. II-3, p. 49-56, août 2014.
- [4] R. M. Haralick, Statistical and structural approaches to texture, *Proceedings of the IEEE*, vol. 67, n° 5, p. 786-804, 1979.
- [5] X. Huang, L. Zhang, et P. Li, Classification and Extraction of Spatial Features in Urban Areas Using High-Resolution Multispectral Imagery, *IEEE Geoscience and Remote Sensing Letters*, vol. 4, n° 2, p. 260-264, avr. 2007.
- [6] C. Iovan, D. Boldo, et M. Cord, Automatic Extraction of Urban Vegetation Structures from High Resolution Imagery and Digital Elevation Model, in *Image analysis: 15th Scandinavian conference, SCIA 2007, Aalborg, Denmark, June 10-14, 2007: proceedings*, Springer, 2007, p. 858-867.
- [7] T.-T. Ngo, C. Collet, et V. Mazet, Détection simultanée de l'ombre et de la végétation sur des images aériennes couleur en haute résolution, *Traitement du signal*, vol. 32, n° 2-3, p. 311-333, oct. 2015.
- [8] M. Pesaresi et J. A. Benediktsson, A new approach for the morphological segmentation of high-resolution satellite imagery, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, n° 2, p. 309-320, févr. 2001.
- [9] SNCF RESEAU, IN7136 Gestion des dépendances vertes, in *Référentiel d'infrastructure*, 2012.
- [10] A. Tong et Y. He, Estimating and mapping chlorophyll content for a heterogeneous grassland: Comparing prediction power of a suite of vegetation indices across scales between years, *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 126, p. 146-167, avr. 2017.