Automatic mapping of urban tree species based on multi-source remotely sensed data

Josselin Aval

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Institut Supérieur de l'Aéronautique et de l'Espace (ISAE)*

Présentée et soutenue le *25/10/2018* **par :**

Josselin AVAL

Automatic mapping of urban tree species based on multi-source remotely sensed data

---

JURY

Paolo GAMBA	Professeur	Rapporteur
Florence TUPIN	Professeure des universités	Rapporteure
Christiane WEBER	Directrice de Recherche	Examinatrice
Anne PUISSANT	Professeure des universités	Présidente du Jury
Xavier BRIOTTET	Directeur de Recherche	Directeur de thèse
Emmanuel ZENOU	Maître de Conférences	Co-directeur de thèse
Sophie FABRE	Chargée de Recherche	Co-encadrante
David SHEEREN	Maître de Conférences	Co-encadrant

École doctorale et spécialité :
AA : Aéronautique Astronautique
Unité de Recherche :
ONERA/DOTA - ISAE-SUPAERO/DISC
Directeur(s) de Thèse :
Xavier BRIOTTET et *Emmanuel ZENOU*
Rapporteurs :
Paolo GAMBA et *Florence TUPIN*

Document language

This PhD document is written in English. According to the Higher Institute of Aeronautics and Space Engineerings (ISAE-SUPAERO) policy regarding the language of the PhD documents, each chapter is also summarized in French.

Acknowledgements

With the many things that happen during a thesis, I know that I cannot be exhaustive. These are key elements for me...

I would like first to thank my supervisors. Thanks to Sophie for always being there when I needed it, for her rigour which helped me to improve the quality of my researches. Thanks to Emmanuel for our discussions about the unified frameworks, for allowing me to teach in the ISAE-SUPAERO. Thanks to David and Mathieu for their significant contribution, while it was not necessarily planed. Thanks to David for his feedbacks on the papers, for his original ideas. Thanks to Mathieu for his rigour on the methodological aspects, for allowing me to review some articles. Thanks to Xavier who has always took his role, from the supervision of the works at the more accurate level, to his proposals for helping me in getting more experience. A particular thank to Sabine and Emmanuel who initiated this thesis based on a need for Toulouse city. Thanks to all of you for our efficient meetings!

Special thanks to Jean for your work on street trees at ISAE-SUPAERO. What a beautiful complicity between us for this work and besides, I think of basketball, ants, etc.

Thank you Manon for your unconditional support, for your curiosity about my thesis. We discussed several times in detail the definition of my research axes, and it helped me a lot. You have always been there, when I wasn't confident, even when you had your own problems, and I will always be grateful for that.

Thanks to my family, my mother Annette and Serge, my father Gérard and Sophie, who know everything of the three parts of my work. Thanks to my beloved brothers Nicolas and Romain for their support. Thanks to my grandmother Jeannine for her constant force of thought. I think of course about my grandfather Pierre whose neurodegenerative disease made me think about the meaning of life and the communication between humans. There are many researches to carry out in these fields. Whereas I could talk with him at the beginning, I speak to him now. But life goes on everywhere, my nephews, Léontine, Chloé, Nathan and Marguerite, are the best examples!

I met several people during this experience. Thanks to Thierry for all, how hard it was when your chair became empty. Thanks to Karine and Étienne for our incredible squash sessions. Thanks to Laure for her kindness. Thanks to Bertrand for his candour. Thanks to Jean, Éric, Guillaume and Guillaume, Florian, Simon, Sandra, Antoine, Jan, Vincent, Nicolas, Yousra, Clémence, Damien and many others for these beautiful moments of shar-

ing. Focusing on the DOTA, I have at least an anecdote with every people of the department. It is not the purpose to detail it here and I intend to remember them with you directly!

Finally, I would like to say a word for Aurélie. Have faith in yourself. The department, and more broadly, needs people like you, able to point out dysfunctions on the human and scientific level, even if it is not pleasant. It only refreshes the landscape. Thank you for all these beautiful moments of sharing and complicity. I will undoubtedly miss you very much.

Funding

This thesis was mainly funded by the French Aerospace Lab (ONERA) and the région Midi-Pyrénées, France. The ISAE-SUPAERO funded two communications, for the 2017 IAC conference, Adelaide, Australia, and for the 2018 GEOBIA conference, Montpellier, France. A master internship was funded by the ISAE-SUPAERO.

Contents

General Introduction

Synthèse en français

Contexte

L'augmentation des zones urbaines se traduit par une densification des villes et un trafic routier accru (Tosics, 2017; Voelcker, 2014), engendrant des phénomènes d'îlot de chaleur urbain et de pollution de l'air (Oke, 2011; Chan and Yao, 2008). L'état de santé des habitants est ainsi dégradé. Entre autres, le cancer du poumon touche de plus en plus d'individus dans certaines villes (Guo et al., 2016). Dans ce contexte, les arbres urbains ont un rôle à jouer. Des études ont montré que les parcs pouvaient constituer des îlots de fraîcheurs et que la qualité de l'air pouvait y être meilleure (Doick et al., 2014; Yang et al., 2005). Aussi, la végétation a un effet relaxant sur les personnes (Tsunetsugu et al., 2013). Par ailleurs, les arbres sont soumis à des conditions difficiles en milieu urbain, notamment les arbres d'alignement fortement taillés et ayant peu d'espace pour se développer. Les gestionnaires des villes ont donc intérêt à utiliser les diverses propriétés des canopées pour concevoir les villes de demain. Ils doivent aussi surveiller l'état de santé des arbres.

Ces éléments montrent que la carte des arbres en milieu urbain (i.e. l'emplacement des troncs, la délinéation des couronnes, l'espèce, la santé, la capacité d'évapotranspiration, etc.), à l'échelle de la ville et mise à jour régulièrement, semble être une information essentielle pour une planification urbaine et un suivi de la végétation efficaces. Il y a deux tâches préliminaires avant de pouvoir vérifier la santé des arbres : la localisation des troncs d'arbres et l'identification des espèces au niveau de l'arbre individuel, les maladies étant liées à l'espèce. C'est l'objet de cette thèse, c'est-à-dire la cartographie des espèces d'arbre en milieu urbain. De nos jours, ce type de procédure est encore effectuée manuellement, via une campagne terrain ou par photo-interprétation, en utilisant Google Earth et des données issues des Systèmes d'Information Géographie (SIG) par exemple. La ville de Toulouse (France) aurait par exemple environ 140 000 arbres répartis sur plus de 100 km^2 selon les estimations des gestionnaires de la végétation urbaine. Une telle procédure n'est

donc pas appropriée parce qu'elle est fastidieuse et ne peut donc pas être réalisée sur de grandes zones.

Potentiel de la télédétection pour surveiller les arbres urbains

Les données de télédétection sont intéressantes pour l'automatisation de la cartographie des espèces d'arbres en milieu urbain. Alors que la délinéation des couronnes d'arbre a été étudiée à partir d'imagerie optique, de Modèles Numériques de Surface normalisés (nMNS), etc. (Zhen et al., 2016), la classification des espèces d'arbres a été explorée en utilisant des données multispectrales, hyperspectrales, etc. (Fassnacht et al., 2016) . En particulier, les données hyperspectrales aéroportées permettent d'extraire des caractéristiques spectrales des couronnes des arbres, liées aux constituants et à la structure du feuillage, donc à l'espèce (Jacquemoud et al., 2009). En outre, des caractéristiques de texture et de structure 3D liées à l'espèce peuvent être extraites à partir de données panchromatiques (PAN) et nDSM (Iovan et al., 2008; Dalponte et al., 2014). Pour effectuer la cartographie des espèces d'arbres en milieu urbain, les approches conçues pour les forêts naturelles sont souvent utilisées (Alonzo et al., 2014; Zhang and Hu, 2012; Liu et al., 2017). Cependant, l'environnement urbain a des caractéristiques spécifiques qui doivent être considérées :

- Grande diversité des espèces
- Différentes structures d'arbres : arbres d'alignement, arbres de parc, etc.
- Divers développements d'arbres pour une même espèce en raison de l'élagage et de conditions spécifiques (peu d'espace pour la croissance, îlot de chaleur urbain, pollution de l'air, etc.)
- Chevauchement modéré des couronnes d'arbres
- Beaucoup d'ombres

Objectifs

L'objectif de cette thèse est de cartographier les espèces d'arbres en milieu urbain à partir de plusieurs sources de données aéroportées (hyperspectrales, PAN, nDSM) et d'informations contextuelles.

I. Évaluer la performance de sources de données aéroportées individuelles hyperspectrales, PAN et nDSM, et la performance des sources fusionnées (chapitre I).

II. Optimiser le schéma de fusion à partir des meilleures sources (chapitre II).

III. Tirer profit des informations contextuelles sur les différentes structures d'arbres en milieu urbain en considérant les données SIG (chapitre III).

Organisation du document

Ce document est organisé suivant les articles qui ont été rédigés pendant le travail de thèse (2 articles acceptés avec révision, liés aux premier (I) et troisième (III) objectifs, et un autre article en préparation concernant le deuxième objectif (II)). Il y a ainsi un chapitre de contribution par article : I, II et III, respectivement. Chaque article est résumé en français au début de chaque chapitre, avant d'être présenté (sections I.1, II.1 et III.1). Le chapitre suivant (A) est l'état de l'art. Il peut y avoir des redondances entre les discussions des articles et la discussion générale présentée à la fin de ce document.

English part

Context

Urban areas

The world urban population is growing rapidly, causing a significant expansion of urban areas. Whereas the proportion of humans living in cities was 2% in 1800s and 29% in 1950s, at least 50% of the humans are living in the urban environment since 2009 (Cunningham, 2018). By 2030, the world urban population will rise to nearly 5 billion, and at the same time the urban land cover will raise by 1.2 million km^2 (Seto et al., 2012). Among others, this expansion of urban areas induces an increase in road traffic and in urban densification. According to (Voelcker, 2014), there will be 2 billion vehicles by 2035 against 1.2 billion in 2014. Only 2.5% of those 2 billion will be battery electric while 8% will be hybrid-electric or natural-gas powered. About the densification, the example of the dynamically expanding Amsterdam (The Netherlands) can be cited, with its obligation to build 300 thousand new housing units by 2040 (Tosics, 2017).

As already observed, these important changes will cause in a more significant way the air pollution of the cities (Chan and Yao, 2008) and the urban heat island phenomenon (Oke, 2011) (Figure 1). These effects have many consequences on our health, but also at a more global scale on the Earth, through climate changes which will modify the living conditions of humans and other species (Seinfeld and Pandis, 2016). Focusing on the state of health of the inhabitants, the results are not encouraging. The number of people with respiratory and cardiac diseases is increasing, for example lung cancer in China whose rate was 47.5 per 100,000 for men and 22.2 per 100,000 for women in 2009, and expected to rise because of the increasing air pollution (Guo et al., 2016). People can suffer from serious illnesses such as heat exhaustion and heat stroke when they are exposed to extreme heat, in particular older adults and young children (EPA, 2017). The urban heat island phenomenon contributed around 50% of the total heat-related mortality during the 2003 heatwave in the West Midlands, United Kingdom, according to (Heaviside et al., 2016).

In order to deal with these major issues, one solution is to reduce air pollution and heat island effect.

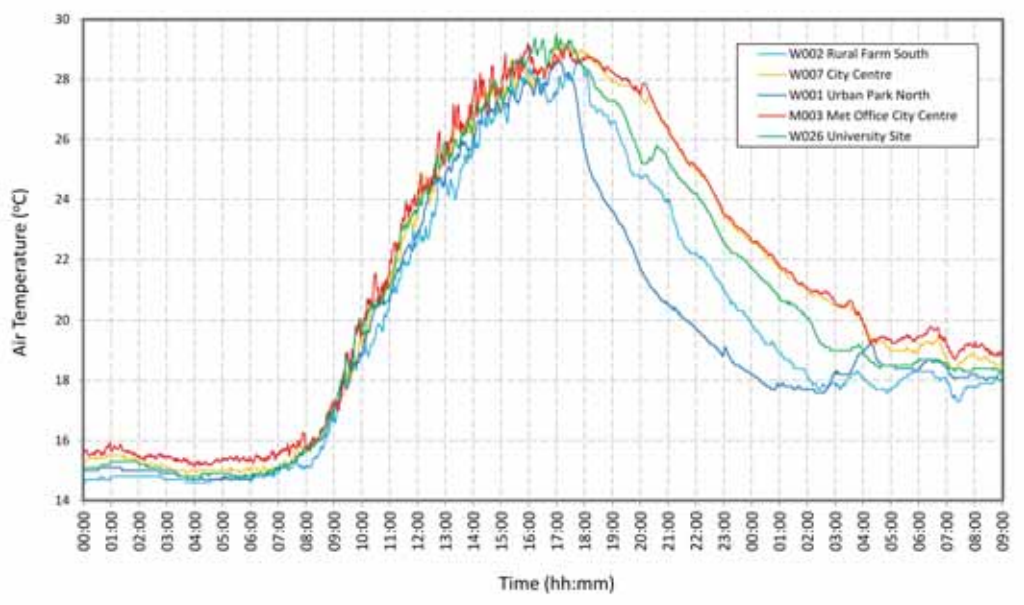

Figure 1: Illustration of the urban heat island effect at Birmingham, United Kingdom (from (MetLink, 2017)). This is the development of Birmingham's urban heat island on the night of the 22nd July 2013, during a heat wave. Especially in the evening, the air temperature in the city centre remains higher than in an urban park (around 5 °C at 10 pm), for example, because the densely packed buildings in the centre retain more heat.

Urban canopies

Urban trees contribute to several ecosystem services (Jones, 2014). Tree infrastructures, especially park trees, can locally improve the air quality (Yang et al., 2005) (Figure 2) and decrease the air temperature (Doick et al., 2014) during heatwave in dense and polluted cities. As an example, the study of (Yin et al., 2011) aims at quantifying air pollution attenuation within urban parks in Shanghai, China. After monitoring the sulfur dioxide (SO_2) and the nitrogen dioxide (NO_2) from six parks in Pudong District, they demonstrate that park vegetation can remove large amount of airborne pollutants: 9.1% of Total Suspended Particles (TSP) removal in summer. As another example, the work of Doick et al. (2014) studies the role of one large greenspace in mitigating London's nocturnal urban heat island. They conclude that trees help regulate air temperature and combat the urban heat island effect, with an observed cooling of up to 4 °C over 400 m distance from the park on single nights. All these properties depend on the tree species as highlighted by

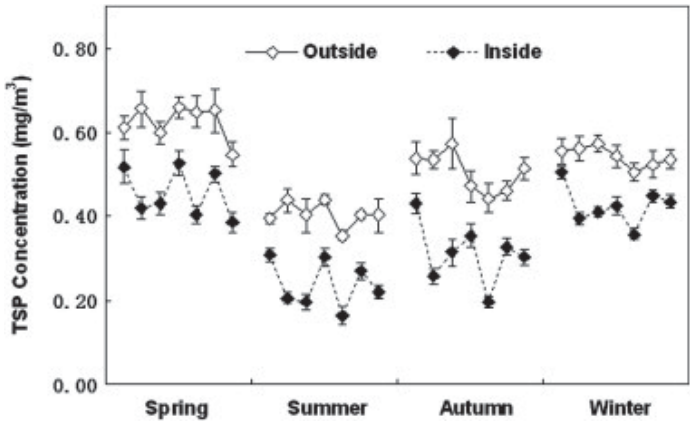

Figure 2: Illustration of the improvement of air quality thanks to park trees. TSP stands for Total Suspended Particles. While "Outside" refers to outside the park, "Inside" refers to inside the park (from (Yin et al., 2011)).

Figure 3.

Moreover, urban trees promote biodiversity, have a relaxing psychic action and contribute to aesthetics (Chiesura, 2004). As an example, the study of (Tsunetsugu et al., 2013) characterizes the physiological and psychological effects of viewing urban forest landscapes in Japan. They find that the subjects exhibited significantly lower heart rate and diastolic blood pressure in the forested areas. Following on from this phenomenon, the work of (Song et al., 2013) quantifies the effects of walking on young people in urban parks during winter. The heart rate was significantly lower. Walking in the urban park improved mood and decreased negative feelings and anxiety according to the results of questionnaires proposed to the subjects.

There is no ambiguity on the fact that the urban managers have an interest in taking advantage of the many properties of urban canopies to design future cities: air quality improvement, heat island effect reduction, relaxing psychic action, etc.

But urban trees suffer from hard conditions in the urban environment. In particular, they can be affected by specific diseases depending on species. For example, *Ceratocystis platani* is a fungus responsible for the Canker stain of *Platanus* trees. In Europe, it was allegedly introduced to Marseille, France in 1945 from infested wooden crates of US troops containing military equipment (Vigouroux, 2014). At Forte dei Marmi in Italy, 90% of

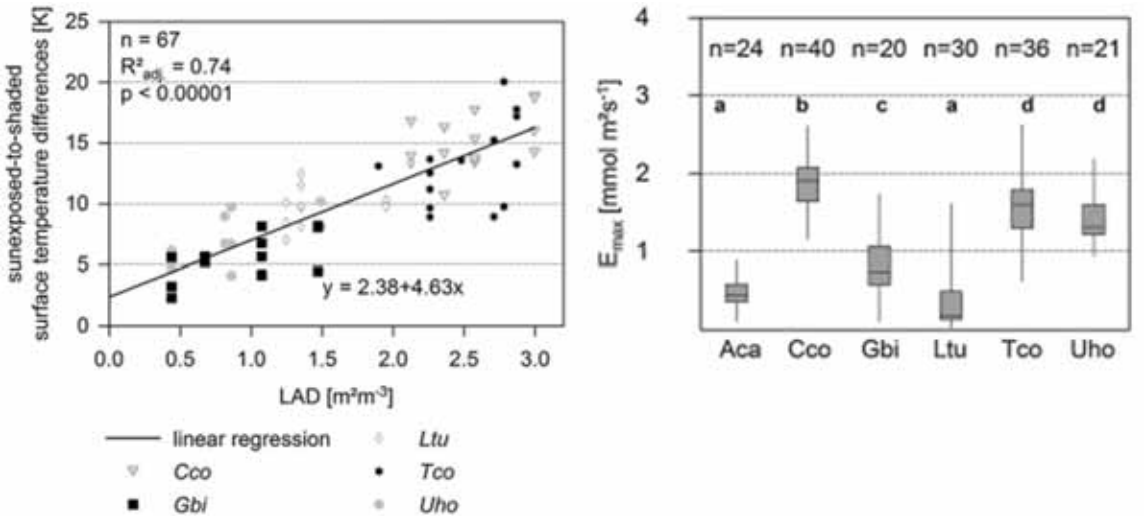

Figure 3: Freshness island properties of the trees in function of the LAD (Leaf Area Density in this particular context) (left) and the species (right). E_{max} stands for the rate of transpiration (from (Gillner et al., 2015)).

the *Platanus* trees died in 1972-1991. As another example, *Aesculus hippocastanum* can be affected by the horse-chestnut leaf miner (Paterska et al., 2017) which necroses its foliage, making it characteristic (Figure 4). Indeed, the leaf miner, which can more easily travel in urban areas, attacks the parenchyma of the leaf. In cities, trees are often planted as alignment trees and belong to the same tree species, thus easing the transmission.

Focusing on the alignment trees (or street trees, Figure 5), they have little space for growth and are pruned, most often to be adapted to the constraints of the sites. As a case in point, a pruned lime tree (*Tilia*) has a life expectancy of 150 years against 800 years without constraint (Baraton, 2014; Fini et al., 2015). In order to highlight the crucial place of these trees in the urban environment, the example of Paris, France can be cited with nearly 100,000 alignment trees (about half of the trees). They cover around 700 km of roads and concern approximately 1600 roads out of 6000.

These examples demonstrate that the urban managers have to take care of the trees by designing an efficient monitoring of their health.

Beside the mentioned elements, the urban tree map (i.e. the location of tree trunks, crown delineations, species, health, capacity of evapotranspiration, etc.), at the city scale, updated

Figure 4: On the left: horse-chestnut leaf miner from (Wikipédia, 2018). In the middle: a leaf attacked by the leaf miner from (Jardinier, 2016). On the right: canopy of an *Aesculus hippocastanum* attacked by the leaf miner.

Figure 5: Illustration of the urban tree structures extracted from (Dépêche, 2013). Example of Toulouse. On the left, the well known *Platanus* alignment trees highly represented in the South of France. On the right, park trees.

with regular time basis, appears to be an essential information for an effective urban planning and vegetation monitoring. There are two preliminary tasks before reaching the health of trees: the tree trunks localization and the species identification at an individual level, the diseases being related to the species. This is the purpose of this PhD thesis, in other

words the urban tree species mapping. Nowadays, this type of procedure is still carried out manually, by field campaign or by photointerpretation, considering Google Earth and Geographic Information System (GIS) data for instance. The city of Toulouse (France), for example, would have approximately 140,000 trees spread over more than 100 km^2 according to estimates of vegetation managers, such a procedure is not appropriate because it is tedious and thus cannot be carried out over large areas.

Remote sensing: a useful tool to monitor urban trees

Remotely sensed data are of interest for automating the urban tree species mapping. Indeed, remote sensing of urban areas and remote sensing of vegetation have been carried out for decades (Jensen and Cowen, 1999; Ustin and Gamon, 2010). Regarding the trees, while tree crown delineation has been studied based on optical imagery, normalized Digital Surface Model (nDSM), etc. (Zhen et al., 2016), tree species classification has been explored using multispectral, hyperspectral, multitemporal data, etc. (Fassnacht et al., 2016). In particular, airborne hyperspectral data allow spectral features of the tree crowns to be extracted, related to the foliage components and structure, and therefore to the species (Jacquemoud et al., 2009). Also, textural and 3D structural features that are related to the species can be extracted at the crown level based on airborne panchromatic (PAN) and nDSM data (Iovan et al., 2008; Dalponte et al., 2014). Focusing on the urban tree species mapping, approaches designed initially for natural forests are often used in the urban context (Alonzo et al., 2014; Zhang and Hu, 2012; Liu et al., 2017). However, the urban environment has distinctive features that have to be considered:

- High species diversity
- Different tree structures: street trees, park trees, etc.
- Various tree developments for a same species due to pruning and specific conditions (little space for growth, urban heat island, air pollution, etc.)
- Moderate overlap between the tree crowns
- A lot of shadows

The latter make the tree species classification difficult, because of the great species diversity on the one hand, and due to the different tree developments (especially between the street trees and the others) on the other hand, increasing the intra-species variability. However, the specific distribution of the urban trees, as street trees for instance, constitutes a prior knowledge that could be integrated in the methodologies. Also, the moderate

overlap between the tree crowns makes the crown delineation easier in comparison to dense natural forest conditions. In order to improve the existing urban tree species mapping methods, one way is to be aware of these distinctive features when developing new approaches. In particular, the complementarity of several remote sensing data sources, potentially multitemporal, can be investigated in order to deal with the great species diversity issue, while the contribution of contextual information can be explored to alleviate the high intra-species variability issue.

Objectives

The objective of this PhD thesis is to map urban tree species based on several airborne data sources (hyperspectral, PAN, nDSM) and contextual information.

I. Assess the performance of individual airborne hyperspectral, PAN and nDSM sources, and the performance of the fused ones (chapter I).

II. Optimize the fusion scheme based on the best sources (chapter II).

III. Take advantage of the contextual information about the different urban tree structures considering GIS data (chapter III).

Figure 6 illustrates the contributions of this PhD work.

Thesis organization

This document is organized according to the papers that have been written during the PhD work (2 papers accepted, related to the first (I) and third (III) objective, and another paper under review concerning the second objective (II)). There is then one contribution chapter per paper: I, II and III, respectively. Each paper is summarized in French in the beginning of each chapter, before being introduced (sections I.1, II.1 and III.1). The next chapter (A) is the state-of-the-art. There can be some redundancies between the discussions of the papers and the General discussion presented at the end of this document.

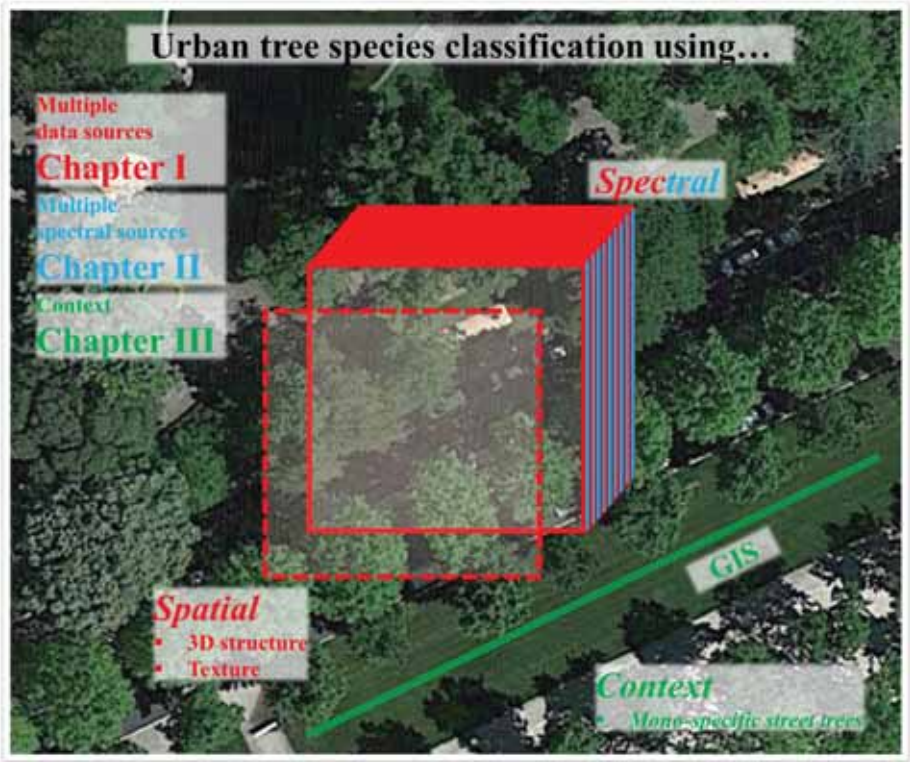

Figure 6: Overall diagram of the PhD thesis contributions.

Chapter A

State-of-the-art

Synthèse en français

Pour la cartographie des espèces d'arbres en milieu urbain, il existe des approches orientées pixel où chaque pixel des images est classé, et des approches orientées objet où chaque objet détecté (arbre individuel, amas d'arbres, etc.) est classé. Nous nous focalisons sur la classification supervisée parce que certaines espèces sont connues à l'avance. En ce qui concerne les méthodes orientées objet, nous ne considérons que le niveau de l'arbre individuel pour répondre à l'objectif de cette thèse. Ces méthodes nécessitent que les couronnes d'arbres soient préalablement délinéées avant que les espèces ne soient classées.

Méthodes de délinéation des couronnes

Concernant les méthodes de délinéation, un grand nombre de travaux a déjà été effectué comme l'atteste l'état de l'art de (Zhen et al., 2016). Il existe différents groupes de méthodes. En particulier, le matériel disponible pour cette thèse nous oriente vers l'utilisation de méthodes utilisant des rasters. Les études de (Iovan et al., 2008; Alonzo et al., 2014; Ardila et al., 2012) constituent des exemples d'utilisation de ces méthodes pour la délinéation des arbres urbains, et montrent que des précisions de détections de l'ordre de 80% peuvent être obtenues. Les principales erreurs se produisent dans les cas de chevauchements importants entre les couronnes, un problème bien connu dans la littérature (Zhen et al., 2016). Mais le chevauchement modéré dans les villes par rapport aux forêts naturelles nous pousse à choisir une méthode standard pour effectuer la délinéation des couronnes. Sans différences significatives entre les approches mentionnées, celle développée dans (Iovan et al., 2008) est choisie car elle est efficace en

termes de temps de calcul et facile à utiliser.

Méthodes de classification des espèces

Pour la classification, deux étapes sont en général effectuées: une extraction de caractéristiques et une classification supervisée. Ensuite, il est possible d'adopter une approche de fusion pour améliorer les performances. Un état de l'art des méthodes de classification des espèces d'arbres est fourni dans (Fassnacht et al., 2016). Selon la technologie des capteurs, différentes informations peuvent être utilisées pour classer les espèces. Alors que les capteurs multispectraux ou hyperspectraux permettent de modéliser les caractéristiques spectrales des espèces d'arbres, les caractéristiques spatiales structurales ou texturales peuvent être extraites des données PAN et nDSM, respectivement, mais aussi grâce à des mesures RADAR (Radiation Detection And Ranging). De plus, les caractéristiques temporelles caractérisant la phénologie de la végétation sont accessibles à partir de séries temporelles. Des informations contextuelles peuvent également être intégrées. Parce que les données RADAR ne sont pas disponibles pour cette étude, nous ne considérons pas les travaux. L'utilisation des caractéristiques temporelles n'est pas détaillée car le coût d'images adéquates est trop élevé. Premièrement, les mesures d'évaluation de la classification des espèces utilisées dans ce document sont détaillées.

Métriques d'évaluation de la classification des espèces

Afin d'évaluer les résultats des méthodes et de comparer leurs performances, les métriques suivantes qui peuvent être calculées à partir de la matrice de confusion (exemple pour 3 espèces dans la Table A.1) sont habituellement considérées : Précision Globale (OA), κ, Précision du Producteur (PA), la Précision de l'Utilisateur (UA) et le F-score. Nous les utilisons toutes dans le cadre de ces travaux. La matrice de confusion, donc les métriques associées, ne peut évidemment être calculée que si la vérité de terrain est disponible.

L'OA et le κ sont des métriques globales et sont calculées comme suit (Figure A.1) :

- $OA = p_0 = 100 \cdot \frac{TS_1+TS_2+TS_3}{Total}$
- Sachant que $p_e = 100 \cdot (\frac{Total(s_1)}{Total} \cdot \frac{Total(\widehat{s_1})}{Total} + \frac{Total(s_2)}{Total} \cdot \frac{Total(\widehat{s_2})}{Total} + \frac{Total(s_3)}{Total} \cdot \frac{Total(\widehat{s_3})}{Total})$, alors $\kappa = 100 \cdot \frac{p_0-p_e}{1-p_e}$

Dans le cas général de N espèces :

- $OA = 100 \cdot \frac{\sum_{i=1}^{N} TS_i}{Total}$

Table A.1: Description de la matrice de confusion (cas de 3 espèces). "Vraie" se réfère à l'espèce d'arbre réellement dans la scène tandis que "Prédite" se réfère à l'espèce d'arbre prédite par la méthode considérée. Par exemple, TS_1 (ie "espèce vraie 1") fait référence aux cas où la méthode prédit l'espèce 1 ($\widehat{s_1}$) alors que l'espèce réelle est bien l'espèce 1 ($\widehat{s_1}$) (c'est-à-dire une vraie prédiction). Comme autre exemple, FS_3 correspond aux cas où la méthode prédit à tort l'espèce 3 (mauvaise prédiction).

		Prédite			
		$\widehat{s_1}$	$\widehat{s_2}$	$\widehat{s_3}$	**Total**
Vraie	s_1	TS_1	FS_2	FS_3	$Total(s_1)$
	s_2	FS_1	TS_2	FS_3	$Total(s_1)$
	s_3	FS_1	FS_2	TS_3	$Total(s_1)$
	Total	$Total(\widehat{s_1})$	$Total(\widehat{s_2})$	$Total(\widehat{s_3})$	$Total$

- $p_e = 100 \cdot \sum_{i=1}^{N} \frac{Total(s_i)}{Total} \cdot \frac{Total(\widehat{s_i})}{Total}$

La PA, l'UA et le F-score sont des métriques spécifiques à chaque espèce et sont calculées de la manière suivante, par exemple pour la ième espèce (s_i) :

- $PA_{s_i}(\%) = 100 \cdot \frac{TS_i}{Total(s_i)}$
- $UA_{s_i}(\%) = 100 \cdot \frac{TS_i}{Total(\widehat{s_i})}$
- $F - score_{s_i}(\%) = 100 \cdot \frac{2 \cdot PA_{s_i} \cdot UA_{s_i}}{PA_{s_i} + UA_{s_i}}$

Extraction de caractéristiques et classification supervisée

Pour l'extraction de caractéristiques, la transformation Minimum Noise Fraction (MNF) est choisie pour extraire les caractéristiques spectrales parce que c'est une approche bien connue qui est efficace à la fois pour améliorer la performance par rapport à l'utilisation directe de la réflectance spectrale, et réduire la dimension des données (Fassnacht et al., 2014; Ghosh et al., 2014a). Par ailleurs, l'utilisation de caractéristiques structurales pour modéliser avec précision la structure 3D des espèces d'arbres est plutôt récente et encourageante (Alonzo et al., 2014; Dalponte et al., 2012). Par conséquent, des rapports de hauteur similaires à ceux développés dans (Alonzo et al., 2014; Dalponte et al., 2012) et adaptés aux données nDSM sont considérés pour l'extraction des caractéristiques structurales. En outre, les paramètres de Haralick sont couramment utilisés dans la littérature

pour modéliser la texture des espèces d'arbres et ont démontré des résultats encourageants (Franklin et al., 2000; Iovan et al., 2008). Ils sont donc sélectionnés pour l'extraction de la texture. Pour la classification supervisée, des échantillons d'apprentissage directement extraits des images et fondés sur des mesures spectrales terrain sont considérés (Alonzo et al., 2014; Nidamanuri and Zbell, 2011). Dans d'échantillons d'apprentissage terrain, les indices spectraux sont préférés pour faire face au changement d'échelle (Cho et al., 2008). Enfin, les algorithmes Support Vector Machine (SVM) et Random Forest (RF) sont choisis parce qu'ils ont déjà démontré de bonnes performance dans la littérature (Féret and Asner, 2013).

Fusion

Chaque ensemble de caractéristiques (spectrales, structurales, texturales, etc.) contribue à l'identification de l'espèce, il y a donc intérêt à les combiner pour obtenir des cartes d'espèces d'arbres plus précises. Une manière habituelle consiste alors à combiner plusieurs sources de données / caractéristiques / algorithmes de classification, c'est-à-dire à considérer un cadre de fusion. Cependant, le terme fusion est ambigu car différentes "choses" peuvent être fusionnées. Pour éviter toute ambiguïté, nous avons résumé les stratégies de fusion possibles en fonction des informations à fusionner dans la Figure A.1. Deux cadres de fusion peuvent être utilisés: fusion de niveau des caractéristiques (Alonzo et al., 2014; Dalponte et al., 2012) et fusion au niveau de la décision (Stavrakoudis et al., 2014; Engler et al., 2013). Dans notre contexte, la fusion du niveau de la décision et l'ensemble de classifieurs se réfèrent à la même méthodologie. D'une part, la fusion au niveau des caractéristiques concatène de nombreuses caractéristiques d'intérêt et le vecteur de caractéristiques résultant est classé. D'autre part, la fusion de niveau de la décision considère plusieurs classifieurs (extraction de caractéristiques et classification supervisée comme montré dans la Figure A.1), selon un critère, et les prédictions de ces classifieurs sont combinées à travers une règle de décision. En se concentrant sur la Figure A.1, nous avons distingué deux catégories de fusions au niveau de la décision: les approches par sources de données et les approches par classifieurs. Pour la première, un classifieur est défini pour chaque source de données alors que les classifieurs sont choisis selon un autre critère pour la deuxième (par exemple donnant le même ensemble de caractéristiques mais différents algorithmes de classification supervisée, le même algorithme mais différents ensembles de caractéristiques, etc.). Cette dernière approche est donc très flexible. En ce qui concerne le contexte de classification des espèces d'arbres, la fusion au niveau des caractéristiques est toujours préférée pour sa simplicité à implémenter sauf dans les études de (Stavrakoudis et al., 2014) et (Engler et al., 2013) où des fusions au niveau de la décision sont considérées, par source de données et par classifieur respective-

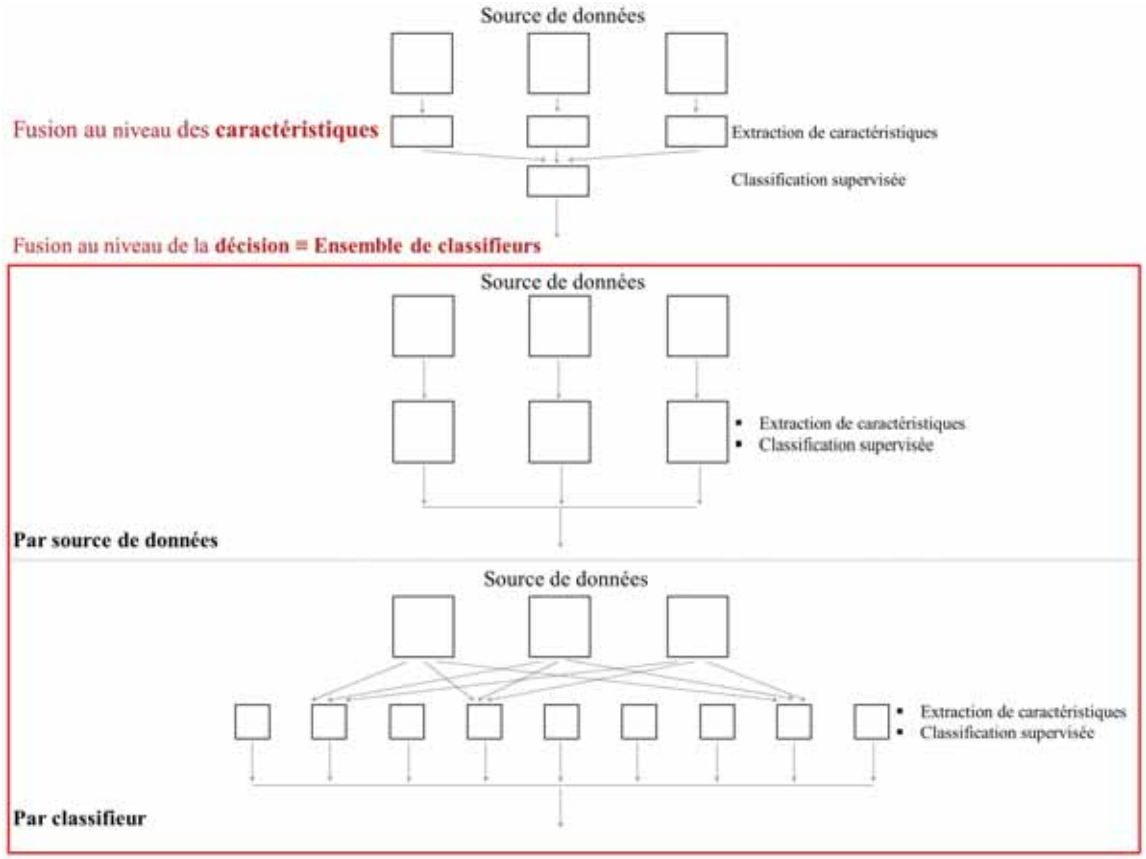

Figure A.1: Définition des stratégies de fusion. Pour la fusion au niveau de la décision par classifieur, toutes les flèches ne sont pas dessinées afin de ne pas surcharger le diagramme.

ment.

La combinaison de l'information spectrale et structurale a démontré un intérêt pour la classification des espèces d'arbres, mais il n'y a pas d'amélioration significative (Alonzo et al., 2014; Dalponte et al., 2012). La contribution de l'information texturale suit le même comportement (Franklin et al., 2000), sauf qu'il y a beaucoup moins d'études pour ce type de fusion. Il n'y a pas de travail qui fusionne l'information spectrale, structurale et texturale, alors que cela conduirait à de meilleurs résultats, puisque des améliorations ont été indiquées pour les combinaisons spectrales / structurales et spectrales / texturales, bien que légères. De plus, la complémentarité de ces caractéristiques n'a pas été évaluée, alors

que des sources non complémentaires ne peuvent logiquement pas améliorer la performance de la meilleure d'entre elles. Alors que la fusion au niveau des caractéristiques est principalement utilisée dans la littérature (Fassnacht et al., 2016), la fusion au niveau de la décision a déjà démontré son potentiel (Stavrakoudis et al., 2014; Engler et al., 2013), et aucune comparaison n'a été effectuée afin de sélectionner la meilleure approche pour un cas de classification d'espèces d'arbres donné. Cela nous amène à identifier la meilleure stratégie de fusion orientée objet (au niveau des caractéristique ou de la décision) en tirant profit de la complémentarité de plusieurs sources de données hétérogènes aéroportées pour améliorer la cartographie des espèces d'arbres en milieu urbain. En premier lieu, une fusion au niveau de la décision fondée sur les sources de données (un classifieur par source) semble être un candidat d'intérêt afin d'évaluer la contribution de chaque source de données. Les échantillons d'apprentissage sont directement extraits des images à partir de couronnes délimitées manuellement (chapitre I).

Afin d'améliorer ces méthodes de classification multi-source, la deuxième partie de cette thèse est consacrée à l'exploration du potentiel d'une fusion au niveau de la décision fondée sur des classifieurs. En particulier, il y a intérêt à extraire les caractéristiques de telle sorte qu'elles optimisent la précision de classification, par exemple par espèce si chaque classifieur est dédié à la prédiction d'une espèce particulière. Classiquement, l'apprentissage des modèles de classification est effectué à partir de pixels des images (Fassnacht et al., 2016; Alonzo et al., 2014). Nous allons analyser la qualité des cartes générées quand l'apprentissage est effectué à partir de mesures terrain (aux échelles de la canopée et de la feuille). L'intérêt est d'utiliser l'approche de classification résultante dans un contexte opérationnel où des pixels d'apprentissage ne sont pas disponibles au sein des images (Nidamanuri and Zbell, 2011). Dans le but de s'affranchir du changement d'échelle, introduit quand l'apprentissage est effectué à partir de mesures terrain, les cartes obtenues en utilisant des indices de végétation sont évaluées (Cho et al., 2008). Par conséquent, il est intéressant de développer une approche d'ensemble de classifieurs à partir de mesures spectrales terrain en utilisant des indices de végétation (chapitre II).

D'autres améliorations des méthodes proposées pourraient être obtenues en tenant compte des informations contextuelles. La détection des arbres d'alignement permettrait de régulariser la prédiction des espèces d'arbres dans les alignements urbains. Deux cadres sont possibles pour cette détection: méthode Individual Tree Crown Delineation and Detection (ITCD) classique + information *a priori* (Koch et al., 2014), et les cadres unifiés (Perrin et al., 2006). En particulier, les cadres unifiés ne présentent pas les inconvénients des approches classiques en termes de diffusion des erreurs. Deuxièmement, les cadres unifiés conduisent à des performances similaires. La méthode du contour actif est

préférable pour une délinéation précise des couronnes, mais n'est pas appropriée pour modéliser l'interaction entre les arbres (organisation spatiale et caractéristiques communes parmi les arbres), une composante essentielle pour la détection des arbres d'alignement. Ainsi, les Marked Point Process (MPP) sont sélectionnés. Le travail de Wen (Wen et al., 2017) au niveau de l'amas d'arbres n'est pas compatible avec le MPP au niveau de l'arbre individuel mais certaines des caractéristiques utilisées pour discriminer les arbres d'alignement pourraient être considérées, en particulier la distance aux routes. Ceci nous conduit à développer une méthode MPP de détection des arbres d'alignement à partir de données aéroportées et d'informations contextuelles (chapitre III).

English part

Remote sensing introduction

Urban areas

Remote sensing opens the way to automate urban tree species mapping thanks to its high spatial resolution available from the past and future generations of sensors. The spatial resolution is a key parameter for monitoring urban areas made of complex arrangements of objects (buildings, roads, vegetation, mixture, etc.). In a perspective of individual tree mapping, the spatial resolution has to be much smaller than the size of the tree crowns (diameter ranging from a few meters up to more than 15 meters), otherwise the individuals are not discernible. Today, satellites such as GeoEye-1, QuickBird, WorldView 1-4, IKONOS, Pléiades, KOMPSAT or TripleSat provide images with a spatial resolution of around 1 m or finer (INFO, 2018). GeoEye-1 has for example a spatial resolution of 46 cm in panchromatic (PAN) mode, 1.84 m in multispectral mode (Corporation, 2018). Airborne systems followed by Unmanned Aerial Systems (UAS) give logically access to better spatial resolution (e.g. 20 cm in the work of (Feng et al., 2015) based on high-resolution UAS imagery), but the images cover is smaller. The spatial resolution available from remote sensing sensors is therefore suitable for characterizing the urban trees at an individual level.

Many researchers have studied the urban environment thanks to remotely sensed data (Gamba et al., 2005; Weber and Puissant, 2003; Puissant et al., 2014). As examples of remote sensing use in the urban context, the satellite data applications of (Hester et al., 2008) and (Shackelford and Davis, 2003) can be cited. The work of (Hester et al., 2008) produces a six-category urban land-cover map from QuickBird imagery. In a per-pixel classification framework, they obtain a performance of nearly 90%. As another example, the study of (Shackelford and Davis, 2003) investigates the potential of a combined fuzzy pixel-based and object-based approach for the classification of IKONOS multispectral data over urban areas. In particular, they aim at discriminating spectrally similar road and building urban land cover classes by using both spectral and spatial information. They get classification maps with an accuracy of 84%. Focusing on UAS-based uses, an urban flood mapping is carried out in (Feng et al., 2015). The results of this study demonstrate that UAS can provide an appropriate platform for the accurate extraction of inundated areas (87% of accuracy) (Figure A.2).

Figure A.2: Flood mapping results obtained in the work of (Feng et al., 2015). Left: study site with visible inundated areas. Right: resulting map with the detected inundated areas colored in red.

Vegetation

The spectral resolution is particularly important when attempting to classify vegetation species or monitor the state of health of vegetation (Ustin and Gamon, 2010). In a perspective of tree species identification, an high spectral resolution measurement is required because of the similarities between the spectral traits of the numerous species under consideration (Dalponte et al., 2009). An hyperspectral system is then a candidate of interest for estimating the species, while previously mentioned satellite devices are less appropriate from the point of view of the spectral resolution. The future HYPXIM instrument will be capable of resolving few hundred of spectral bands (from 0.4 to 2.5 µm), however with a 8-meter spatial resolution (Briottet et al., 2011). The currently operational hyperspectral sensor Hyperion has a spatial resolution of 30 m (USGS, 2011). Knowing that, for a sensor with the same technology an improvement in spectral resolution implies a decline in spatial resolution, airborne hyperspectral sensors have both suitable spatial and spectral resolutions for species classification at the individual tree level.

Spectral imagery of vegetation has been carried out for decades, in particular for estimating biophysical and biochemical parameters of vegetation (Ustin and Gamon, 2010;

Jacquemoud et al., 2009) (Figure A.3). As an illustration of biophysical parameter es-

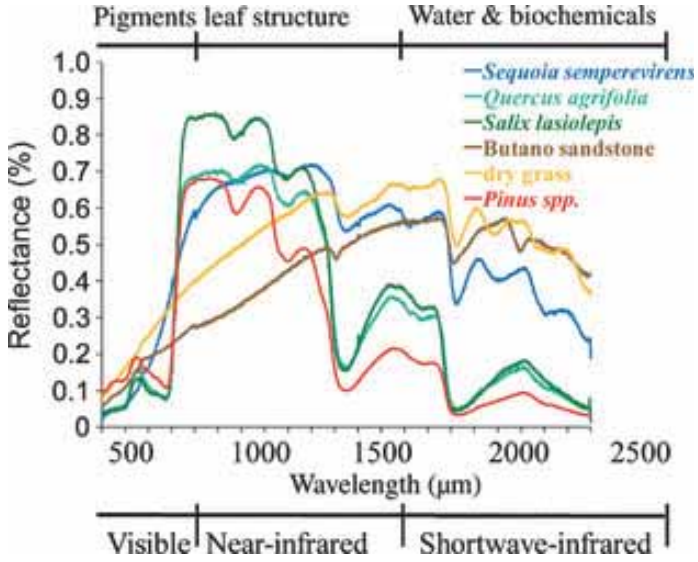

Figure A.3: Spectral reflectance of different live and dry plant foliage and soil (from (Ustin and Gamon, 2010)).

timation, the work of (Asrar et al., 1984) aims at estimating the Leaf Area Index (LAI) from field spectral measurements over wheat canopies, because some plant growth models require estimates of such quantity. Regarding biochemical parameter estimation, the signature analysis of leaf reflectance spectra carried out in (Gitelson and Merzlyak, 1996), aimed at extracting relevant vegetation indices for improving the remote sensing of chlorophyll, is another example. From field level, the results of these studies are intended to be applied on a larger scale thanks to satellite and airborne data (Gong et al., 2003; Smith et al., 2003). These plant parameters depend on the species making it an information of interest in the remote sensing community (Martin et al., 1998). Focusing on the Figure A.3, it is clear that each species has a particular spectral behaviour (spectral features), because of the specific foliar component contents (Jacquemoud and Baret, 1990) and the specific foliage structure (Verhoef, 1984).

Other features can be related to the species such as spatial (textural, structural, etc.), temporal and contextual features. For instance, airborne PAN sensors allow a very high spatial resolution (order of magnitude of 10 cm) measurement of the reflected radiation integrated over the visible spectral range (Iovan et al., 2008) and gives information about the spatial arrangement of the foliage within the tree crowns (textural features), which is also related

to the species (Zhang and Hu, 2012). Moreover, the airborne nDSMs allow a very high spatial resolution measurement of the height (order of magnitude of 10 cm) (Dalponte et al., 2015) and gives information about the 3D structure of trees (structural features), which may differ among species.

Focusing on the urban context, mapping the extents of urban tree canopy using aerial or satellite imagery is currently operational (MacFaden et al., 2012). However, the resulting maps rarely provide information about species, LAI, etc. (Alonzo et al., 2014). This is why urban tree species mapping methods need to be developed, based on the many features mentioned above.

A.1 Urban tree species mapping methods

In this state-of-the-art, we focus on the urban tree species mapping. However, as approaches designed initially for natural forests are often used in urban areas, the tree species mapping in natural forests is also reviewed. Because certain methodologies such as decision level fusion are rarely used in the tree species mapping context, the classification of other objects than the trees is also considered.

For urban tree species mapping, there are pixel-based and object-based approaches, both considered in this state-of-the-art. For pixel-based approaches, each pixel of the images is classified, while each detected object (individual tree, forest stand, etc.) is classified for object-based ones. We focus on supervised classification because some of the species are previously known. Regarding object-based methods, we consider only the individual tree level in order to fit the objective of this PhD thesis. The individual tree-based frameworks, for which the obtained maps are particularly suitable from a user perspective, then requires the tree crowns to be previously delineated before the species are classified. This object-based approach is illustrated in Figure A.4. The first step aims at delineating the tree crowns, the second one is intended to identify the species of the trees, by using the information from within the crowns (spectral, spatial, temporal, etc. information). To summarize in this framework, the species classification step requires the crown delineation knowledge.

Crown delineation methods are presented in section A.1.1 whereas species classification methods (pixel and object-based ones) are reviewed in section A.1.2. For crown delineation, our state-of-the-art is organized according to the review of (Zhen et al., 2016). Focusing on species classification, the species classification assessment metrics used in this

Figure A.4: Illustration of the urban tree species mapping procedure. Left: crown delineation. The white outline polygons correspond to tree crown delineations. Right: species classification. Each color refers to a specific species. For example, *Platanus x hispanica* trees are represented in cyan whereas *Tilia tomentosa* ones are yellow.

document (state-of-the-art and contribution chapters) are first described (section A.1.2.1). Then, feature extraction, supervised classification and fusion approaches are reviewed (sections A.1.2.2, A.1.2.3 and A.1.2.4, respectively).

A.1.1 Crown delineation methods

A review on the state-of-the-art Individual Tree Crown Delineation and Detection (ITCD) methods was published in 2016 (Zhen et al., 2016). The identified approaches are summarized in Figure A.5. As no Light Detection And Ranging (LiDAR) data are available in this PhD thesis, the point cloud-based algorithms (second method group in Figure A.5) are not detailed. Also, as there is no prior information on the crown size or on the stand density in the urban environment, none of the related algorithms (third method group in Figure A.5) are considered. Moreover, the tree shape reconstruction approaches (fourth

Method Group	Example of Specific Algorithms/Methods	Advantages	Disadvantages
Raster-based method	Treetop detection: local maximum, image binarization, template matching Crown delineation: valley-following, region-growing, watershed segmentation GEOBIA-based method	• Well developed; • Easy to be used and improved; • Easy to use multiple data sources if using GEOBIA-based method; • Easy to use multiple scales if using GEOBIA-based method.	• Missing information or causing potential errors from extraction, interpolation and smoothing procedures.
Point cloud-based method	K-means clustering technique Voxel-based single tree segmentation	• Easy to use 3D information; • Easy to reflect canopy structure; • Could detect understory trees or small trees.	• Harder to implement than raster-based method; • Greatly depend on point density of LiDAR data.
Methods combining raster, point, and *a priori* information	Classic ITCD algorithms + prior information (crown size/stand density) Imagery + point cloud	• Benefit to use historical data; • Easy to integrate remotely sensed data and GIS data; • Could use both spectral and height information.	• Depend on prior information; • May have difficulties in registration of different data sources.
Tree shape reconstruction	Convex hull Alpha shape Superquadrics Hough transform	• Further delineation of crown; • Provide 3D individual tree profiles; • Provide information for estimating foliage and photosynthetic activity; • Allow further application of modeling tree growth	• Difficult to implement; • Sometimes depend on successful segmentation of single trees; • Difficult to collect field data for validation

Figure A.5: Summary of the Individual Tree Crown Delineation and Detection (ITCD) approaches from (Zhen et al., 2016)

method group in Figure A.5) are not explored because it is not in the scope of our study. This leads us to review the raster-based approaches, described in the next section.

A.1.1.1 Raster-based approaches

The raster-based approaches use rasters for crown delineation, whatever the data source. While the first group of raster-based methods consists of valley-following (Leckie et al., 2003), region-growing (Adeline, 2014) and watershed segmentation (Chen et al., 2006), Geographic Object-Based Image Analysis (GEOBIA)-based methods (Suárez et al., 2005) form the second group. All these methods can be applied on a Canopy Height Model (CHM) derived from stereoscopic acquisitions. For urban environment, the region growing approach developed in (Iovan et al., 2008) is a standard illustration of the raster-based

methods. Based on a Digital Surface Model (DSM) derived from stereoscopic acquisitions (spatial resolution of 20 cm), they delineate 78% of the trees in an area located in Marseille, France. Watershed segmentation is used in the study of (Alonzo et al., 2014) for delineating trees in Santa Barbara, California, with a CHM (spatial resolution of 25 cm). Whereas 83% of the watershed segments contain a single tree stem, this accuracy decreases to 55% when assessed on a highly complex urban forest setting, i.e. an area containing several cases of significant crown overlaps. The GEOBIA approach of (Ardila et al., 2012) can be cited with their multiple segmentation scales for delineating tree crowns in an urban area (Delft, The Netherlands). With high resolution imagery (QuickBird, spatial resolutions of 2.4 m and 0.6 m for multispectral and PAN modes, respectively), they successfully delineate 70%–80% of the trees (illustrated in Figure A.6).

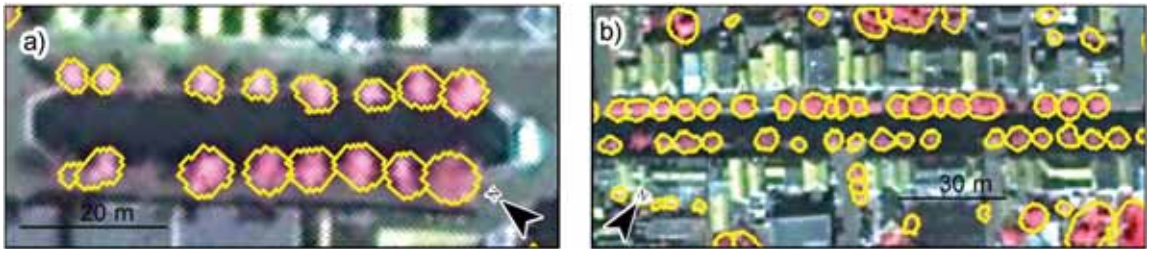

Figure A.6: Illustration of the results obtained from (Ardila et al., 2012). The yellow outline polygons correspond to the estimated tree crowns.

As for the previous work, the main errors occur for cases of significant overlaps, a well known issue in the literature (Zhen et al., 2016). However, the highlighted accuracies are quite high (around 80% whatever the spatial resolution), because there is moderate overlap in cities in comparison to natural forests. This is then reasonable to choose a standard raster-based approach for delineating the tree crowns. Without significant differences between the mentioned approaches, the one developed in (Iovan et al., 2008) is chosen for its efficiency in terms of computational burden and easiness.

A.1.2 Species classification methods

A review on the state-of-the-art tree species classification approaches is provided in (Fassnacht et al., 2016). Depending on the technology of the sensors, different information can be used to classify the species. While multispectral or hyperspectral sensors allow spectral features of the tree species to be modelled, spatial features such as structural or textural

Table A.2: Description of the confusion matrix (case of 3 species). "True" refers to the real tree species in the scene while "Predicted" refers to the predicted tree species by the method under consideration. For instance, TS_1 (i.e. "true species 1") refers to cases where the method predicts the species 1 ($\widehat{s_1}$) whereas the real species is indeed species 1 ($\widehat{s_1}$) (i.e. a true prediction). As another example, FS_3 corresponds to cases where the method wrongly predicts the species 3 (wrong prediction).

		Predicted			
		$\widehat{s_1}$	$\widehat{s_2}$	$\widehat{s_3}$	**Total**
True	s_1	TS_1	FS_2	FS_3	$Total(s_1)$
	s_2	FS_1	TS_2	FS_3	$Total(s_1)$
	s_3	FS_1	FS_2	TS_3	$Total(s_1)$
	Total	$Total(\widehat{s_1})$	$Total(\widehat{s_2})$	$Total(\widehat{s_3})$	$Total$

ones can be extracted from PAN and nDSM data, respectively, but also thanks to RAdio Detection And Ranging (RADAR) measurements. Moreover, temporal features characterizing the phenology of vegetation are accessible from time series. Contextual features can also be integrated. Our work is based on optical data, hence the methods related to RADAR data are not listed in this document. The exploitation of time series is not addressed in this work owing to the necessity of high spatial and spectral resolutions for the species cartography at tree level. Once these features are extracted, a supervised classification is carried out for identifying the species based on training samples. Based on these elements, data fusion can be employed in order to improve the classification performance of the individual methods. First, the species classification assessment metrics used in this document are detailed.

A.1.2.1 Species classification assessment metrics

In order to assess the results of the methods and compare their performances, the following metrics that can be derived from the confusion matrix (case of 3 species in Table A.2) are usually considered: Overall Accuracy (OA), κ, Producer Accuracy (PA), User Accuracy (UA) and F-score. We use them all in this work. The confusion matrix, thus the related metrics, can obviously be computed only if ground truth is available.

OA and κ are overall metrics and are derived such as (Figure A.2):

- $OA = p_0 = 100 \cdot \frac{TS_1+TS_2+TS_3}{Total}$

- Let $p_e = 100 \cdot (\frac{Total(s_1)}{Total} \cdot \frac{Total(\widehat{s_1})}{Total} + \frac{Total(s_2)}{Total} \cdot \frac{Total(\widehat{s_2})}{Total} + \frac{Total(s_3)}{Total} \cdot \frac{Total(\widehat{s_3})}{Total})$, then $\kappa = 100 \cdot \frac{p_0 - p_e}{1 - p_e}$

In the general case with N species:

- $OA = 100 \cdot \frac{\sum_{i=1}^{N} TS_i}{Total}$
- $p_e = 100 \cdot \sum_{i=1}^{N} \frac{Total(s_i)}{Total} \cdot \frac{Total(\widehat{s_i})}{Total}$

PA, UA and F-score are species-specific metrics and are derived such as, for instance for the *i*th species:

- $PA_{s_i}(\%) = 100 \cdot \frac{TS_i}{Total(s_i)}$
- $UA_{s_i}(\%) = 100 \cdot \frac{TS_i}{Total(\widehat{s_i})}$
- $F - score_{s_i}(\%) = 100 \cdot \frac{2 \cdot PA_{s_i} \cdot UA_{s_i}}{PA_{s_i} + UA_{s_i}}$

A.1.2.2 Feature extraction

The feature extraction is an essential step in order to highlight the features that can discriminate the tree species.

Spectral features Although the spectral reflectance can be directly used for the species classification, it is generally subject to specific processing in order to get more discriminative spectral features, through feature reduction techniques or transformation of the spectral reflectance.

Feature reduction can be applied to spectral reflectance in order to reduce the dimensionality of the data and to reduce potential noise: feature extraction (Principal Component Analysis (PCA) (Abbasi et al., 2015), Minimum Noise Fraction (MNF) (Ghosh and Joshi, 2014), etc.) and / or feature selection (Genetic Algorithm (GA) (Fassnacht et al., 2014), Support Vector Machine (SVM) wrapper (Fassnacht et al., 2014), etc.). The study of (Fassnacht et al., 2014) is intended to compare feature reduction algorithms for classifying tree species with hyperspectral data (HyMap) on three central European test sites (Demmin, Karlsruhe and Merzalben, Germany). Spatial resolutions of 3 m, 4 m and 5 m are used. Among the feature reduction techniques, the use of the MNF components systematically leads to the best accuracies with 91%, 96% and 59.6% for the three test

		PLS feature select.		SVM wrapper		Genetic Algorithm		Reference all bands		Reference MNF	
	Classifier	SVM	RF	SVM	RF	SVM	RF	SVM	RF	SVM	RF
Demmin	Input dataset	REFL	REFL	CONTREM	CONTREM	CONTREM	CONTREM	-	-	MNF	MNF
	Nr. of input bands	20	25	15	20	20	25	125	125	20	20
	Min	0.650	0.629	0.750	0.742	0.723	0.695	0.729	0.643	0.848	0.820
	Median	0.729*	0.720	**0.816***	0.803	0.802*	0.785	0.798*	0.742	**0.910***	0.882
	Max	0.807	0.830	0.891	0.875	0.875	0.873	0.862	0.821	0.923	0.952
	P-value	0.0384		0.0236		< 0.0001		< 0.0001		< 0.0001	
Karlsruhe	Input dataset	SAGO	SAGO	CONTREM	CONTREM	SAGO	SAGO	-	-	MNF	MNF
	Nr. of input bands	20	30	20	20	20	30	125	125	10	15
	Min	0.710	0.720	0.725	0.665	0.726	0.720	0.680	0.553	0.924	0.919
	Median	0.806	0.808	0.817*	0.777	**0.826***	0.808	0.783*	0.660	**0.960***	0.953
	Max	0.889	0.880	0.886	0.856	0.916	0.879	0.875	0.775	1.0	0.989
	P-value	0.9008		< 0.0001		< 0.0001		< 0.0001		0.0019	
Merzalben	Input dataset	SAGO	SAGO	SAGO	CONTREM	SAGO	CONTREM	-	-	MNF	MNF
	Nr. of input bands	25	20	15	10	15	25	125	125	25	10
	Min	0.392	0.375	0.441	0.386	0.396	0.427	0.375	0.356	0.506	0.507
	Median	0.509*	0.492	**0.559***	0.529	0.528	0.526	0.510*	0.466	0.596	**0.629***
	Max	0.640	0.599	0.675	0.619	0.678	0.641	0.623	0.583	0.687	0.724
	P-value	0.0003		< 0.0001		0.9680		< 0.0001		< 0.0001	

Figure A.7: Illustration of the results obtained in the feature reduction algorithms comparison (Fassnacht et al., 2014). REFL corresponds to the reflectance. SAGO refers to the smoothed first-order derivative based on the Savitzky–Golay filter. CONTREM stands for the continuum removal transformation.

sites respectively, in comparison to 79.8%, 78.3% and 51% with all the spectral bands (Figure A.7). MNF transformation has also led to the best performance for classifying tree species and vegetation classes in the works of (Ghosh et al., 2014b; Zhang and Xie, 2012).

Transformation of the spectral reflectance enhances the pigment absorption features and reduces the effects of the soil background (derivative (Ghiyamat et al., 2013), Continuum Removal (CM) (Fassnacht et al., 2014), vegetation indices (Clark and Roberts, 2012) etc.). Whereas some studies find an interest in using derivative spectra (Datt, 2000) (at the leaf level), other ones demonstrate that the direct use of the reflectance spectra leads to better results (Ghiyamat et al., 2013) (at the airborne canopy level). In particular, the latter study aims at discriminating 6 tree species including different ages of Corsican and Scots Pines located in the Thetford Forest, Britain. They use airborne hyperspectral data with a spatial resolution of 5 m (HyMap). While an accuracy of 66.9% is obtained with the reflectance, the consideration of the first and second derivative spectra declines the performance to 64.8% and 53% respectively, when using the Jeffreys–Matusita distance in a single end-member approach. They conclude that the reflectance spectra is more stable compared to

derivative spectra, which is consistent as airborne data can be particularly noisy in comparison to field data (e.g. in (Datt, 2000) where better accuracies are obtained with derivative spectra). Regarding the application of vegetation indices, none of the reviewed studies in (Fassnacht et al., 2016) demonstrated a clear advantage of their use. Nevertheless, other works show a benefit in using such indicators (Erudel et al., 2017).

At this stage, the MNF transformation is chosen for extracting the spectral features because it is a well-known approach for both improving the performance in comparison to the direct use of the spectral reflectance, and reducing the data dimension.

Structural features Many structural features (statistics (Dalponte et al., 2012), profiles (Zhang and Hu, 2012)) can be considered for modeling the 3D tree species structure. In general, this information is used in conjunction with spectral information. (Alonzo et al., 2014) classify 29 species in Santa Barbara, California, USA, with hyperspectral AVIRIS (spatial resolution of 3.7 m) and LiDAR (22 pulse/m^2) data by defining structural features from LiDAR point cloud, followed by feature level fusion with the hyperspectral data. In particular, features such as the median height of returns in crown, the crown width at median height of returns in crown, or the ratio of crown height to width are considered. Their fused results compared to those obtained with hyperspectral data alone lead to an Overall Accuracy (OA) improvement of 4.2 percentage point (pp). Focusing on the framework proposed by (Dalponte et al., 2012), they aim at identifying 7 species in a mountain area situated in the Southern Alps with AISA Eagle hyperspectral data (spatial resolution of 1 m) and LiDAR data (8.6 pts/m^2). Similar features to those computed by (Alonzo et al., 2014) are chosen. An OA improvement of 8.9pp is obtained with the inclusion of the structural features. Other studies of the literature point out the interest of extracting structural features such as those mentioned above (Holmgren et al., 2008; Ørka et al., 2009). In general, the species discrimination power of these features is low in comparison to the spectral information but improves the results.

In conclusion, the use of structural features in order to model accurately the 3D structure of the tree species is rather recent and encouraging. Therefore, height ratios similar to those developed in (Alonzo et al., 2014; Dalponte et al., 2012) and adapted to nDSM data are considered for extracting the structural features.

Textural features Different methods exist to define textural features like Grey Level Cooccurrence Matrix (GLCM) (Coburn and Roberts, 2004) or Wavelet Transform (WT) (Rajpoot and Rajpoot, 2004). Unlike the structural features for which there is no standard method dedicated for their extraction, probably because of their recent use, the texture is

exclusively modelled with the Haralick features derived from the GLCM in the tree species classification context. As for the structural features, the textural ones are in general used in conjunction with spectral information. An exception is however the work of (Iovan et al., 2008) which classifies 37 trees of 2 species in Marseille, France, from airborne multispectral data with a spatial resolution of 20 cm by computing Haralick's features from the GLCM. They get an overall accuracy of 100% in this simple study case. The study of (Johansen and Phinn, 2006) is another example of use of the Haralick parameters. In particular, the contrast, dissimilarity, entropy, homogeneity and variance are extracted from IKONOS imagery for mapping 10 tree species in Australian tropical savannah, but the benefit of using texture in addition to spectral information is not assessed. Several works use the GLCM-based parameters (Franklin et al., 2000; Ghosh et al., 2014b). In a more original way, the study of (Zhang and Hu, 2012) aims at classifying 6 species in Toronto, Canada, from multispectral data with a spatial resolution of 6 cm. They consider a radiometric profile along a path in the solar plane at the object scale in addition to multispectral reflectances, vegetation indices and texture information. In fact, this feature could be viewed as a structural feature. When adding this information, an overall accuracy improvement of approximately 10 pp is obtained. Similarly to the structural features, the discriminative power of these features is low in comparison to that of the spectral information but improves the results.

In summary, the GLCM-based parameters are commonly used in the literature to model the texture of the tree species and have demonstrated encouraging results. They are thus selected for textural feature extraction.

Contextual features The use of contextual features is not new in remote sensing methods, with the example of the GEOgraphic Object-Based Image Analysis (GEOBIA) community (Blaschke, 2010) which has already demonstrated the interest of integrating prior information (Platt and Rapoza, 2008). There are many uses of prior knowledge in the remote sensing community (Forestier et al., 2012; Heinzel et al., 2011; Zhu et al., 2001). For example, the study of (Heinzel et al., 2011) uses information about the crown size in order to improve delineation results. Prior knowledge about the objects diameter, area, perimeter, etc. are considered for improving object-based classification in urban areas by (Forestier et al., 2012). However, none of the reviewed studies takes advantage of such knowledge for urban tree species identification, whereas the urban environment has distinctive features (great species diversity, different tree structures, different tree developments for a same species, etc.). In particular, contextual information about the different urban tree structures (i.e. street trees, park trees, etc.), can be used as prior knowledge in order to improve the urban tree species classification performance. For instance, the detec-

tion of street trees has several advantages. In addition to being useful to urban managers for the specific monitoring of these trees, such information could be considered to improve classification. Knowing that the urban alignments are often monospecific or bispecific, for example in 57.7% and 24.2% of the Paris roads respectively (Rol-Tanguy et al., 2010), it is reasonable to consider a species regularization step based on the assumption that the alignments are often of the same species in order to improve the classification results. On the other hand, the different tree developments between the street trees and the other ones suggests differentiating these urban tree structures for improving the classification procedure. There is therefore an interest in identifying the trees that belong to urban alignments.

There is no such method in the literature. However, this objective can be divided in two sub-objectives: the individual tree crown delineation and the determination of membership in an alignment. From this point of view, there are many works related to the first task, as reviewed in section A.1.1. These works are useful to define the appropriate framework for individual tree crown detection in urban alignment. At this stage, we have to focus on the Individual Tree Crown Delineation and Detection (ITCD) methods that allow a prior information to be modelled (third method group in figure A.5). Indeed, trees in urban alignment are connected to each other, thus their discrimination requires the inclusion of contextual information (modelled via the so-called prior component of the algorithms). In particular, there are two categories of frameworks: classic ITCD method + prior information (figure A.5), and unified ones. Concerning the determination of membership in an alignment, the work of (Wen et al., 2017) is the only one, detailed in section A.1.2.2. This study is an example of definition of the discriminative features of street trees.

ITCD methods using a prior information Regarding classic ITCD method + prior information, prior information is included in classic ITCD methods in order to improve their performance. In particular, the most useful information that can be incorporated are the expected crown size and stand density (Zhen et al., 2015; Koch et al., 2014). About the crown size, the example of (Heinzel et al., 2011) can be cited. They apply their approach on two study areas (Poland), in a forest context. A nDSM (spatial resolution of 0.5 m) derived from LiDAR data (7 pts/m^2) is considered. A preliminary classification of the crown size is carried out before this information is used as prior knowledge for crown delineation based on watershed. They get an average improvement of 30pp, with reached accuracies ranging between 64% and 88%. Other works show improvements using the crown size (Chen et al., 2006; Zhen et al., 2014). Concerning the stand size, the work of (Hauglin et al., 2014) can be taken as an example (Norway). From data similar to those used in the previous example, a crown delineation based also on watershed is first applied, before the stand density (sample plots are used for estimating the stem number)

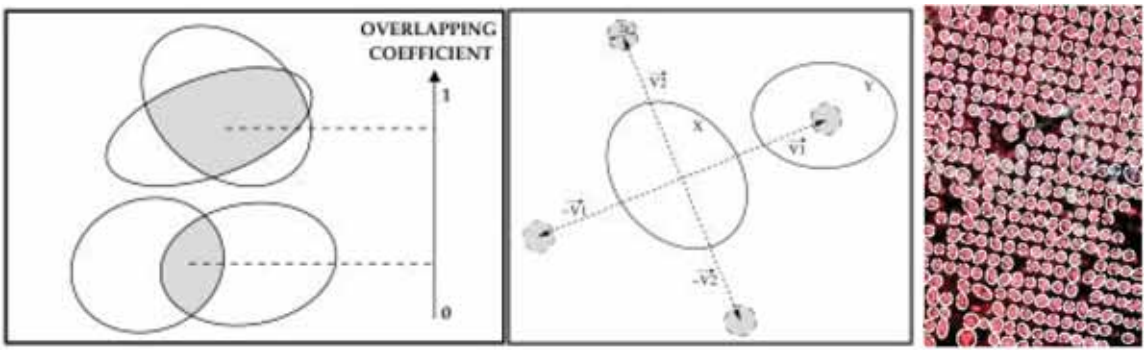

Figure A.8: Illustration of the prior energy of a MPP model (left) and results (right) of Perrin (Perrin et al., 2006). The prior energy can model alignments as shown on the right part of the prior energy illustration.

helps guiding the delineation. Even if the contribution of this auxiliary data is not assessed in this study (not the main objective), it seems to be helpful regarding the authors analysis. Other studies such that of (Ene et al., 2012) use the stand density.

In conclusion, these frameworks allow a prior information to be modelled in a simple way, via the inclusion of an additional processing before or after the classical ITCD algorithm, in other words consider several *ad hoc* steps rather than an unified model. This can cause errors since the errors associated to each step are spread through the framework in a non-reversible way (Horvath et al., 2009).

On the other hand, unified frameworks aim at combining the individual tree crown delineation and the prior information in an unified model. In particular, the Marked Point Process (MPP) framework is dedicated to find objects within an image by minimizing an energy including a data term and a prior term. The data term models how the objects fit the image while the prior term (or interaction term depending on the case) models the link between the objects (overlapping, proximity, similarity, etc., illustrated in figure A.8, left) (Van Lieshout, 2000). This framework has been used for several applications, including the crown delineation. Focusing on this application, the contributions of Perrin can be highlighted through several papers (Perrin et al., 2006, 2005, 2004) (illustrated in figure A.8). In the MPP, the shape of the delineation has to be assumed preliminarily, this is the so-called mark (a circle or an ellipse in general). The work of (Larsen et al., 2011) compares the performance of six tree crown delineation algorithms (including a MPP approach) when applied to six study sites. Colour Infrared Film (CIR) aerial photos are used

(spatial resolution varying from 10 cm to 50 cm). Different forest conditions are considered. Whereas the MPP approach is the best for the plantation scene (similar to the urban tree alignments) with matching score values up to 99%, it can decrease to 35% in case of overlapping trees and oblique acquisition configuration.

While the main advantage of the MPP is the possible inclusion of a prior information in an unified model (via the prior term), its main disadvantage is the shape of the crown that has to be assumed in advance, not allowing to delineate accurately the crowns. Even if the mentioned studies indicate errors in case of significant overlap, a more complex data term definition could help to deal with theses cases.

The active contour model is another framework, aimed at delineating objects outlines within an image by minimizing an energy including an internal term, an image term and a constraint term. The internal term can be viewed as a prior on the shape of the contour (continuity, smoothness, etc.), similar to the prior term of the MPP (at the contour level instead of the inter-objects level). The image term models how the contour fits the image, closer in signification to the data term of the MPP. Finally, the constraint term allows a user intervention to be considered for guiding the contours (via a user interface for example) (Kass et al., 1988). This framework has been applied for several applications such as the crown delineation. The study of (Lin et al., 2011) aims at delineating trees from three sample plots in Alishan National Scenic Area, Taiwan, based on a CHM derived from LiDAR data (5 pts/m^2). As an initialization step, they compute bottom up erosion (identifies stand candidates) followed by top down dilation (estimates the crown periphery). An active contour algorithm is then applied giving an average detection accuracy of 76%. Other works use active contour for this purpose (Ke et al., 2010). The Higher-order active contours (HOACs), improvement of the initial active contour framework in order to include non-trivial prior knowledge about region shape without constraining topology, has been developed by (Rochery et al., 2006) and used for crown delineation. Indeed, the paper of (Horvath et al., 2009) proposes a method for delineating the crowns of poplar stands, in France, from colour infrared aerial image (spatial resolution of 50 cm). They found that the HOAC framework is better than the classical active contour with improvements of correct detections varying from 1pp to 12pp according to the complexity of the case.

While the main advantage of the active contour is the shape of the crown that can be accurately delineated, its main disadvantage is the prior information that only concerns the contour shape, not the interaction between objects.

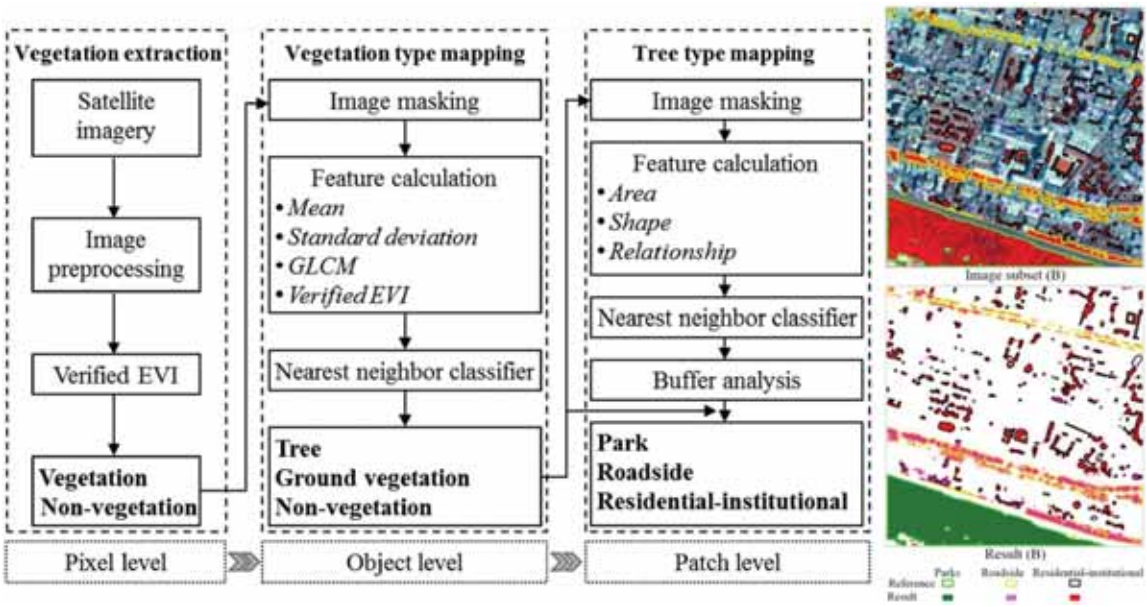

Figure A.9: Illustration of the workflow (left) and results (right) developed by Wen (Wen et al., 2017).

Patch-level approach of Wen The work of Wen (Wen et al., 2017) is particularly relevant because it aims at classifying the urban canopies (patch-level classification) in three classes (park, roadside and residential-institutional canopies). The roadside class corresponds to our urban alignments. In particular, their method is applied on two areas: Shenzhen and Wuhan (China). WorldView-2 satellite imagery (spatial resolution of 2 m for the multispectral mode) as well as the road network (from OSM) are used. From a methodological point of view, their framework consists of three main steps (presented in figure A.9, left): vegetation extraction (vegetation / non-vegetation at the pixel level), vegetation type mapping (tree / ground vegetation / non-vegetation at the object level different to tree level: multiple crowns) and tree type mapping (park / roadside / residential-institutional at the patch level: alignment scale). For the second and last steps, segmentation (Baatz and Schape, 2000) and supervised classification (k-NN) are carried out in order to identify the classes under consideration. Focusing on the last step which is of interest for this chapter, specific spectral, textural, shape and contextual features are considered for characterizing these classes. Among these numerous features, the perimeter-area ratio, the related circumscribing circle and the distance to road of the patches are computed. Regarding the results, F-score values of 76%, 89% and 87% are obtained for park, roadside

and residential canopies respectively (illustrated in figure A.9, right).

In such a patch-level framework, there are confusions between the street trees and the other populations of trees because of the spatial connections between the canopies (i.e. proximity between roadside and park ones for example). Moreover, the assessment metrics used in this study are computed in terms of patch whose reference data are not clearly described through the paper, which does not allow the latter confusions to be correctly quantified.

A.1.2.3 Supervised classification

Once these features (spectral, textural, structural, contextual, etc.) are extracted, they are subject to classification from training samples, through the use of a supervised classification algorithm.

Training samples The training samples can be extracted directly from the images, or from an external database based on field measurements for examples (Fassnacht et al., 2016). It is traditionally done directly from the image level, either from the image under consideration itself, or from other images of the same data type, to get rid of the change of scale and because field measurements are not always available. In particular, many studies consider manually delineated trees for extracting the training samples (Alonzo et al., 2014). While it is useful to consider this assumption in order to compare different classification strategies, it is not realistic for an operational point of view. Instead, learning examples extracted from field measurements are relevant candidates. There is no work dedicated to the classification of tree species based on leaf or canopy training samples using optical imagery (Fassnacht et al., 2014). However, when focusing on the thermal infrared domain (8.0 - 13.5 μm), the study of (da Luz and Crowley, 2010) uses the SEBASS airborne sensor (spatial resolution of 1 m) in order to analyse and map canopy spectral features in the State Arboretum of Virginia, near Boyce, Virginia. They success in classifying up to 20 species based on laboratory-measured leaf spectra. To do so, they use the original spectral reflectance. Focusing on the optical domain, previous studies show that field and airborne spectral reflectances are often incomparable, especially in the case of leaf level measurements because of the variability of the canopy structure (Roberts et al., 2004). Faced with this problem, spectral features such as vegetation indices can be both discriminative and invariant to the change of scale (Cho et al., 2008) (Figure A.10). Focusing on Figure A.10, while the majority of the indices changes significantly from leaf to canopy scale, indices such as Curvature Index (CI), Gitelson and Merzlyak Index (GMI) or Photochemical Reflectance Index (PRI) are invariant for certain species. As an exam-

Vegetation index	Formula	Biophysical significance	Reference
Normalised difference vegetation index (NDVI)	$(R_{830}-R_{670})/(R_{830}+R_{670})$	Canopy greenness, LAI, fraction of photosynthetically active radiation	Rouse et al., 1974; Tucker, 1979
Carter index (CI)	R_{760}/R_{695}	Chlorophyll content	Carter, 1994
Gitelson and Merzylak index (GMI)	R_{750}/R_{700}	Chlorophyll content	Gitelson and Merzlyak, 1997
Vogelman index (VOG)	R_{740}/R_{720}	Chlorophyll content	Vogelmann et al., 1993
Photochemical reflectance index (PRI)	$(R_{531}-R_{570})/(R_{531}+R_{570})$	Conversion of xanthophylls-cycle pigments, photosynthetic light-use efficiency, LAI	Gamon et al., 1992; Peñuelas et al., 1995
Carotenoid reflectance index (CRI)	$R_{800}(1/R_{520}-1/R_{550})$	Carotenoids (alpha- and beta-xanthophylls), indicator of plant stress	Getilson et al., 2002

Species	NDVI	CI	GMI	VOG	PRI	CRI
Hedera	-2.22^{*}	-0.90^{ns}	-1.24^{ns}	-3.38^{**}	-4.95^{**}	-0.44^{ns}
Rhodo-dendron	-7.43^{**}	-8.73^{**}	-7.40^{**}	-5.66^{**}	7.44^{**}	-10.98^{**}
Prunus	-4.11^{**}	-4.00^{**}	-4.45^{**}	-7.94^{**}	-5.26^{**}	-3.39^{**}
Corylus	-8.03^{**}	-8.45^{**}	-4.86^{**}	-3.41^{**}	2.22^{*}	-9.11^{**}
Malus	-2.02^{*}	-0.5^{ns}	0.88^{ns}	2.33^{*}	6.47^{**}	-3.94^{**}
Aesculus	-4.78^{**}	-4.2^{**}	-3.49^{**}	-5.08^{**}	1.34^{ns}	-5.23^{**}

Figure A.10: List of indices (top) and two-sample t-test for differences between leaf and canopy vegetation indices (bottom) based on the study of (Cho et al., 2008). $* = p < 0.05, ** = p < 0.01, ns = not significant (p > 0.05)$ from (Cho et al., 2008).

ple, PRI index is invariant for *Aesculus* and found to have high potential to discriminate the tree species in the work of (Cho et al., 2008). In another context than the tree species classification (crop mapping), the use of ground-based spectral references has already proven its potential (Nidamanuri and Zbell, 2011).

To conclude, training samples both directly extracted from the images and based on field spectral measurements are considered. In the case of field measurements, spectral indices

are preferred in order to deal with the change of scale.

Supervised classification algorithms An overview over advantages and disadvantages of most commonly applied classification algorithms is provided in (Fassnacht et al., 2016), with a more detailed review provided by (Lu and Weng, 2007). According to (Fassnacht et al., 2016), the choice of the algorithm is of low importance when the requirements of the classifier in terms of data preprocessing are reached. In particular, using the non-parametric Support Vector Machine (SVM) and Random Forest (RF) does not require any distributional assumption, and these algorithms have already demonstrated good performance in the literature (Féret and Asner, 2013; Sheeren et al., 2016).

Therefore, SVM and RF algorithms are chosen.

A.1.2.4 Fusion

Each set of features (spectral, structural, textural, contextual, etc.) contributes in the species identification, there is thus an interest in combining them in order to improve tree species maps. A usual way is then to combine multiple data sources / features / classification algorithms, i.e. to consider a fusion framework. However, the term fusion is ambiguous as it can be achieved at different levels. To avoid any ambiguity, we have summarized the possible fusion strategies depending on the information to fuse in Figure A.11. Two possible fusion frameworks can be used: feature level fusion (Alonzo et al., 2014; Dalponte et al., 2012) and decision level fusion (Stavrakoudis et al., 2014; Engler et al., 2013). In our context, the decision level fusion and the ensemble classifier refer to the same framework. On the one hand, the feature level fusion stacks many characteristics of interest and the resulting feature vector is classified. On the other hand, the decision level fusion considers several classifiers (feature extraction and supervised classification as shown in Figure A.11), according to a criterion, and the predictions of these classifiers are combined through a decision rule. Focusing on Figure A.11, two categories of decision level fusions are distinguished: data sources-based and classifier-based approaches. For the first one, a classifier is defined for each data source while the classifiers are chosen according to another criterion for the second one (leading for instance to the same set of features but different supervised classification algorithms, the same algorithm but different sets of features, etc.). The latter approach is therefore very flexible. Regarding the tree species classification context, the feature level fusion is always preferred for its simplicity to implement except in the studies of (Stavrakoudis et al., 2014) and (Engler et al., 2013) where decision level fusions are considered, data sources-based and classifier-based approaches respectively, both detailed below. In general, spectral information is used in

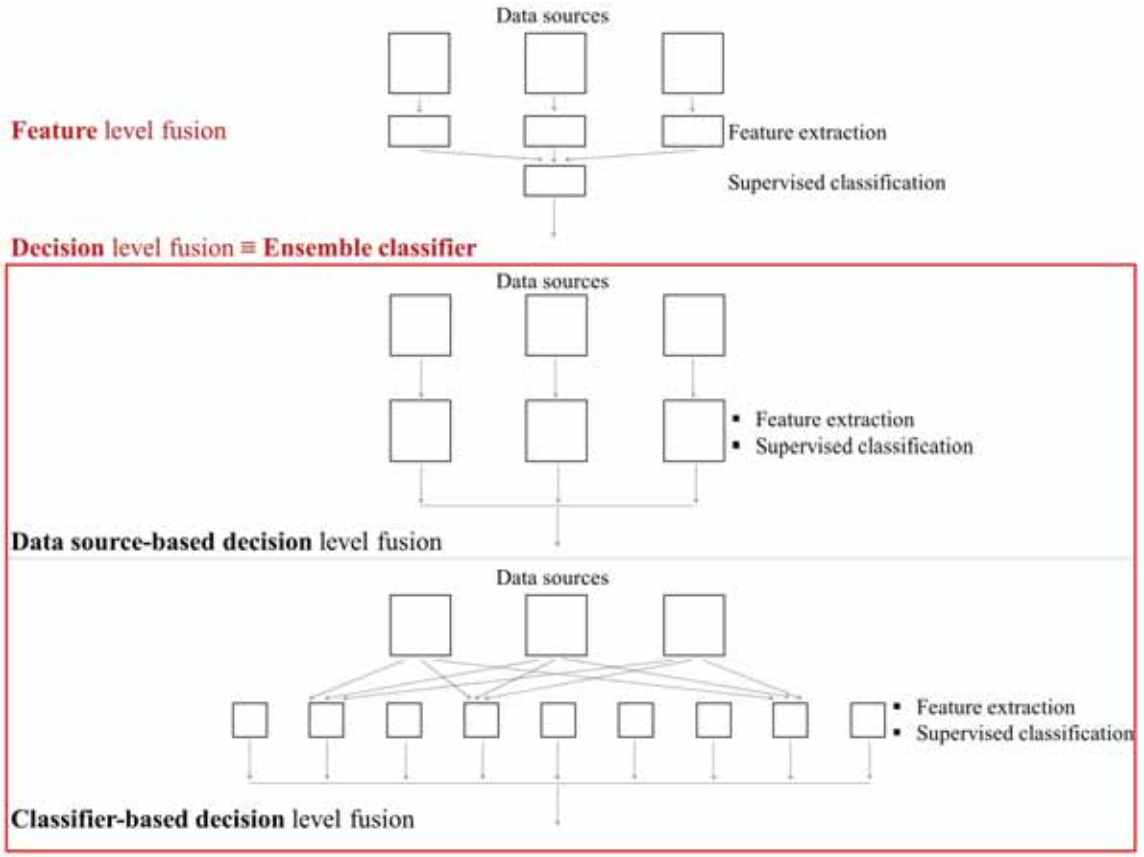

Figure A.11: Fusion strategies definition. Focusing on the classifier-based decision fusion, all the arrows are not drawn in order not to overload the diagram.

conjunction with others (textual and structural information). In particular, there are many works which fuse hyperspectral and LiDAR data (17 out of the 29 relevant cases listed in the recent review of (Fassnacht et al., 2016)).

Feature level fusion Focusing on the fusion of spectral and structural features, the studies of (Alonzo et al., 2014) and (Dalponte et al., 2012) reflect the state-of-the-art spectral / structural information combinations. In an urban context, the work of (Alonzo et al., 2014) aims at classifying 29 tree species in Santa Barbara, California. For that task, they use hyperspectral AVIRIS data (spatial resolution of 3.7 m) and airborne LiDAR data (22 pulse/m^2). They stack the spectral and structural features at the hyperspectral pixel

level, knowing that the structural characteristics are computed at the object level (i.e. subsampling of the structural ones). In an object-based approach, the resulting feature vectors are classified within the objects thanks to a Canonical Discriminant Analysis (CDA) supervised classification algorithm, followed by a majority vote (feature level fusion, Figure A.12, and results illustrated in Figure A.13). Whereas they get an Overall Accuracy (OA) value of 83.4% with the hyperspectral data alone, an increase of 4.2pp is obtained with the inclusion of LiDAR-based metrics (illustrated in Figure A.12). This improvement occurs for species crowns either small, i.e. having few pixels (6 species on the basis of 8 for which improvements higher than 10pp have been observed), or morphologically unique, e.g. particularly tall species such as *Washingtonia robusta* (2 species concerned). In the work of (Dalponte et al., 2012), they aim at identifying 7 species in a mountain area situated in the Southern Alps. In particular, they use AISA Eagle hyperspectral data (spatial resolution of 1 m) and LiDAR data (8.6 pts/m^2). They stack the spectral and structural features at the hyperspectral pixel level, knowing that this time the structural characteristics are computed at the pixel level. A SVM supervised classification is then applied in a pixel-based approach (feature level fusion). An OA value of 83% is obtained with the fusion while the use of the hyperspectral data alone leads to an OA score of 74.1% (-8.9pp). Other studies demonstrate the interest of using the spectral and structural information together (Dalponte et al., 2008; Jones et al., 2010) (+2pp in terms of OA).

Regarding the fusion of spectral and textural features, compared to the spectral / structural information combinations which include hyperspectral data for modelling the spectral features, the spectral features used in the studies which fuse spectral / textural information are based on multispectral data. It is expected that the contribution of the textural information would be more significant when adding to multispectral data in comparison to hyperspectral data, as hyperspectral data are more powerful for tree species classification when its spatial resolution is similar. Thus, the following accuracies have to be carefully considered. The works of (Zhang and Hu, 2012) and (Franklin et al., 2000) illustrate the spectral / textural combinations. The first one aims at classifying 6 species in Toronto, Canada, from multispectral data with a spatial resolution of 6 cm. A knowledge-based decision tree including spectral and textural features is used as classification algorithm in an object-based manner, which can be viewed as a feature level fusion. When adding textural features, an OA improvement of approximately 10pp is obtained. Otherwise, the study of (Franklin et al., 2000) aims at identifying pure and mixed wood forest stands species in Alberta and New Brunswick, Canada. Multispectral data with a spatial resolution of 1 m are used. Maximum Likelihood (ML) supervised classification with a pixel-based approach is applied (feature level fusion). While the addition of the textural features lead to an OA improvement of 5pp for Alberta, an increase of 12pp is obtained for New Brunswick. Other

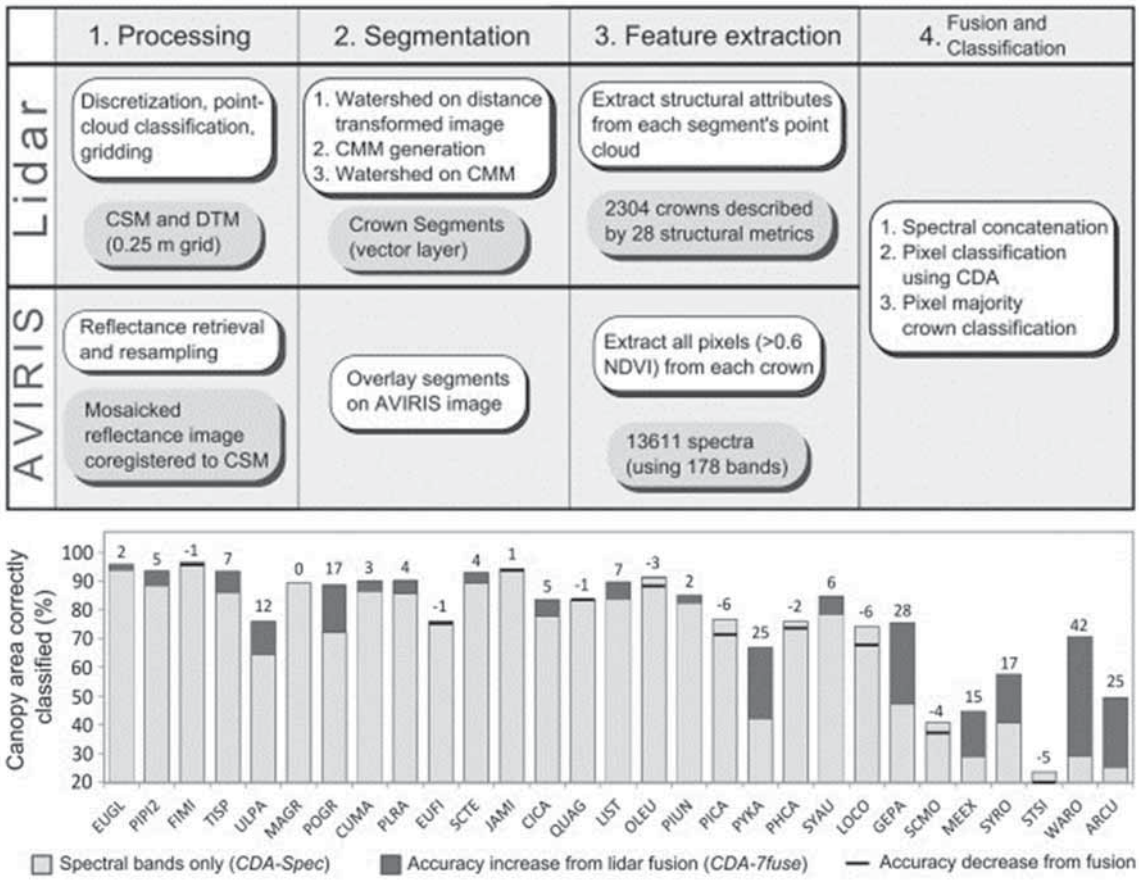

Figure A.12: Illustration of the hyperspectral / LiDAR fusion proposed by (Alonzo et al., 2014) for the classification of 29 tree species in an urban environment. Top: classification framework. Bottom: results.

studies (Johansen and Phinn, 2006; Mallinis et al., 2008) lead to similar conclusions.

In conclusion for feature level fusion, even if there is a benefit in fusing spectral / structural or spectral / textural information for tree species classification, the mentioned approaches fail to substantially improve the performance. This can be caused by several reasons. First of all, the structural or textural features are not enough complementary to the spectral ones. Secondly, the fusion strategy does not allow to take advantage of this complementarity. No work analyses the complementarity of the sources subject to fusion, and feature level fusion is generally preferred to decision level fusion because of its simplicity to implement

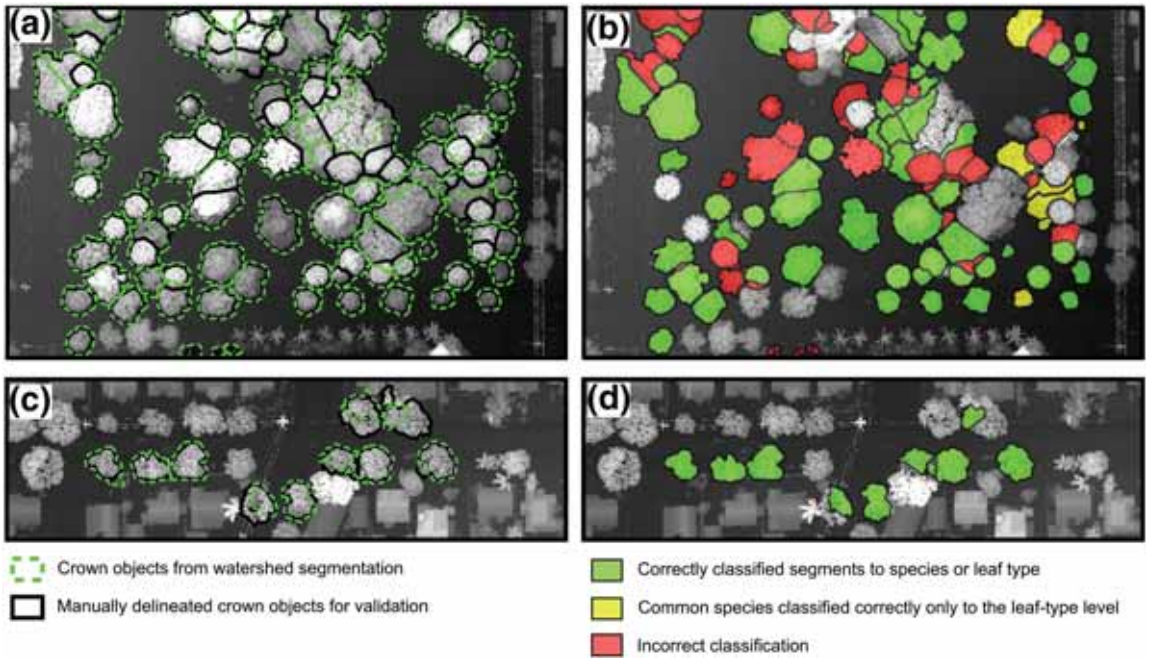

Figure A.13: Urban tree map obtained in (Alonzo et al., 2014).

whereas no comparison of the strategies has been carried out.

Decision level fusion Even if the feature level fusion framework is preferred for tree species classification, the method developed in (Stavrakoudis et al., 2014) constitutes an example of data sources-based decision level fusion, in a pixel-based framework. In particular, they aim at classifying 6 species in the southern and southwestern slopes of Mount Cholomontas in Macedonia. Multispectral QuickBird data (spatial resolution of 2.4 m) and hyperspectral EO-1 Hyperion data (spatial resolution of 30 m) are used, which differs from the previous state-of-the-art examples where spectral information were used in conjunction with others (structural or textural). After classifying the pixels of the two images with a SVM supervised classification, the resulting membership probabilities corresponding to Hyperion are resampled to the QuickBird spatial resolution, and a decision rule is applied for predicting the species (Figure A.14). While the OA was of 66.5% and 65.7% for the hyperspectral and the multispectral data, it reaches 78.9% with fusion strategy (illustrated in Figure A.14). In another context than tree species classification (land cover mapping), other studies have demonstrated the efficiency of data sources-based decision

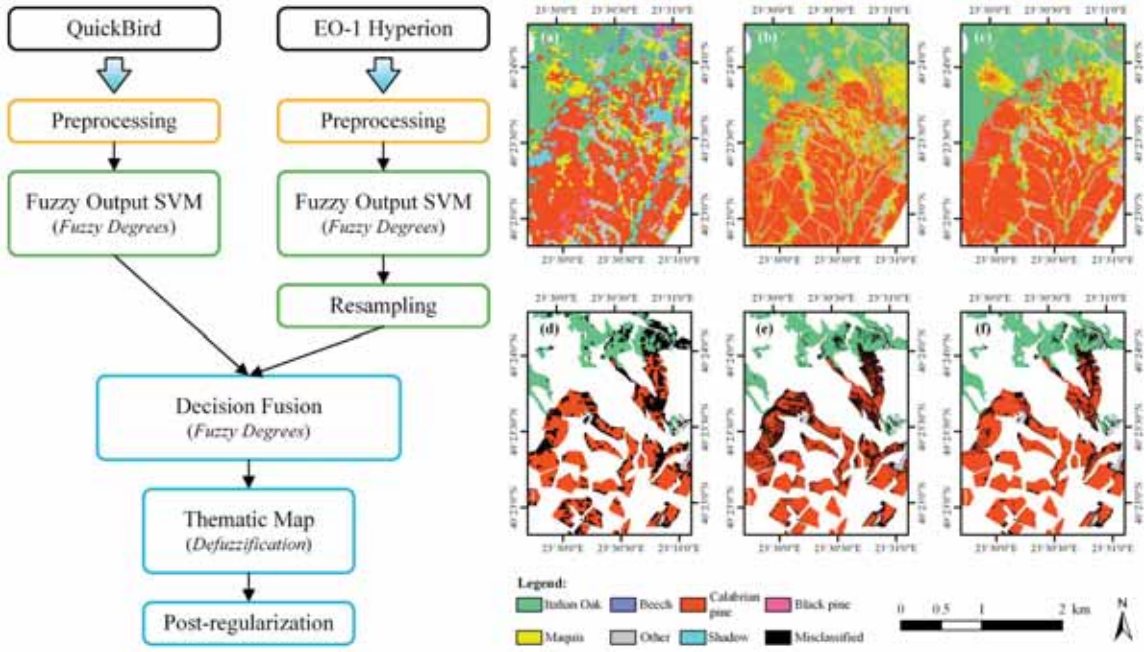

Figure A.14: Decision level fusion developed in (Stavrakoudis et al., 2014) for the classification of 6 tree species in a mountain area. Left: classification framework. Right: results. Focusing on the results, (a), (b) and (c) correspond to the Hyperion, QuickBird and fusion performance maps. (d), (e) and (f) highlight the misclassifications in each case, by comparing the predicted species with the ground truth available in this study.

level fusion, such as (Abbasi et al., 2015).

To conclude, no analysis of the complementarity of the sources involved in the fusion has been carried out. The fusion frameworks are not compared to the standard feature level one. Although having one classifier associated to a particular data source allows the contribution of each one to be assessed, this remains an arbitrary choice.

The study of (Engler et al., 2013) is an example of classifier-based decision level fusion. They aim at classifying 6 tree species in North-Eastern Switzerland, based on high-resolution aerial imagery (50 cm spatial resolution) and topo-climatic variables (5 m spatial resolution). In particular, they use different classification algorithms while keeping the same set of features, and find this ensemble classifier better than individual approaches.

In another context than tree species classification (land cover mapping), the work of (Ceamanos et al., 2010) demonstrates that it is preferable to use an ensemble classifier approach instead of a standard one when dealing with hyperspectral data. The reason given is related to the Hughes effect (curse of dimensionality) (Hughes, 1968) that is particularly significant when the number of training samples is much lower than the number of spectral bands when focusing on hyperspectral signals. This effect is naturally less important in the case of an ensemble classifier approach. With AVIRIS hyperspectral data (spatial resolution of 3.7 m), the ensemble methods developed in the study of (Ceamanos et al., 2010) can lead to OA improvements up to 5pp in comparison to the commonly used approach (Figure A.15, right). Especially, the hyperspectral data are decomposed into few data sources

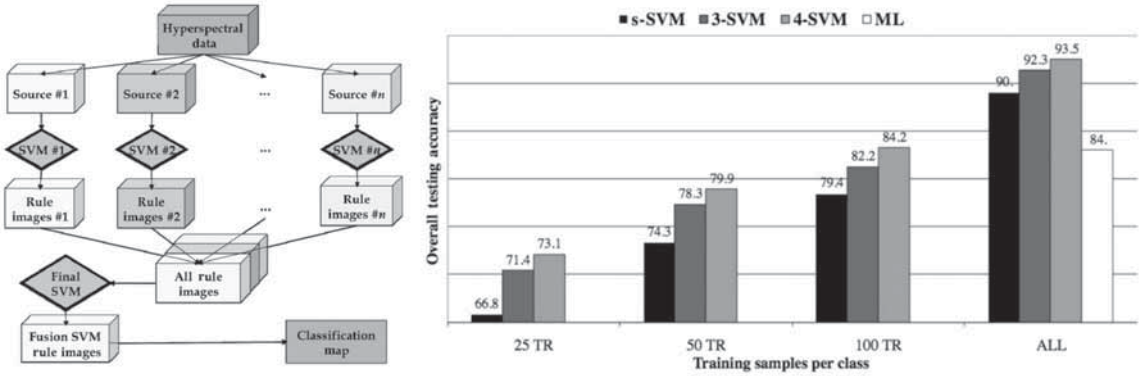

Figure A.15: Ensemble method and results from the work of (Ceamanos et al., 2010). s-SVM stands for the standard approach (single classifier) while 3-SVM and 4-SVM refer to the cases where 3 and 4 SVMs are considered. ML stands for Maximum Likelihood.

according to the similarity of the spectral bands. Then, each source is classified by a SVM before a decision rule is applied (illustration in Figure A.15, left). Other studies lead to similar conclusions (Wang et al., 2009; Xia et al., 2015, 2017) (urban land cover applications).

From a global point of view, a review of the ensemble classifier theory is provided in (Kuncheva, 2004). In particular, the decision rule used for combining the predictions of each classifier is the critical point of these approaches. Several strategies can be used: majority vote, weighted majority vote, naive Bayes combination, multinomial methods, decision templates, Dempster-Shafer combination, etc. According to (Kuncheva, 2004), the simple weighted average class of rules have been most widely used due to their simplicity and consistently good performance. The decision level fusion framework has been

used many times in another context than tree species classification (Zhang, 2010). For instance, the work of (Tupin et al., 1999) is intended to identify classes in Synthetic Aperture Radar (SAR) images, based on Dempster-Shafer theory (Shafer, 1976) applied to several structure detectors. The method proposed in (Chanussot et al., 1999) is dedicated to detect linear feature in SAR satellite data applied to road network extraction based on fuzzy fusion techniques. Using IKONOS imagery, the study of (Fauvel et al., 2006) proposes an ensemble classifier approach for the classification of urban remote sensing images based on a fuzzy decision rule.

To conclude regarding the mentioned classifier-based decision level fusion studies, the feature extraction step is not carried out in such a way that it directly optimizes the classification accuracy. Another drawback is that the complementarity of the classifiers is not optimized. Finally, a simple weighted average decision rule is chosen.

Summary and selected approach

Selected methods

Focusing on the tree crown delineation, this is reasonable to choose a standard raster-based approach for delineating the tree crowns. Without significant differences between the mentioned approaches, the one developed in (Iovan et al., 2008) is selected for its efficiency in terms of computational burden and easiness.

Regarding the species classification, in particular the feature extraction approaches, the MNF transformation and spectral indices are chosen for extracting the spectral features, while Haralick parameters will be used to model the tree species texture. About the 3D structure, height ratios adapted to nDSM data will be computed based on the previous studies. Regarding the supervised classification task, training samples both directly extracted from the images and from field measurements are considered. SVM and RF algorithms are chosen. A simple weighted average decision rule is considered for the fusion.

Finally the Overall Accuracy (OA), κ, Producer Accuracy (PA), User Accuracy (UA) and F-score metrics, are all considered in this work.

Developed methods

Concerning the fusion, the combination of spectral and structural information have demonstrated a potential for tree species classification, but without substantial improvement. Tex-

tural information contribution follows the same behaviour, except that there are far fewer studies for this type of fusion. There is no work that fuses spectral, structural and textural information, whereas it would lead to better results, as improvements have been stated for the spectral / structural and spectral / textural information combinations, although slight. Moreover, the complementarity of these features has not been assessed, whereas non complementary sources cannot logically improve the performance compared to the best of them. Whereas the feature level fusion is widely used in the literature, the decision level fusion has already demonstrated its potential, and no comparison has been carried out in order to select the best approach for a given tree species classification case. This leads us to identify the best object-based fusion strategy (feature or decision level) taking advantage of the complementarity of several heterogeneous airborne data sources for improving the urban tree species mapping. To begin, a data sources-based decision level fusion seems to be a candidate of interest in order to assess the contribution of each data source. The training samples are directly extracted from the images based on manually delineated crowns (chapter I).

In order to improve these multi-source classification methods, the second part of this PhD thesis will explore the potential of a classifier-based decision level fusion. In particular, there is an interest in extracting the features in such way that they optimize the classification accuracy, for example per species if each classifier is dedicated to the prediction of a particular species. Classically, the training step of the classification models is based on target image data (Fassnacht et al., 2016; Alonzo et al., 2014). We will analyse the quality of the maps generated when the training is carried out from field measurements (canopy and leaf levels). The interest is the use the resulting classification approach in an operational context where target image training samples are not available (Nidamanuri and Zbell, 2011). In order to overcome the problem of change of scale, introduced when the training is based on field measurements and not images, the maps obtained by using vegetation indices are assessed (Cho et al., 2008). Therefore, there is an interest in developing an ensemble of species-specific classifiers based on field spectral measurements using vegetation indices (chapter II).

Further improvements of the proposed methodologies could be obtained through the consideration of contextual information. The detection of the street trees would allow the tree species predictions to be regularized within urban alignments. Two frameworks are possible for this detection: classic ITCD method + prior information, and unified ones. In particular, the unified frameworks do not have the disadvantages of the classic + prior information ones in terms of spread of errors. Secondly, the unified ones leads to similar performance. Active contour method is better for an accurate delineation of crowns but is

not appropriate for modelling the interaction between the trees (spatial organization and common features among the trees), an essential feature for detecting trees in urban alignment. Thus the MPP are selected. Regarding the work of Wen, its patch level approach is not compatible with the MPP but some of the features used to discriminate the street trees could be considered, especially the distance between the tree and the road. This leads us to develop a MPP method for detecting the street trees based on airborne data and contextual information (chapter III).

Chapter I

Urban tree species classification from multiple airborne data sources

Synthèse de l'article en français

L'objectif de ce papier est d'identifier la meilleure stratégie de fusion orientée objet qui exploite la complémentarité de plusieurs sources de données aéroportées hétérogènes pour améliorer la classification de 15 espèces d'arbres en milieu urbain (Toulouse, France).

Jeu de données

Afin de sélectionner la meilleure approche de fusion, des classifications mono et multi-source sont d'abord testées sur un site de référence où 15 espèces d'arbres sont préalablement identifiées et les couronnes d'arbres délimitées manuellement (un total de 194 arbres). Ensuite, cette approche est introduite dans un processus automatique (délimitation de la couronne et classification des espèces) pour classer les espèces d'un site test, indépendant du site de référence utilisé pour l'apprentissage. Dans ce cas, la méthode n'est évaluée que pour des arbres d'alignement d'espèces majoritaires : *Tilia tomentosa* et *Platanus x hispanica*.

Parce que les données aéroportées hyperspecrales, PAN et nDSM sont *a priori* complémentaires pour la classification des espèces, ces données sont considérées à travers les approches de fusion testées. Alors que les données hyperspectrales sont destinées à mettre en évidence les caractéristiques spectrales des espèces d'arbres, les données PAN et nDSM sont utilisées pour modéliser leurs propriétés texturales et structurales, respectivement. Les principales caractéristiques du jeu de données sont décrites dans le tableau

TABLE I.1 : Principales caractéristiques du jeu de données. "N" fait référence au nombre de bandes spectrales. GSD signifie Ground Sampling Distance.

	VNIR	**SWIR**	**PAN**	**nDSM**
Source	HySpex VNIR-1600	HySpex SWIR-320m-e	CAMV2	CAMV2
Quantité	réflectance spectrale	réflectance spectrale	compte numérique	hauteur
GSD	0.4 m	1.6 m	0.14 m	0.125 m
Intervalle	0.4 - 1 μm	1 - 2.5 μm		
N	160	256		

III.2, après que les prétraitements géométrique et radiométrique aient été effectués. Le nDSM est obtenu à partir d'acquisitions stéréoscopiques du système CAMV2.

Méthode proposée

Concernant la classification mono-source, les composantes MNF sont calculées pour chaque pixel afin de constituer les vecteurs de caractéristiques spectrales Visible Near Infrared (VNIR) et Short Wavelength Infrared (SWIR). Alors que les vecteurs de caractéristiques texturales sont constitués de paramètres de Haralick dérivés de la matrice de cooccurrences de niveaux de gris (GLCM), plusieurs rapports de hauteur composent les vecteurs de caractéristiques structurales. Ces caractéristiques texturales et structurales sont calculées à l'échelle de la couronne. Une fois ces caractéristiques construites, elles sont respectivement classées dans les couronnes d'arbres grâce à un algorithme de classification supervisée (SVM et RF testés). Cela permet de calculer un profil de décision fondé sur les probabilités d'appartenance et soumis à une règle de vote majoritaire, permettant de prédire une espèce pour chaque arbre et chaque type de données.

Une approche de fusion au niveau de la décision est ensuite proposée sur la base d'un profil de décision constitué des profils de décision combinés de chaque type de données. Trois règles de décision sont prises en compte pour prédire l'espèce. La source VNIR est considérée comme une référence puisqu'elle donne la meilleure classification mono-source. Une méthode de fusion standard au niveau des caractéristiques est considérée comme une autre référence. Au lieu de combiner les résultats de classification comme pour la fusion au niveau de la décision, toutes les caractéristiques (spectrales, texturales et structurales) sont concaténées dans un vecteur de caractéristiques pour chaque pixel

TABLE I.2 : OA (%) et κ (%) par source. Les scores engras associés d'une étoile (*) font référence au meilleur score parmi les sources.

	VNIR	**SWIR**	**PAN**	**nDSM**	**Baseline**	**Fusion**
OA (%)	75	69	43	35	73	**77***
κ (%)	72	65	35	27	69	**74***

VNIR, ainsi soumis au même traitement que dans la classification mono-source.

Résultats

Les principaux résultats des classifications mono-source et multi-source sur le site de référence sont présentés dans Table I.2. Pour la classification mono-source, le VNIR est le meilleur avec une valeur d'OA de 75%, suivi par le SWIR (69%). Le PAN et le nDSM conduisent à des valeurs d'OA de 43% et 35%. Les composantes MNF montrent de meilleures précisions que l'ensemble des bandes spectrales (+9pp pour le SWIR avec SVM)). En conclusion, les données hyperspectrales sont le principal moteur de la précision tandis que les données PAN et nDSM contribuent marginalement. En outre, il est avantageux d'utiliser des composantes MNF pour le traitement de données hyperspectrales.

En ce qui concerne la classification multi-source, l'approche de fusion au niveau de la décision proposée améliore légèrement la performance (+2pp) tandis que la fusion standard diminue l'OA du VNIR (-2pp). En particulier, l'analyse de complémentarité des accords de prédiction démontre que la fusion proposée tire profit des cas de complémentarité, par exemple lorsque le VNIR se trompe mais qu'au moins une autre source a raison, mais les cas où le VNIR est correct alors que le SWIR est faux ne sont pas bien gérés par la stratégie proposée. Concernant l'analyse de la complémentarité par espèce, la fusion au niveau de la décision améliore la performance pour 8 espèces sur 15, incluant les deux espèces majoritaires (*Tilia tomentosa* et *Platanus x hispanica*). Pour le site test, la fusion au niveau de la décision conduit à une valeur d'OA de 63% contre 55% pour le VNIR.

Discussions

Ces résultats démontrent que les caractéristiques spectrales sont le principal moteur de la précision de classification, ce qui est cohérent avec la littérature (Fassnacht et al., 2016). La réduction MNF améliore les performances, comme mentionné par (Fassnacht et al.,

2014). Évaluée à partir des communalités MNF des bandes spectrales (Bailey et al., 2002), la contribution du red edge semble être particulièrement utile dans notre contexte (Dalponte et al., 2009). D'autre part, l'analyse spatiale de la performance démontre que les arbres d'alignement monospécifiques sont bien identifiés alors que les principales erreurs se produisent pour les parcs. Ainsi, la discrimination parc / arbres d'alignement pourrait être intéressante pour appliquer un traitement spécifique aux arbres de parc.

En ce qui concerne les caractéristiques texturales et structurales, elles contribuent marginalement, ce qui est cohérent avec les travaux antérieurs (Franklin et al., 2000; Alonzo et al., 2014). Globalement, les caractéristiques texturales et structurales permettent de bien classer les arbres d'alignement. Comme ces caractéristiques ne permettent pas de classer l'ensemble des espèces (15 ici) avec une grande précision, il est probablement intéressant de les utiliser d'une autre manière, par exemple en discriminant des groupes d'espèces ayant des textures similaires, de manière hiérarchique.

L'analyse de complémentarité réalisée dans cette étude démontre que les sources sont complémentaires mais que cette complémentarité est faible. Dans ce contexte, une fusion standard au niveau de la caractéristique (Alonzo et al., 2014) ne permet pas d'améliorer les performances du VNIR alors que l'approche proposée y succède, mais avec de légères améliorations (Stavrakoudis et al., 2014). Ce n'est *a priori* pas la conséquence d'une règle de décision non optimale, mais plutôt les sources individuelles qui ne sont pas assez complémentaires. En particulier, la méthode au niveau des caractéristiques souffre de l'effet de Hughes (fléau de la dimension) (Hughes, 1968). La complémentarité des sources doit être optimisée, par exemple en définissant une source spécifique par espèce.

Conclusions

L'objectif est d'identifier la meilleure stratégie de fusion orientée objet qui tire profit de la complémentarité de plusieurs sources de données hétérogènes aéroportées pour améliorer la classification de 15 espèces d'arbres dans une zone urbaine (Toulouse, France). Les données hyperspectrales aéroportées VNIR et SWIR, PAN et nDSM sont prises en compte. Les stratégies de fusion au niveau de la décision et au niveau des caractéristiques sont comparées lorsqu'elles sont appliquées à un site de référence, et la meilleure est introduite dans un processus automatique de prédiction des espèces d'arbres dans un site test. L'approche VNIR et l'approche de fusion au niveau des caractéristiques sont choisies comme méthodes de référence.

En ce qui concerne les résultats, les sources VNIR et SWIR sont les meilleures avec des

valeurs d'OA de 75% et 69%, respectivement. En particulier, il y a un intérêt à classer les espèces avec les composantes MNF. Le PAN et le nDSM conduisent à des valeurs d'OA de 43% et 35%. Pour la classification multi-source, la fusion au niveau de la décision proposée améliore légèrement les performances du VNIR (77% au lieu de 75%). Cette légère amélioration est due à la faible complémentarité des sources, plutôt qu'à la règle de décision.

Une amélioration des méthodes est nécessaire. Premièrement, il est nécessaire d'optimiser la complémentarité des sources, par exemple en définissant une source spécifique par espèce. Cela pourrait se faire en utilisant des indices spectraux, en plus des bandes spectrales et des composantes MNF. D'un autre côté, il semble prometteur de distinguer les arbres de parc des arbres d'alignement, comme le démontre l'analyse spatiale. Enfin, la définition d'échantillons à partir de mesures sur le terrain pourrait permettre de rendre la méthode plus opérationnelle.

English part: First article

The first paper is included in the next section[1].

I.1 Object-based fusion for urban tree species classification from heterogeneous data: hyperspectral, panchromatic and nDSM

Abstract: This study aims at identifying the best object-based fusion strategy that takes advantage of the complementarity of several heterogeneous airborne data sources for improving the classification of 15 tree species in an urban area (Toulouse, France). The airborne data sources are: hyperspectral Visible Near-Infrared (160 spectral bands, spatial resolution of 0.4 m) and Short-Wavelength Infrared (256 spectral bands, 1.6 m), panchromatic (14 cm) and normalized Digital Surface Model (12.5 cm). Object-based feature and decision level fusion strategies are proposed and compared when applied to a reference site where the species are previously identified during ground truth collection. This allows the best fusion strategy to be selected with a view to introducing the method in an automatic process (tree crown delineation and species classification) on a test site, independent of the reference site used for learning. In particular, a decision level fusion is selected: Visible Near-Infrared and Short-Wavelength Infrared classifications use Minimum Noise Fraction components at the original spatial resolution and Support Vector Machine, whereas panchromatic and normalized Digital Surface Model classifications use respectively Haralick's and structural features computed at the object scale, and Random Forest. After the computation of a decision profile for each source at the object level based on the classification algorithms membership probabilities, these decision profiles are combined and a decision rule is applied to predict the species. Focusing on the reference site, the Visible Near-Infrared exhibits the best performances with F-score values higher than 60% for 13 species out of 15. The Short-Wavelength Infrared is the most powerful for 3 species with F-score greater than 60% for 7 common species with the Visible Near-Infrared. The panchromatic and normalized Digital Surface Model contribute marginally. The best fusion strategy (decision fusion) does not improve significantly the overall accuracy with 77% (kappa = 74%) against 75% (kappa = 72%) for the Visible Near-Infrared but in general, it improves the results for cases where complementarities have been observed. When

[1] **J. Aval**, S. Fabre, E. Zenou, D. Sheeren, M. Fauvel and X. Briottet. Object-based fusion for urban tree species classification from heterogeneous data: hyperspectral, panchromatic and nDSM. International Journal of Remote Sensing, 2018.

applied to the test site and assessed for the two majority species (*Tilia tomentosa* and *Platanus x hispanica*), the selected approach gives consistent results with an overall accuracy of 63% against 55% for the Visible Near-Infrared.

Keywords Tree species classification; urban remote sensing; hyperspectral; panchromatic; nDSM; spatial information; object-based; complementarity; decision fusion.

I.1.1 Introduction

In urban areas, trees can impact the microclimate, promote biodiversity, have a relaxing psychic action and contribute to aesthetics (Jones, 2014). During heatwaves in dense and polluted cities, tree infrastructures can locally decrease the temperature (freshness islands) (ADEUS, 2014) and improve the air quality (Yin et al., 2011). These properties depend on the capacity of evapotranspiration and on the crown volume of the trees which are related to the tree species. Urban trees are also subject to special conditions and can be particularly affected by some diseases. For example, *Ceratocystis platani* is a fungus responsible for the Canker stain of *Platanus* trees. In Europe, it was allegedly introduced to Marseille, France in 1945 from infested wooden crates of US troops containing military equipment (Vigouroux, 2014). At Forte dei Marmi in Italy, 90% of the Platanus trees died from 1972-1991. In cities, trees are often planted as street trees and belong to the same tree species, thus the transmission of this type of disease is easier. The current struggle against these species specific diseases is based on several key elements, including the monitoring. Beside these statements, tree species information is essential for the management of urban trees. Nowadays, the operational procedure for tree species classification is based on field campaign which does not allow large areas to be covered on a regular basis. For example, the city of Toulouse, France, would have approximately 140,000 trees spread over more than 100 km^2 according to estimates of vegetation managers, such a procedure is not appropriate.

Remote sensing opens the way to produce maps of tree species automatically (Fassnacht et al., 2016; Shojanoori and Shafri, 2016; Sheeren et al., 2016; Li et al., 2015). Thanks to airborne hyperspectral sensors it is possible to acquire the spectral reflectance of vegetation volumes with a spatial resolution of an order of magnitude of 1 m (Dalponte et al., 2009) which is related to the foliar components (Jacquemoud and Baret, 1990) and the foliage structure (Verhoef, 1984), and therefore to the species. Airborne panchromatic (PAN) sensors allow a very high spatial resolution measurement of the reflected radiation integrated over the visible spectral range (order of magnitude of 10 cm) (Iovan et al., 2008) and gives information about the spatial arrangement of the foliage within the tree

crowns, which is also related to the species (Zhang and Hu, 2012). The airborne normalized Digital Surface Models (nDSM) allow a very high spatial resolution measurement of the height (order of magnitude of 10 cm) (Dalponte et al., 2015) and gives information about the 3D structure of trees, which may differ between species (Van Leeuwen et al., 2010). Remote sensing gives encouraging results in tree species classification (Shojanoori et al., 2018), but in urban environment it remains a challenging task because of the large tree diversity (species, age, life conditions, pruning, etc.) (Welch, 1982; Alonzo et al., 2013) with potentially a small number of individuals per species.

Depending on the technology of the sensors, different parameters can be used to classify the species. While Very High Resolution (VHR) imagery is required for studying the urban environment, the high spectral resolution available from hyperspectral data is more suitable than the one of multispectral data when dealing with several tree species (Dalponte et al., 2012). The consideration of airborne hyperspectral sensors, combining both high spatial and spectral resolution, is therefore preferable for urban tree species identification although VHR satellite data would be cheaper and would allow the monitoring to be carried out more frequently. Focusing on these hyperspectral data, feature reduction can be applied to spectral reflectance in order to reduce the dimensionality of the data and to reduce potential noise: feature extraction (Principal Component Analysis (PCA) (Abbasi et al., 2015), Minimum Noise Fraction (MNF) (Ghosh and Joshi, 2014), etc.) and / or feature selection (Genetic Algorithm (GA) (Fassnacht et al., 2014), Support Vector Machine (SVM) wrapper (Fassnacht et al., 2014), etc.). There are also transformations of the spectral reflectance to enhance the pigment absorption features and to reduce the effects of the soil background (derivative (Ghiyamat et al., 2013), Continuum Removal (CM) (Fassnacht et al., 2014), vegetation indices (Clark and Roberts, 2012) etc.). In particular, the MNF performs well in many studies (Fassnacht et al., 2014).

Regarding PAN and nDSM data, different methods exist to define textural features (Grey Level Cooccurrence Matrix (GLCM) (Coburn and Roberts, 2004), Wavelet Transform (WT) (Rajpoot and Rajpoot, 2004)) and structural features (statistics (Dalponte et al., 2012), profiles (Zhang and Hu, 2012)). In general, these data are used in conjunction with spectral information. Focusing on object-based classification, (Iovan et al., 2008) classify 37 trees of 2 species in Marseille, France, from airborne multispectral data with a spatial resolution of 20 cm by computing GLCM and Haralick's features at the object scale. They have an overall accuracy of 100% which can be expected in such a simple case. (Zhang and Hu, 2012) classify 6 species in Toronto, Canada, from multispectral data with a spatial resolution of 6 cm by considering a radiometric profile along a path in the solar plane at the object scale, in addition to multispectral reflectances, vegetation indices

and texture information. When adding this information, an overall accuracy improvement of approximately 10 pp is obtained. (Alonzo et al., 2014) classify 29 species in Santa Barbara, California, USA, with hyperspectral AVIRIS (Airborne Visible Infrared Imaging Spectrometer, spatial resolution of 3.7 m) and LiDAR (Light Detection And Ranging, 22 pulse/m^2) data by defining structural features from LiDAR point cloud at the object scale, followed by feature level fusion. Their fused results compared to those obtained with hyperspectral data alone lead to an overall accuracy improvement of 4.2 pp.

Because hyperspectral, PAN and nDSM data contribute differently in species classification, combining them can lead to better performances (Alonzo et al., 2014). In a classification process, data fusion can mainly be achieved at two levels: feature (fusion of feature vectors) (Alonzo et al., 2014) or decision (fusion of classification results) (Stavrakoudis et al., 2014). As said previously, (Alonzo et al., 2014) classify 29 species in Santa Barbara by using a feature level fusion where hyperspectral data at the pixel scale and structural features at the object scale are stacked in a feature vector at the pixel scale (spatial subsampling of structural features). (Stavrakoudis et al., 2014) classify 4 species in Macedonian forests with hyperspectral (Hyperion, spatial resolution of 30 m) and multispectral (Quickbird, spatial resolution of 2.4 m) data by using a pixel-based decision-level fusion based on SVM membership probabilities. While the overall accuracy was of 66.5% and 65.7% for the hyperspectral and the multispectral data, it reaches 78.9% with fusion strategy. In all the studies, there are mainly combinations of two sources while using several sources should improve the performance more. In addition, there is often partial complementarity analysis of the data sources, whereas non complementary sources cannot logically improve the performance compared to the best of them. Concerning the fusion strategy, feature level fusion is often used for object-based classification whereas all the features are not necessarily computed at the same spatial scale, precisely when working with heterogeneous data as highlighted in Figure I.1. Particularly, these features have to be resampled at the same scale to be stacked in the same feature vector, requiring quality registration. This implicitly assigns weights to each feature which have no basis in a perspective of species classification. And more features can lead to a decrease in accuracy because of the Hughes effect (Hughes, 1968). Moreover, the classification algorithm is the same for all the sources, whereas one might be more appropriate for certain sources. On the other hand, a decision level fusion does not have these drawbacks but requires a decision rule in order to weight the different sources.

As a conclusion, only few works focused on the case of the object-based fusion from several heterogeneous airborne data sources for the classification of tree species, by considering different fusion strategies and by assessing the complementarity of the sources.

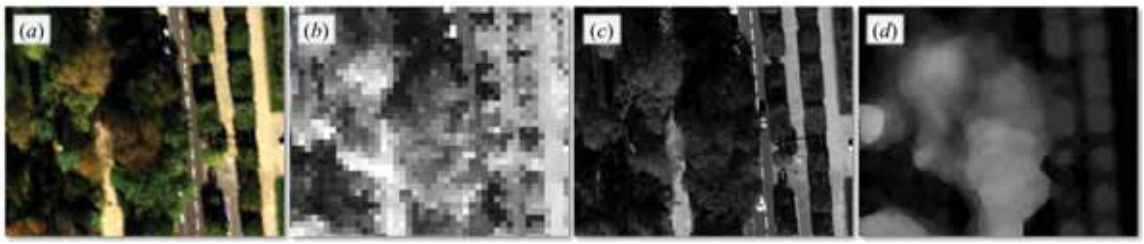

Figure I.1: Example of heterogeneous airborne data sources for the same observed area. From left to right, hyperspectral Visible Near-Infrared (VNIR) image (a), hyperspectral Short-Wavelength Infrared (SWIR) image (b), panchromatic image (c) and a nDSM (d) obtained from stereoscopic acquisitions.

The present work aims at identifying the best object-based fusion strategy taking advantage of the complementarity of several heterogeneous airborne data sources for improving the classification of 15 tree species in an urban area (Toulouse, France). This includes a first stage where feature and decision level fusion strategies are compared when applied to a reference site, and a second stage where the best fusion strategy is introduced in an automatic process in order to classify the tree species of a test site, independent of the reference site used for learning. More precisely, the following questions are addressed:

1. What unique information content is available in each data source?
2. What is the best fusion strategy that takes advantage of the sources complementarities?

I.1.2 Materials

I.1.2.1 Study area

The study area is located in Toulouse, France (43.6 °N, 1.44 °E). Toulouse is the fourth city in France with about 500,000 inhabitants. It has a temperate climate with oceanic, Mediterranean and continental characteristics. In general, it has mild winters, wet springs with thunderstorms, dry and warm summers and sunny autumns. Toulouse should have approximately 140,000 trees according to estimates of vegetation managers, with at least 50 species. These trees are distributed along streets, in parks and in private properties. Our experimental analysis of the images is carried out in two study sites (Figure I.2, (a)): a reference site with a tree reference map (detailed in Section II.1.2.4) for applying several

mono-source and multi-source classification approaches and selecting the best ones, and a test site, independent from the reference site used for learning, for assessing the potential of the selected approaches through a completely automatic process (Section II.1.2.5).

Figure I.2: (a) Overall view of the downtown part of Toulouse from Google Earth. The yellow rectangles indicate the two study areas: reference and test sites. (b) Reference site with park and street trees represented on the French Aerospace Lab airborne VNIR data. The colored polygons indicate the delineations and the species of the inventoried trees in the reference site. (c) Test site with park and street trees, especially composed of the two majority species street trees (*Tilia tomentosa*, *Platanus x hispanica*) represented on the VNIR. The white outline polygons correspond to the automatic tree crown delineation results.

I.1.2.2 Airborne data

The airborne data were acquired on October 24, 2012 at 11:00 UT (Universal Time) during the UMBRA campaign (Adeline et al., 2013) organized by the French Aerospace Lab (ONERA) and the French Mapping Agency (IGN). The sun zenith angle was approximately 60 °. Concerning the measurement devices, the HySpex Visible Near-Infrared (VNIR) and Short-Wavelength Infrared (SWIR) (Köhler, 2016) and CAMv2 (Souchon et al., 2010) systems were installed on board an aircraft and the flight height was approximately 2,000 m over the study area. The VNIR and SWIR systems consist of hyperspectral push broom cameras with respectively 160 spectral bands (0.4 μm - 1 μm) and 256 spectral bands (1 μm - 2.5 μm). The CAMv2 system has a PAN matricial camera. Regarding the spatial resolution, the VNIR camera has a pixel Field Of View (FOV) of 0.18 mrad and 0.36 mrad across and along track whereas the SWIR camera has a pixel FOV of 0.75 mrad. This results in spatial resolutions of 0.4 m and 0.8 m across and along track for the VNIR and 1.6 m for the SWIR. For the CAMv2, it results in a spatial resolution of 0.14 m and the images were acquired in stereoscopic configurations with an overlap of 80% to build a DSM.

I.1.2.3 Preprocessing

The French Mapping Agency (IGN) provides us with an orthorectified and georeferenced DSM with a spatial resolution of 0.125 m. Then, the hyperspectral and PAN data are registered on the DSM by defining Ground Control Points (GCP) using QGIS software and *gdalwarp* from GDAL. Nearest neighbor resampling is applied in order to preserve the original spectral data. Also, the Thin Plate Spline (TPS) transformation (Duchon, 1977) is applied for its ability to correct the deformations locally. Because the VNIR pixels have rectangular shapes with the longer side along track, a square grid with a spatial resolution of 0.4 m (minimum between the rectangle sides) is chosen to preserve the original data. For the SWIR and the PAN data, the spatial resolutions of 1.6 m and 0.14 m are kept. Visual assessment suggests that the error related to registration quality is less than a pixel for all the data set. Furthermore, the hyperspectral data are atmospherically corrected to deal with spectral reflectances with the COCHISE platform (Poutier et al., 2002) based on MODTRAN and assuming a flat scene. Spectral bands are not taken into account where the Signal-to-Noise Ratio (SNR) was low (due to atmospheric water absorption (1.339 μm - 1.423 μm, 1.79 μm - 1.952 μm) and due to low signal and spectral sensivity (2.444 μm - 2.499 μm) for examples). Finally, the DSM is normalized to produce a nDSM thanks to a Digital Terrain Model (DTM) with a spatial resolution of 25 m and knowing that the ground of the study area is flat.

I.1.2.4 Reference site and tree reference map

On the reference site, a tree reference map is built from an existing inventory delivered by Toulouse city and from a field campaign (Figure I.2, (b)). In this reference map, there are 194 trees with an unbalanced sample which is representative of the urban environment in Toulouse (Table I.3). The trees are delineated manually on the PAN data, ensuring that the selected pixels belong to trees. This avoids distorting the assessment of the classification because of delineation errors for example. The number of pixels per species and per data source is given in Table I.3. The species code (Table I.3) will be used in Sections I.1.4 and I.1.5 (results and discussions).

Table I.3: Main characteristics of the trees in the reference map.

Species scientific name (species code)	**Tree type**	**Stem count**	**Canopy area (m^2)**	**VNIR pixel count**	**SWIR pixel count**	**PAN pixel count**	**nDSM pixel count**
Acer negundo (A.n.)	Broadleaf	6	202	1077	41	9929	12325
Acer platanoides (A.p.)	Broadleaf	6	204	1084	39	10035	12458
Aesculus hippocastanum (A.h.)	Broadleaf	23	1433	7936	335	71237	88367
Cedrus atlantica (C.a.)	Coniferous	6	1036	5954	283	52119	64584
Celtis australis (C.au.)	Broadleaf	10	1032	5856	264	51721	64082
Celtis occidentalis (C.o.)	Broadleaf	9	1340	7750	382	67656	83792
Fagus sylvatica (F.s.)	Broadleaf	10	529	2897	118	26229	32559
Juglans nigra (J.n.)	Broadleaf	12	904	5021	220	45011	55867
Liquidambar styraciflua (L.s.)	Broadleaf	6	334	1836	75	16565	20562
Liriodendron tulipifera (L.t.)	Broadleaf	11	436	2510	94	22983	28534
Platanus x hispanica (P.h.)	Broadleaf	26	2082	11730	521	104134	129107
Taxus baccata (T.b.)	Coniferous	7	226	1185	39	11024	13732
Tilia platyphyllos (T.p.)	Broadleaf	14	570	3075	106	28064	34380
Tilia tomentosa (T.t.)	Broadleaf	41	1424	7650	292	70082	86980
Ulmus glabra (U.g.)	Broadleaf	7	1169	6736	323	58918	72962

I.1.2.5 Test site and tree crown delineation

This test site is independent from the reference site and far from it (Figure I.2, (a)). It is mainly composed of two majority species street trees (*Tilia tomentosa* and *Platanus x hispanica*, a species highly represented in the south of France), easily identifiable by visual

interpretation with the help of Google Street View (and checked by Toulouse city). This visual information is exploited to generate the reference classified product used to assess the performance of the proposed method (Figure I.2, (c)).

In order to automate the processing chain, the trees are now delineated automatically (Figure I.2, (c)). In particular, three masks (vegetation, shadow and height) are generated and combined (geometric intersection, i.e. logical and) to generate a high vegetation mask without shadow. For the vegetation mask, the Normalized Difference Vegetation Index (NDVI) index (Rouse Jr et al., 1974) is computed for each pixel of the VNIR image from a red and an infrared bands (643 nm and 788 nm respectively). About the shadow mask, the spectral reflectance cannot be retrieved in shadows, as the atmospheric correction method is based on a flat scene hypothesis. To avoid errors from these shadow regions, the associated pixels are masked by using the following literature index: $I = 1/6(2R + G + B + 2(\text{NIR}))$ (Nagao et al., 1979). This index is used for its efficiency and simplicity. Above thresholds determined automatically with the Otsu method (Otsu, 1975), the pixels are considered as vegetation pixel and pixel in the sun, respectively. Regarding the height mask, all the pixels with a nDSM value higher than 5 m are filtered (the minimum height value of the trees in Toulouse according to urban managers). Then, the region growing-based delineation method developed by (Adeline, 2014) and inspired by the work of (Iovan et al., 2008) is chosen because such approach is commonly used in the literature (Zhen et al., 2016). In particular, a Canopy Height Model (CHM) is derived from the high vegetation mask. The principle of the algorithm is then to choose the highest pixel of the CHM as the first pixel of the first delineated tree. Then the height is decremented and the corresponding pixel is either assigned to that first tree if it is at a distance less than 2 m here as in (Adeline, 2014), or assigned to a new tree, and so on. The produced delineation map allows localizing the trees for which species have to be defined.

I.1.3 Methods

I.1.3.1 Classification framework

The main steps of our method are detailed in Figure I.3. Different feature extraction techniques and classification algorithms are considered (Section I.1.3.2). Then, object-based feature and decision level fusion strategies are proposed and compared when applied to the reference site (Section I.1.3.3). This allows the best fusion strategy to be selected with a view to mapping the trees of the test site. The feature level fusion is based on the spatial resampling of the feature vectors on the VNIR spatial resolution. Concerning the decision level and due to the large heterogeneity of our data set, a classification is conducted for each source independently (VNIR, SWIR, PAN and nDSM sources) (mono-source

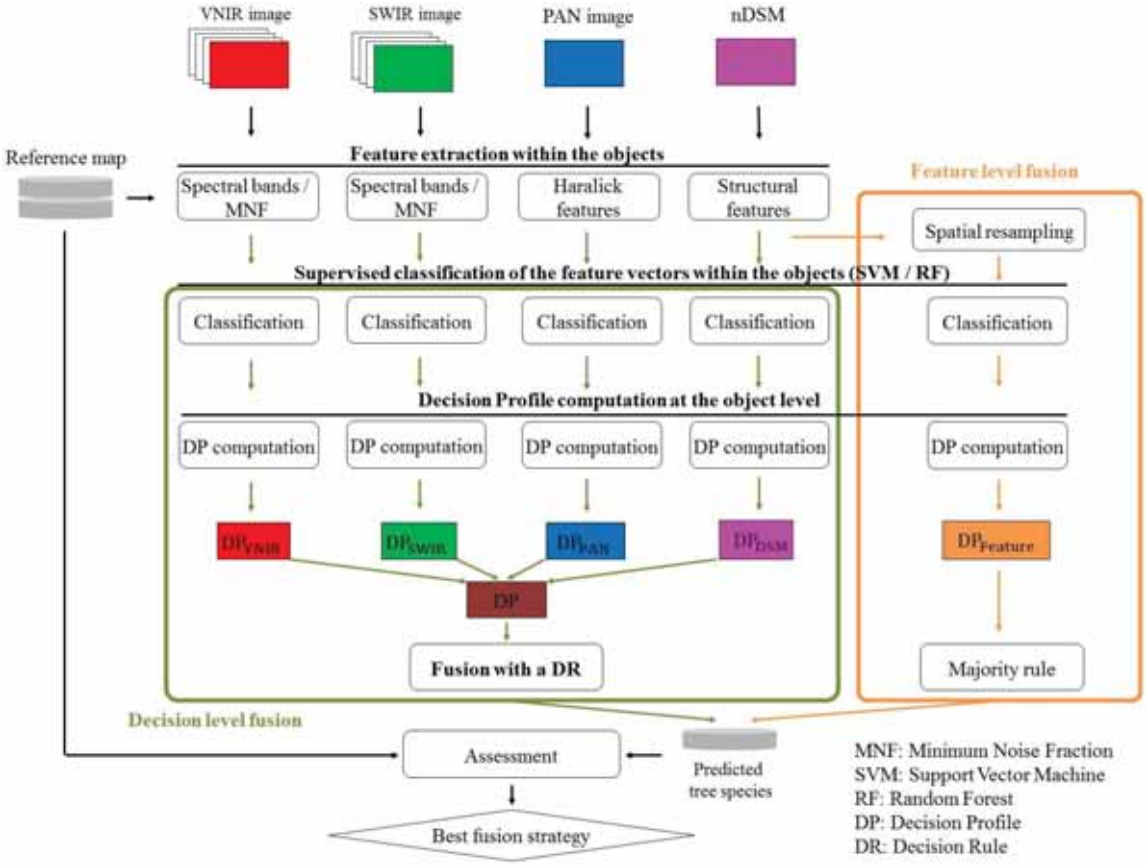

Figure I.3: Main steps of the proposed framework.

classification). The results are then combined to estimate the species (multi-source classification). For the reference site, the assessment of the classifications is carried out through comparison between the predicted species and the reference map species. The training and testing sets are independent as explained in Section I.1.3.3. Regarding the test site, the training is carried out on the whole reference site and the results are assessed only for the two majority species street trees. Below, the mono-source classification is described with the specific processing for each source followed by the multi-source classification.

I.1.3.2 Mono-source classification

Feature extraction within the objects For the VNIR and SWIR sources, the feature vector is made up of either all the spectral bands or MNF components (Figure I.3). The

MNF reduction is used for dimensionality reduction and to reduce potential noise. The MNF reduction improves performance in a number of studies according to (Fassnacht et al., 2016). Here, 30 and 15 Noise-Adjusted Principal Components (NAPC) are selected for the VNIR and SWIR sources after testing 5, 10, 15, 20, 25 and 30 as in (Fassnacht et al., 2014). The training set is used to estimate the MNF model. Concerning the spatial scale of the features, a feature vector is computed for each pixel within the crowns. Indeed, the study of (Alonzo et al., 2014) demonstrates that for manually delineated urban tree crowns, the pixel majority ("winner-take-all") approach is better than the consideration of a single crown-mean spectrum for example, especially with limited training data.

For the PAN source, the feature vector is made up of Haralick features, used to characterize the texture of the tree species, derived from the GLCM computed at the crown scale (Figure I.3). The GLCM is commonly used in the literature to model the texture (Haralick et al., 1973; Franklin et al., 2000). Eight Haralick features are derived from the GLCM (mean, standard deviation, homogeneity, dissimilarity, entropy, second angular moment, contrast and correlation). A distance and an orientation are necessary to compute the GLCM. The distance is related to the frequencies of the texture The orientation is related to the anisotropy of the object. An orientation of 90° is arbitrarily chosen because a tree is a priori isotropic in terms of texture. A distance of one pixel is chosen to keep the finer details.

For the nDSM source, the feature vector is made up of structural features computed at the crown scale (Figure I.3). These features are used to characterize the 3D structure of the tree species. Particularly, structural features similar to those developed in (Alonzo et al., 2014; Dalponte et al., 2012) and adapted to nDSM are used. If h is the height within the crown and A the crown area, these structural features are: $h_{\max}$, A, $\frac{A}{h_{\max}}$, $Ah_{\max}$, $h_{\min}$, h_{mean}, $h_{\max} - h_{\min}$, $h_{\max} - h_{\text{mean}}$, h_{std}, $\frac{h_{\max} - h_{\min}}{h_{\max}}$, $\frac{h_{\max} - h_{\text{mean}}}{h_{\max}}$, $\frac{h_{\text{std}}}{h_{\max}}$. The subscripts max, min, mean and std refer to maximum, minimum, mean and standard deviation of h.

Supervised classification of the feature vectors within the objects The resulting feature vectors within the objects are used for the classification. For this step, two supervised classification algorithms, SVM and Random Forest (RF), are considered (Figure I.3) in order to ensure the stability of the results. No other algorithm is considered because the choice of the classifier is not the main purpose of this study. Using these non-parametric algorithms, no distributional assumption is required and these algorithms have already demonstrated good performance in the literature (Féret and Asner, 2013; Sheeren et al., 2016). For the SVM, the one-vs-one multiclass strategy is used for its computation time which is better than in the case of the one-vs-rest (Fassnacht et al., 2016). Adequate val-

ues of the hyperparameters are estimated after testing ranges of values using a grid search by 2-fold cross-validation because the number of samples is limited for some classes like *Acer platanoides* or *Liquidambar styraciflua* (Table II.3).

In particular, the membership probabilities are used because they give an evaluation of the certainty of the predictions (Stavrakoudis et al., 2014; Kuncheva, 2004). The method of computing these membership probabilities depends on the classification algorithm and the selected methods are commonly used in the literature. For the SVM in the binary case, they can be calibrated using the Platt scaling (Platt et al., 1999) which consists of a logistic regression on the SVM scores fit by a cross-validation on the training data. (Wu et al., 2004) extended this to the multiclass case which is used here. For the RF, the membership probabilities are computed as the mean predicted class probabilities of the decision trees in the forest. The class probability of a single tree is the fraction of samples of the same class in a leaf (leaf: end node of the decision tree) (Li, 2013). The membership probabilities are computed for each feature vector within the objects.

Decision profile computation at the object level From the classification outputs, the obtained membership probabilities $p_{s,n}(c)$ are combined to compute a Decision Profile (DP) $\boldsymbol{d}_s$ per source s (VNIR, SWIR, PAN, nDSM and feature level approach, Figure I.3). $p_{s,n}(c)$ are computed per species c (class) for each feature vector n within the crown and for each source s. A source certainty $p_s(c)$ that the tree belongs to the species c is then computed as if each feature vector votes: $p_s(c) = \frac{1}{N_s}\sum_{n=1}^{N_s} p_{s,n}(c)$. The object is classified as the species corresponding to the maximum $p_s(c)$ by the source of interest: $c = \arg\max_c p_s(c)$ (winner-take-all). The decision profile is defined such as $\boldsymbol{d}_s = \{p_s(c), 1 < c \leqslant 15\}$ knowing that there are 15 species. In particular, there are N_s feature vectors within the crown for the source of interest: $N_{(\text{VNIR})}$ and $N_{(\text{SWIR})}$ are the number of VNIR and SWIR pixels within the crown, $N_{(\text{PAN})}$ and $N_{(\text{nDSM})}$ are equal to 1 because the textural and structural features are computed at the object scale (illustrated in Figure I.4).

I.1.3.3 Multi-source classification

Feature level fusion The feature vectors need to be spatially resampled before the supervised classification. In order to compare the feature and decision level fusion strategies with the same features, we chose to resample each feature vector on the VNIR spatial resolution (the smallest in terms of feature spatial scale) to get for each VNIR pixel, the nearest (in terms of spatial distance) SWIR, textural and structural features. Then, the membership probabilities of the feature vectors within the crowns are estimated and the species of the trees are predicted as in the mono-source classification case (Figure I.3, Sections I.1.3.2

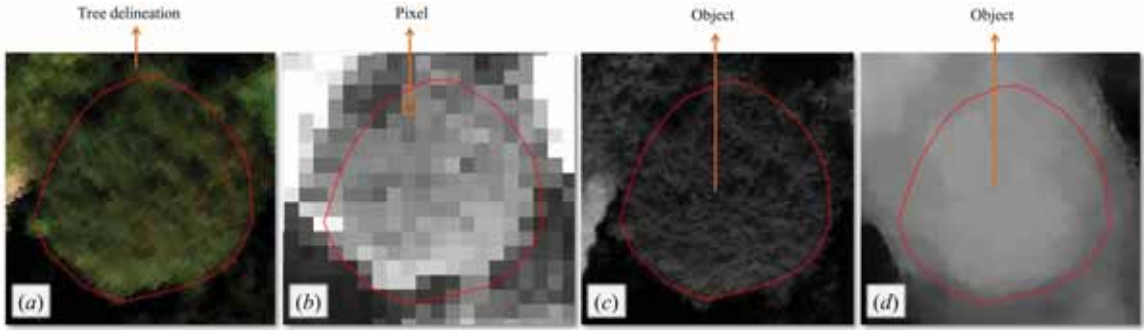

Figure I.4: Illustration for object-based classification. For the VNIR (a) and SWIR (b) sources, each pixel within the crown is classified. For the PAN (c) and nDSM (d) sources, the crown is directly classified as the textural and structural features are computed at the object scale.

and I.1.3.2), i.e. through a pixel majority (winner-take-all) approach.

Decision level fusion The decision profile is represented as a matrix of 4 rows (4 sources) and 15 columns (15 species), each element is $p_s(c)$ (Section I.1.3.2) (Kuncheva, 2004):

$$\mathbf{d} = \begin{pmatrix} \boldsymbol{d}_{(\text{VNIR})} \\ \boldsymbol{d}_{(\text{SWIR})} \\ \boldsymbol{d}_{(\text{PAN})} \\ \boldsymbol{d}_{(\text{nDSM})} \end{pmatrix} = \begin{pmatrix} p_{(\text{VNIR})}(c=1) & \dots & p_{(\text{VNIR})}(c=i) & \dots & p_{(\text{VNIR})}(c=15) \\ p_{(\text{SWIR})}(c=1) & \dots & p_{(\text{SWIR})}(c=i) & \dots & p_{(\text{SWIR})}(c=15) \\ p_{(\text{PAN})}(c=1) & \dots & p_{(\text{PAN})}(c=i) & \dots & p_{(\text{PAN})}(c=15) \\ p_{(\text{nDSM})}(c=1) & \dots & p_{(\text{nDSM})}(c=i) & \dots & p_{(\text{nDSM})}(c=15) \end{pmatrix} \tag{I.1}$$

This decision profile reflects the certainties of each source towards each species. Then, three Decision Rules (DR) are tested (Table I.4). The first decision rule is nontrainable (Kuncheva, 2004), i.e. it can be directly applied from the decision profile. This rule is chosen because it has been most widely used due to its simplicity and consistently good performance (Kuncheva, 2004). Also, this rule is practical for the interpretation of the results. The second decision rule is trainable, i.e. weights $w(s,c)$ have to be computed on weight estimation sets independent of the training and testing sets. Whereas $p_s(c)$ can be seen as a certainty, $w(s,c)$ allows it to be weighted according to the errors of the source p on the weight estimation sets. For example, $w(s,c)$ can be computed from the confusion matrix of the source s on the weight estimation sets (Stavrakoudis et al., 2014). Here, $w(s,c)$ is chosen as the User Accuracy (UA) of the source s concerning species c, because it reflects the probability that the source s is right when it labels c. The third decision rule is also trainable. It is similar to the second one but more complete in the sense that it

Table I.4: Decision rules for 4 sources and 15 species. The weights w allow the sources to be weighted according to their performances. As an example, if a source has a very poor accuracy for a given species, the contribution of this source for the estimation of that species will be weighted by this accuracy in the decision rule, in order to prevent confusions.

DR	Type	Rule
1	Nontrainable	$c = \arg\max_c \sum_{s=1}^{4} \boldsymbol{d}(s,c)$
2	Trainable	$c = \arg\max_c \sum_{s=1}^{4} w(s,c)\boldsymbol{d}(s,c)$
3	Trainable	$c = \arg\max_c \sum_{s=1}^{4} \sum_{c'=1}^{15} w(s,c,c')\boldsymbol{d}(s,c')$

takes into account the probability that the source s labels c' whereas the species is c, which contributes to c. $w(s,c,c')$ is computed similarly to the UA: for instance, the proportion of cases corresponding to PAN *Platanus x hispanica* predictions, whereas the true species is *Tilia tomentosa*, can be computed from the confusion matrix of the PAN on the weight estimation sets, and is used as $w(s,c,c')$.

Three sets are then considered: training, weight estimation and testing sets as in (Stavrakoudis et al., 2014). The second is used to compute the weights $w(s,c)$ and $w(s,c,c')$. Knowing that, the following validation strategy is chosen to assess the method, i.e. to test the individual sources and the fusion:

1. All the trees of the reference map are split into two sets (2/3 and 1/3 compared to the original set) by keeping the same percentage of the trees per species.

2. The first set is split into two subsets (training and weight estimation subsets) by keeping the same percentage of trees per species. Then the weight estimation set is classified. This step is repeated 15 times because the estimations were stable beyond this value. An average confusion matrix and the weights $w(s,c)$ and $w(s,c,c')$ are computed.

3. Each of the 15 training sets is considered and the testing set is classified.

These three steps are repeated 30 times to ensure a stable result. Monte Carlo is used for random selection of the sets (Dubitzky et al., 2007).

I.1.4 Results

First, the results on the reference site are presented, with the mono-source (Section I.1.4.1) and the multi-source (Section I.1.4.2) classifications respectively. This allows the best fusion strategy to be selected and applied on the test site (Section I.1.4.2).

I.1.4.1 Mono-source classification

The OA (Overall Accuracy), κ and F-score are chosen to assess the performance of the VNIR, SWIR, PAN and nDSM sources overall and per species. For each species, the F-score is computed as the product of the producer and user accuracies (*PA* and *UA*), divided by the sum of the producer and user accuracies ($\frac{(PA)(UA)}{(PA)+(UA)}$) defined and used in (Stavrakoudis et al., 2014) for instance. The Table I.5 summarizes the quantitative performances. The figure I.5 illustrates the qualitative performances of the sources on the reference site.

VNIR sources The VNIR MNF associated to SVM exhibits the best performances with an OA of 75% and a κ of 72% (Table I.5). Compared to the VNIR all source associated to SVM, the MNF reduction slightly improves the OA by 2 pp and the κ by 3 pp. Regarding the classification algorithm comparison, the SVM is better than the RF for the VNIR MNF with OA and κ improvements of 8 pp and 9 pp. Focusing on the best source (VNIR MNF associated to SVM), *Aesculus hippocastanum*, *Cedrus atlantica*, *Juglans nigra* and *Platanus x hispanica* are classified with F-score values better than 80%. On the other hand, *Acer platanoides* is classified with a F-score value of 15%. The performance map (Figure I.5, (a)) is consistent with the OA of 75% mentioned above in the sense that the trees are well classified overall. In particular, the street trees are well classified while the main errors occur for the parks trees. The VNIR MNF associated to SVM is therefore selected for the decision level fusion, and considered as the baseline for next comparisons.

SWIR sources The SWIR MNF associated to SVM is the best SWIR source with an OA of 69% and a κ of 65% (Table I.5). Compared to the SWIR all associated to SVM, the MNF reduction improves the OA and the κ by 9 pp. Concerning the classification algorithm comparison, the SVM is better than the RF for the SWIR MNF with OA and κ improvements of 3 pp. Focusing on the best source (SWIR MNF associated to SVM), *Aesculus hippocastanum*, *Juglans nigra* and *Platanus x hispanica* are classified with F-score values greater than 80%. On the contrary, poor results are obtained for *Acer negundo*, *Acer platanoides* and *Taxus baccata* with F-score values of 29%, 2% and 39% respectively. The performance map (Figure I.5, (b)) shows that the trees are well classified overall. The

Table I.5: F-score (%) per source and per species. OA (%) and κ (%) per source. Bolded scores associated to the star (*) refer to the maximum F-score for this species. The terms "all" and MNF associated to VNIR and SWIR refer to the feature vector type (spectral bands or MNF, figure I.3, Section I.1.3.2). Blue refers to SVM while red refers to RF. The species code is presented in table II.3.

	Source					
	VNIR		SWIR		PAN	nDSM
Species code	all	MNF	all	MNF		
A.n.	39 0	**61*** 24	7 2	29 16	2 0	8 2
A.p.	9 0	**15*** 0	2 0	2 1	5 1	0 1
A.h.	92 89	**95*** 91	87 81	88 89	43 51	21 21
C.a.	**99*** 93	**99*** 94	63 39	73 66	8 8	14 7
C.au.	77 71	**78*** 71	63 33	72 68	28 24	24 18
C.o.	61 48	**65*** 53	40 23	43 42	16 29	27 33
F.s.	58 52	**64*** 63	37 30	51 54	19 32	7 7
J.n.	81 44	81 72	83 24	**84*** 81	11 12	20 18
L.s.	**56*** 27	**56*** 53	24 0	55 52	1 1	1 0
L.t.	66 23	**70*** 42	35 6	54 47	15 13	7 5
P.h.	88 63	88 80	85 54	**90*** 88	51 58	38 40
T.b.	70 44	**72*** 56	13 8	39 20	13 25	25 29
T.p.	64 39	**67*** 63	36 0	62 62	25 24	47 44
T.t.	67 50	70 64	64 30	**73*** 69	50 54	68 71
U.g.	64 43	**70*** 54	44 20	51 49	16 23	26 30
OA (%)	73 55	**75*** 67	60 37	69 66	29 37	31 35
κ (%)	69 50	**72*** 63	56 30	65 62	22 29	24 27

main errors occur for the parks trees. The SWIR MNF associated to SVM is then chosen for the decision level fusion.

PAN sources The PAN associated to RF is the best PAN source with an OA of 37% and a κ of 29% (Table I.5). It only allows the prediction of two species (majority species: *Tilia tomentosa* and *Platanus x hispanica*) with F-score values better than 50%. About the classification algorithm comparison, the RF is better than the SVM. Apart from the two majority species (F-score values of 54% and 58% respectively), the performance is poor overall with the example of *Acer platanoides* and *Liquidambar styraciflua* for which a F-score value of 1% is observed. The performance map of this source (Figure I.5, (c)) is consistent with that statement. In particular, the street trees are well classified while the park ones are mainly poorly classified. The PAN associated to RF is thus selected for the decision level fusion.

nDSM sources The nDSM associated to RF is the best nDSM source with an OA of 35% and a κ of 27%. Only one species (*Tilia tomentosa*) is classified by this source with a F-score value larger than 60%. Regarding the classification algorithm comparison, the RF is better than the SVM. Apart for *Tilia tomentosa*, the performance is poor overall with F-score values of 8%, 0% and 14% for *Acer negundo*, *Acer platanoides* and *Cedrus atlantica*. The performance map of this source (Figure I.5, (d)) reinforces this statement in the sense that the trees are poorly classified overall, especially the park ones. The street trees in the bottom right-hand corner are well classified and correspond to *Tilia tomentosa* trees (F-score value of 74%). The nDSM associated to RF is finally selected for the decision level fusion.

I.1.4.2 Multi-source classification

A classification algorithm has to be chosen in case of feature level fusion in order to classify the feature vectors composed of the previously selected features: VNIR and the SWIR MNF components, Haralick and structural characteristics. As the SVM gives an OA of 73% instead of 59% with the RF, the SVM is selected for next comparisons. The Table I.6 highlights the performance of the fusion strategies in comparison to the VNIR. Whereas the feature level fusion gives lower performance in comparison to the VNIR (73% and 69% against 75% and 72% in terms of OA and κ), the decision level fusion slightly improves the accuracy with OA and κ increases of 2 pp for the second decision rule (Table I.6). Regarding the first decision rule, the OA and κ are increased by 1 pp. The OA of the third rule is increased by 1 pp while its κ remains unchanged. The second decision rule is selected for next comparisons as it has succeeded in improving the performance in comparison to the two others.

Table I.6: OA (%) and κ (%) for the VNIR and the fusion. The performance of the feature and decision-level fusions is showed with for the latter, the three tested Decision Rules (DR).

	VNIR	Fusion			
		Feature	Decision		
			DR1	DR2	DR3
OA (%)	75	73	76	77	76
κ (%)	72	69	73	74	72

Prediction agreement complementarity from the sources toward the fusion When the VNIR and SWIR agree, 62% of the trees are well classified. Among these cases, the decision level fusion identifies 61% (-1 pp) of the trees, whereas the feature level fusion well classifies 56% of these trees (-6 pp). This shows that the decision level fusion mainly allows the original performance of the VNIR and SWIR sources when they agree to be reached. On the other hand, 13% of the well identified trees refer to cases where the VNIR is right while the SWIR is wrong. When applying the fusion, 3% and 6% of these trees are not well classified by the decision and feature level fusion strategies, respectively. This demonstrates that when the VNIR is right alone, or when the VNIR agrees with the PAN or / and the nDSM, the fusion mainly does not succeed in keeping the original accuracy of the VNIR. Finally, there are cases (11%) where the VNIR is wrong and one of the other sources is right, i.e. cases of complementarity with the VNIR. Among these cases, 6% and 5% of the trees are well classified by the feature and decision fusion strategies respectively, showing that the fusion is able to take advantage of the complementarity of the sources.

Complementarity per species from the sources toward the fusion The Figure I.11 highlights the performance of the VNIR and the feature and decision level fusions, per species. For species where the F-score of the VNIR is larger than the F-score of the other sources, i.e. for 12 species out of 15, the feature level fusion slightly improves the F-score (*Liquidambar styraciflua*: $<$1 pp, *Tilia platyphyllos*: $<$1 pp and *Ulmus glabra*: +1 pp), otherwise declines the performance. At the same time, the decision level fusion improves the F-score for five species (*Celtis australis*: +7 pp, *Fagus sylvatica*: +1 pp, *Liquidambar styraciflua*: +2 pp, *Tilia platyphyllos*: +2 pp and *Ulmus glabra*: +15 pp). Overall, the F-score of the VNIR is increased when at least one of the other sources has similar performance (ex: *Tilia platyphyllos* where the VNIR and SWIR have F-score values of 67% and 62% respectively). On the contrary, its F-score decreases when the other sources give

much worse results (ex: *Cedrus atlantica* for which the second best source has a F-score value of 73% instead of 99 % for the VNIR). These decreases can be important when the VNIR has initially small F-score values, which is the case for *Acer negundo* and *Acer platanoides* with declines to 35% and 0% respectively when the decision level fusion is applied, instead of 61% and 15% for the VNIR. On the other hand, when the F-score of the VNIR is smaller than the F-score of one of the other sources, i.e. for cases of complementarity of the other sources (three species for which the SWIR is the best: *Juglans nigra*, *Platanus x hispanica*, *Tilia tomentosa*), the decision level fusion systematically leads to F-score improvements of 10 pp, $<$1 pp and 4 pp respectively for the three considered species. Focusing on the feature level fusion, it declines the results for *Juglans nigra* (-3 pp). Finally, the decision level fusion improves the performance for 8 species out of 15, while the feature level fusion increases the F-score values for only 5 species out of 15, including the two majority species (*Tilia tomentosa* and *Platanus x hispanica*). Further results can be found in Appendices 1 and 2.

As a conclusion, the decision level fusion is selected in order to be applied on the test site.

Classification on the test site The Figure I.7 highlights the confusion matrices of the VNIR and the decision level fusion and the Figure I.8 illustrates the classification map. 93% of the trees are detected in the reference site by the automatic tree crown delineation, and 63% of these trees correspond to a unique crown. By visual assessment on the test site, no tree is omitted but oversegmentations are present, i.e. the delineation algorithm can find several sub-crowns within a single tree. The oversegmentation is not a problem because our objective is to verify that these sub-crowns belong to the right species. The decision level fusion leads to an OA improvement of 8 pp (63% against 55%) in comparison to the VNIR (Figure I.7). These improvements can be seen on the Figure I.8, where 8 more *Tilia tomentosa* street trees (35 instead of 27 on the basis of 83) are detected on the left of the image. In addition, 3 more *Platanus x hispanica* street trees (48 instead of 43 on the basis of 48) are mapped in the bottom left-hand corner. However, the fusion leads to the most important confusion between the real *Tilia tomentosa* trees and the *Platanus x hispanica* species (+8 confusions), as it can be seen on the produced maps (Figure I.8).

I.1.5 Discussions

I.1.5.1 Contribution of the spectral features

These results demonstrate that the spectral features are the main driver of the classification accuracy, a consistent result with the existing literature (Fassnacht et al., 2016). In particular, the MNF reduction improves the performance in comparison to the use of all the spectral bands, as already mentioned by (Fassnacht et al., 2014). In our case, significant improvements are obtained for species with small canopy areas (examples from Table I.5 and related canopy areas available in Table I.3: *Acer negundo*, *Fagus sylvatica*, *Liriodendron tulipifera*), and specifically for the SWIR. Both circumstances (small canopy area and SWIR) lead to a small number of pixels, thus a small number of training examples with a high number of variables when all the spectral bands are considered, incurring the Hughes effect (Hughes, 1968). Therefore, this paper reinforces the interest of using a reduction like the MNF when dealing with high dimensional data such as hyperspectral data.

Regarding the spectral regions of interest, the Figure I.9 highlights the communalities (Bailey et al., 2002) of the spectral bands in the MNF components. This allows the contributions of the spectral bands in the MNF components to be quantified, i.e. to give an idea of the variability of the spectral bands in the data set. About the VNIR data, the wavelength spectral region between 650 and 750 nm corresponding to the red edge presents high communalities compared to the other regions. In the visible region, there are comparatively small communalities but a local peak just after 500 nm. In these regions, the spectral reflectance varies a lot in function of the wavelength, which can be related to the absorption properties of foliar pigments. After the red edge, where the scattering is important and is influenced by the canopy structure (Ustin and Gamon, 2010), the communalities are comparable to those of the visible region. These contributions are consistent with the literature (Fassnacht et al., 2016; Dalponte et al., 2009) but the red edge communality is particularly significant here. In autumn, the phenological change is remarkable and depends on the species (Sheeren et al., 2016). This necessarily has an impact on the reflectance in the red edge because of the important differences of foliar pigments contents at this period. In our context, there is an interest in extracting features based on the red edge spectral bands, as the performance is better with the MNF.

Among the studied species, *Aesculus hippocastanum* is well classified (Table I.5). This is because all the trees of this species in the reference site are affected by the horse-chestnut leaf miner (Paterska et al., 2017) which necroses its foliage, making it characteristic (Figure I.10, (a)). Indeed, the leaf miner attacks the parenchyma of the leaf, explaining the decrease of the reflectance in the near infrared (modification of the structure of the leaf).

In the Visible, the drying up and leaf necrosis increase the reflectance (less chlorophyll pigments thus less absorption). On the contrary, *Acer platanoides* is poorly detected. The leaves of this species become yellow then red in autumn, making the classification difficult. In the reference site, half of all trees within this species have green leaves. One tree has largely yellow leaves and two trees have a lot of red leaves (Figure I.10, (b)). Due to the phenological change, there is a decrease of chlorophylls in favor of carotenoids (yellow, orange) and anthocyanins (red, purple) (Féret et al., 2017). These results highlight the potential of the hyperspectral data for monitoring species subject to specific diseases such as *Aesculus hippocastanum*. It also demonstrates the difficulty of the classification task when working with data acquired during periods with high phenological dynamics.

The spatial analysis of the results shows that the VNIR and SWIR mainly success in identifying the street trees (Figure I.5), and that the main errors occur for park ones. Whereas the park trees have various states, whether in terms of species or in terms of stage of development, the street trees often belong to the same species and are pruned, which is the case here. This explains that better performance for street ones, due to similar spectral traits and high number of individuals per species. Thus, there could be an interest in considering a first step of park / street trees discrimination in the classification framework, in order to apply a specific processing for the park trees.

I.1.5.2 Contribution of the textural and structural features

The textural features contribute marginally in the classification, which is consistent with the existing works (Franklin et al., 2000). However, the results demonstrate that the texture allows the two majority species (*Tilia tomentosa* and *Platanus x hispanica*, Table I.5) to be identified with a high performance similar to that of the VNIR and SWIR. In our study, the trees of these species are mainly distributed as street trees, which can explain such a performance, as mentioned for the spatial analysis of the VNIR and SWIR results. Even if the textural characteristics do not allow the whole set of species (15 here) to be classified with a high accuracy, these results encourage the use of such features in order to classify the street trees, but also to use them in another way, for example by discriminating groups of species with similar textural traits, hierarchically.

Regarding the structural features, they only contribute for the classification of one of the two majority species: *Tilia tomentosa* (Figure I.5). This result is consistent with the literature in the sense that the structural features have marginal contributions (Alonzo et al., 2014), but it is accentuated here with only one species correctly identified. This species is mainly composed of trees which belong to the same alignment. Therefore, these trees

have necessarily the same shape as they are pruned here, explaining the high classification rate for that species. However, this performance is not transposable to other study areas. As for the PAN, the nDSM data is not able to classify the whole set of species accurately and thus needs to be used in another way, for example to discriminate high vegetation from low vegetation, highest from lowest trees, in a hierarchical framework.

I.1.5.3 Best classification algorithm

While the SVM is significantly better than the RF when all spectral bands are used, the results become similar (SVM slightly better) when the MNF reduction is applied. Focusing on the PAN and nDSM-based features, the RF is better, knowing that the dimensions of the textural and structural features are comparable to those of the MNF-based VNIR and SWIR spectral features (around 10). This has to be related with the fact that the performance of the RF is similar to the SVM one when the dimension of the data is small (Pal, 2005). This leads us to consider a decision level fusion with 2 sources based on SVM (spectral) and 2 sources based on RF (textural and structural). This finding encourages the consideration of several supervised classification algorithms when dealing with heterogeneous data, thus heterogeneous dimensions, through a decision level fusion framework. From there, several approaches are possible in order to improve the performance of the method. For instance, the results from both SVM and RF could be fused at the decision level, which would amount to considering 8 sources (4 sources and 2 classification algorithms), similarly to what is carried out in the study of (Engler et al., 2013).

I.1.5.4 Best object-based fusion strategy

In this study, a complementarity analysis of the sources has been carried out in order to highlight the cases where the fusion is theoretically of benefit to the classification (Sections I.1.4.2, I.1.4.2 and I.1.7). This analysis demonstrates that the sources are indeed complementary, but that this complementarity is low. This paper highlights that in such a case, a standard feature level fusion (Alonzo et al., 2014) is not the best strategy to use. The decision level fusion (Stavrakoudis et al., 2014) success in getting better performance compared to the VNIR, considered as the best individual source, but with slight improvements. From that result, there are two ways in order to improve the performance of the fusion. Either we can aim at optimizing the complementarity of the sources, or the fusion strategy has to be improved. In particular, the necessary condition for taking advantage of a fusion is the use of complementary sources, whereas the fusion strategy is not a crucial aspect if this first condition is reached. Thus, the results demonstrate the necessity of building complementary sources before applying any fusion scheme.

Focusing on the fusion strategy, this paper shows that the feature level fusion decreases the performance of the VNIR for at least three reasons. First, the feature level fusion requires an accurate registration of the sources, whereas there subsists a residual error that is intrinsic to the registration of the VNIR and SWIR. Secondly, in our cases, the textural or structural features are mainly useful for few species. Thus, when they are added in the VNIR feature vector, they bring useless information for other species, i.e. a noise that can disturb the classification. Thirdly, adding all these features result in a vector with a higher number of variables, which is more sensitive to the Hughes effect (Hughes, 1968) when a small number of training samples are available. On the other hand, the decision level fusion does not have these drawbacks, but requires the definition of a decision rule in order to weight the different sources, a difficult task to handle. Even if the second decision rule gives the best results, improvements are also obtained with each decision rule in comparison to the VNIR. As a consequence, this paper encourages the use of a decision level fusion instead of a feature level fusion.

The proposed approach is validated on a test site in order to test its robustness. The performance is assessed on the two main species (*Tilia tomentosa* and *Platanus x hispanica*), and the obtained results demonstrate the potential of the proposed framework to automatically map *Platanus x hispanica* trees (Figure I.7, Figure I.8), a species highly represented in the south of France. Nevertheless, the VNIR alone gives already a good performance, which is consistent with the classification results for the reference site. Regarding *Tilia tomentosa*, the performance is not sufficient for an operational purpose. This is mainly explained by different phenological behaviours between the *Tilia tomentosa* trees of the reference site (used for learning) and those of the test site, probably related to distinct types of pruning for these street trees. A more representative training set (spectral library) may be used for improving the performance of such configurations. From a practical point of view, the slight cost-benefit ratio of the fusion encourages the use of the VNIR only.

I.1.5.5 Limits of the proposed approach

The first limit of the proposed approach concerns the size of the training set. While certain species are represented by more than 20 trees, a species such as *Cedrus atlantica* has only 6 trees. Such unbalanced dataset is characteristic of the urban environment, where there is a high species diversity in parks, and a low species diversity from the mono-specific street trees. This reinforces the interest of discriminating these different urban tree infrastructures. This heterogeneity does not seem to impact the classification accuracy in our context as for instance *Cedrus atlantica* is well identified. However, this small training sample size increases the Hughes effect (Hughes, 1968) and rises the risk of overfitting the

classification models (Dalponte et al., 2014). Obviously, the conclusions that are drawn through this study can not be generalized. This is why the proposed method is introduced in an automatic chain for mapping the trees of an independent site, allowing its robustness to be assessed for the two majority species, *Tilia tomentosa* and *Platanus x hispanica*, easily identifiable by photointerpretation. In order to get more training samples, Radiative Transfer (RT) models such as Discrete Anisotropic Radiative Transfer (DART) (Gastellu-Etchegorry et al., 2004) could be considered based on leaf-level spectra.

Whatever the study case, the overall accuracy percentage is quite low from an operational point of view (below 80%). However, previous studies show very variable results when attempting to classify tree species based on hyperspectral data, ranging from 60% to almost 100% according to the recent review of (Fassnacht et al., 2016). Focusing on fusion, the works of (Alonzo et al., 2014) and (Dalponte et al., 2012) reach for instance an accuracy of 83%. The proposed method results are thus in the same range than the state-of-the-art ones, our accuracies being equal to 75% based on hyperspectral data and 77% thanks to fusion. But our studies are difficult to compare. More precisely, whereas certain species are detected with high rates (e.g. *Aesculus hippocastanum*, *Cedrus atlantica*), other ones are poorly classified (*Acer platanoides*, *Liquidambar styraciflua*), declining the overall performance and indicating that an effort has to be made for discriminating these particular tree species.

In this work, we took advantage of a campaign dedicated to another application, thus this autumn acquisition is not part of the research design. During this season, the phenological differences between the species can be advantageous or not to the classification task. Regarding the positive aspect, it is clear that *Aesculus hippocastanum*, particularly attacked by the horse-chestnut leaf miner at the end of summer, is easier to identify compared to the rest of the year. This statement can also be made for *Platanus x hispanica* trees, whose senescence happens sooner than the other species in our context, except *Aesculus hippocastanum*. On the other hand, there are species for which the intraclass variability is increased such as *Acer platanoides* or *Liquidambar styraciflua*, leading to confusions with other species. For instance, phenological differences among the individuals of *Acer platanoides* are remarkable. The Figure I.10 shows significant differences in the visible, through the images, but also in the near infrared region. *Liquidambar styraciflua* has a similar behaviour. As a conclusion, our classification method suffers from phenology for some species, while other cases are easier to classify. Previous works show that the contribution of phenology is more significant in the spring because the duration of phenological evolutions is longer than in the autumn (Sheeren et al., 2016).

I.1.6 Conclusions

The objective of this study is to identify the best object-based fusion strategy that takes advantage of the complementarity of several heterogeneous airborne data sources for improving the classification of 15 tree species in an urban area (Toulouse, France). The airborne data sources are: hyperspectral VNIR (spatial resolution of 0.4 m) and SWIR (1.6 m), PAN (14 cm) and nDSM (12.5 cm). Object-based feature and decision level fusion strategies are proposed and compared when applied to a reference site. Different feature extraction techniques and classification algorithms are compared on the reference site with all the trees manually delimited and for each tree, available species identification. This allows the best fusion strategy to be selected with a view to introducing the method in an automatic process, in order to map the trees and define their species on a test site, independent of the reference site used for learning.

The hyperspectral data is the main driver of the classification accuracy with OA values of 75% and 69% for the VNIR and SWIR respectively for the reference site. The MNF reduction is of benefit to these classifications. Regarding PAN data, the extracted textural features contribute marginally to the classification with an OA of 37%. However, the street trees composed of the two majority species (*Tilia tomentosa* and *Platanus x hispanica*) are well identified. About the structural features derived from the nDSM data, they do not contribute overall with an OA of 35%, demonstrating that it is not appropriate to use this type of information to classify the whole set of species, but interesting to detect trees.

The complementarity analysis of the sources carried out in this study highlights that the complementary of the available sources is low. This is mainly due to the high performance of the VNIR in comparison to the other source performance. In this particular context, a standard feature level fusion declines the performance of the VNIR (73% against 75%), whereas the proposed decision level fusion success in slightly improving the performance (77%) for cases where complementarities have been highlighted. In order to assess the robustness of the selected method and to introduce it in an automatic process, it is tested on a test site and the obtained results are consistent to those of the reference site used for learning. Indeed, the decision level fusion improves the OA by 8 pp (63% against 55%) and allows *Platanus x hispanica* to be well identified. But the VNIR alone gives already a good performance.

Further development is necessary in order to get the best species classification approach. First, it is necessary to optimize the complementarity of the sources, for example by defining a specific source per species and by adopting a hierarchical approach. On the other

hand, as this paper highlights different behaviors of the classifications between park and street trees, it looks promising to try to discriminate these specific tree structures of the urban environment as a first step of the classification process. Finally, the training samples defined directly within the images do not allow *Tilia tomentosa* trees to be correctly identified in the test site. Spectral measurements on the field could be of interest in order to build more representative training samples.

Acknowledgements

Concerning the field data collection, thanks to the city of Toulouse and Jérôme Willm from the National Institute of Agricultural Research. Regarding airborne data, thanks to Karine Adeline and Philippe Deliot for his participation in the UMBRA campaign. Thanks to the members of the Hyperspectral imagery for Environmental urban planning (HYEP) project, funded by the French National Research Agency (ANR). Thanks to Thierry Erudel and Xavier Ceamanos for discussions about supervised classification schemes.

Disclosure statement

The authors declare no conflict of interest.

Funding

The authors would like to thank the French Aerospace Lab and the Région Occitanie for funding this research.

I.1.7 Appendices

Appendix 1: Further results about the complementarity per species from the sources toward the fusion (Figure I.11)

Appendix 2: Spatial complementarity from the sources toward the fusion (Figure I.12)

The Figure I.12 presents the results of the individual sources and the decision level fusion on an area of interest. Regarding trees much better classified by the VNIR than by the other sources, the feature and decision level fusion strategies mainly decline the performance (Figure I.12). Generally, when the performance of the other sources are similar or higher than the VNIR, the fusion results in higher scores. In particular, there are *Juglans*

nigra trees aligned along the diagonal from the top left-hand corner to the bottom right-hand corner. For some of these trees, the SWIR provides better results than the VNIR and allows the accuracy to be improved by the fusion, especially by the decision level fusion. On the other hand, the vertical street trees on the right correspond to *Tilia tomentosa* species, one of the two majority species. These trees are well identified by all the sources overall (especially by the PAN and the nDSM). The fusion takes advantage of this complementarity and all of these trees are well identified.

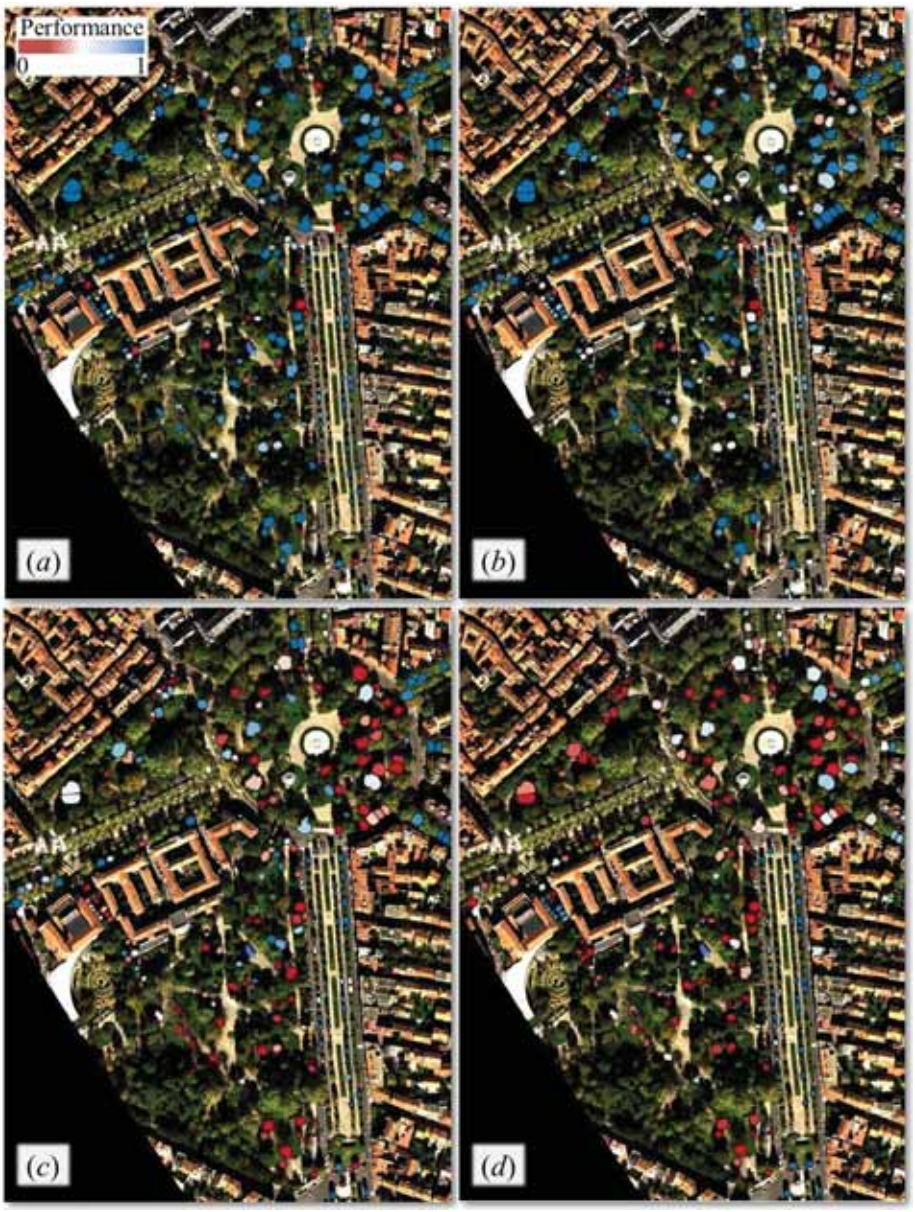

Figure I.5: Performance map per source. For each tree and each source, the performance is computed as the number of right species predictions for this tree, divided by the number of times this tree has been classified in the validation strategy (Section I.1.3.3). By way of example, a red, white or blue polygon indicates a tree which is never, one in two or always well classified, respectively. VNIR (a), SWIR (b), PAN (c) and nDSM (d) refer to the VNIR MNF associated to SVM, the SWIR MNF associated to SVM, the PAN associated to RF and the nDSM associated to RF, respectively.

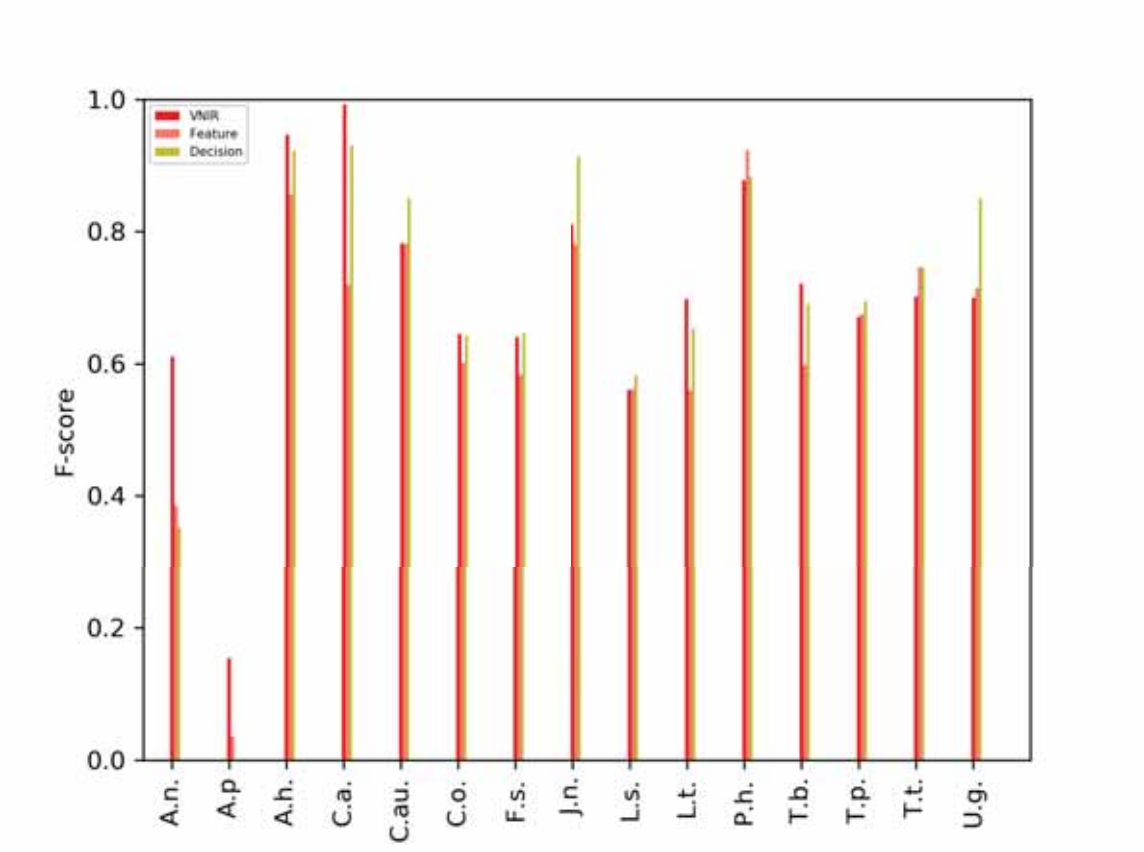

Figure I.6: F-score per species for the VNIR and for feature and decision fusion strategies. The x-axis corresponds to the species code (Table I.3).

(a)

		Predicted		
		T.t	**P.h**	**Other**
True	**T.t**	27	0	56
	P.h	0	45	3
OA (%)	55			

(b)

		Predicted		
		T.t	**P.h**	**Other**
True	**T.t**	35	8	40
	P.h	0	48	0
OA (%)	63			

Figure I.7: Confusion matrices of the VNIR (a) and the fusion (b) on the test site for the two main species (*Tilia tomentosa* and *Platanus x hispanica*). The term ”other” refers to the other species for which the ground truth is not available. The species code is presented in Table I.3.

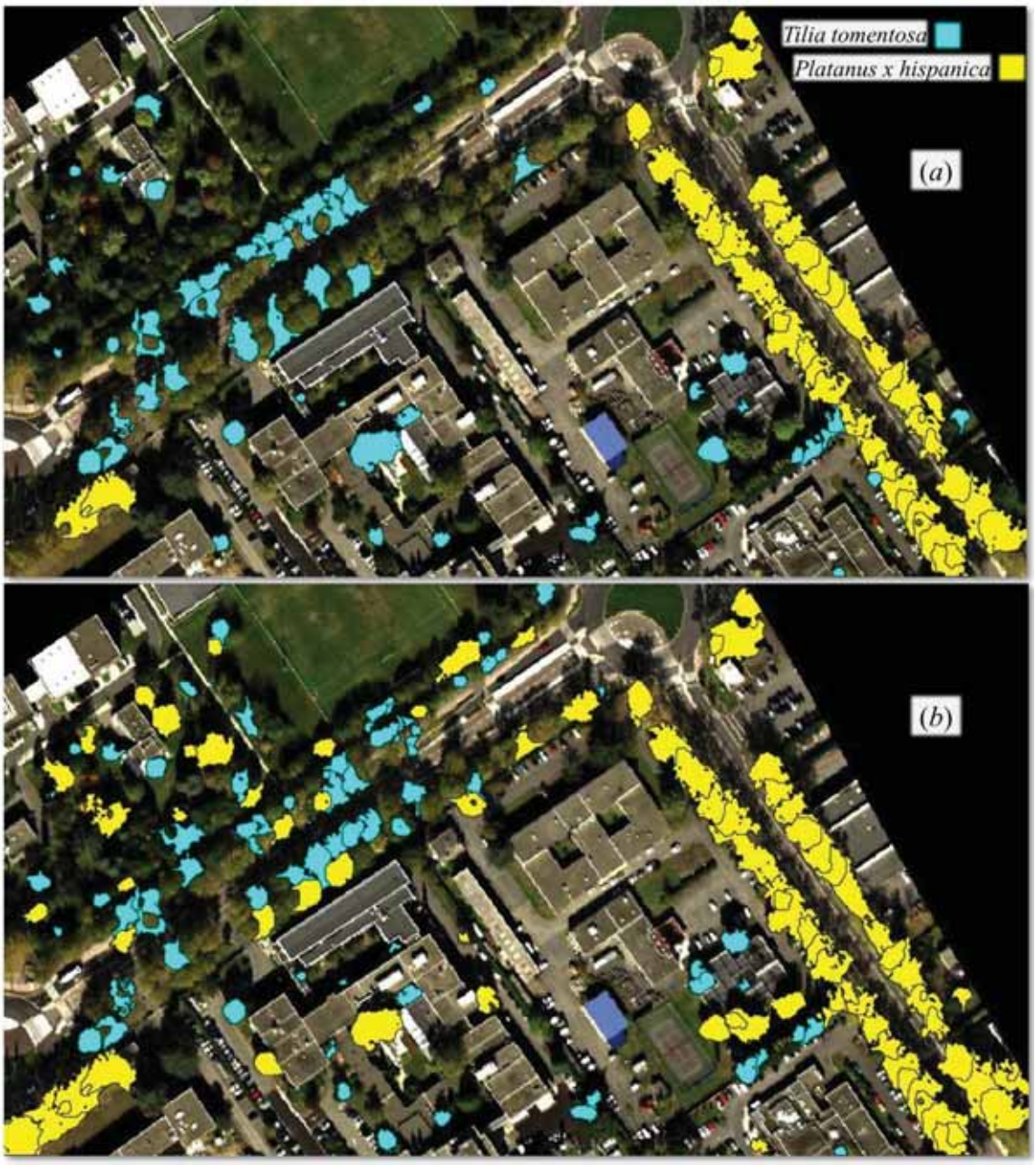

Figure I.8: Map of the VNIR (a) and the fusion (b) on the test site for the two majority species (*Tilia tomentosa* and *Platanus x hispanica*).

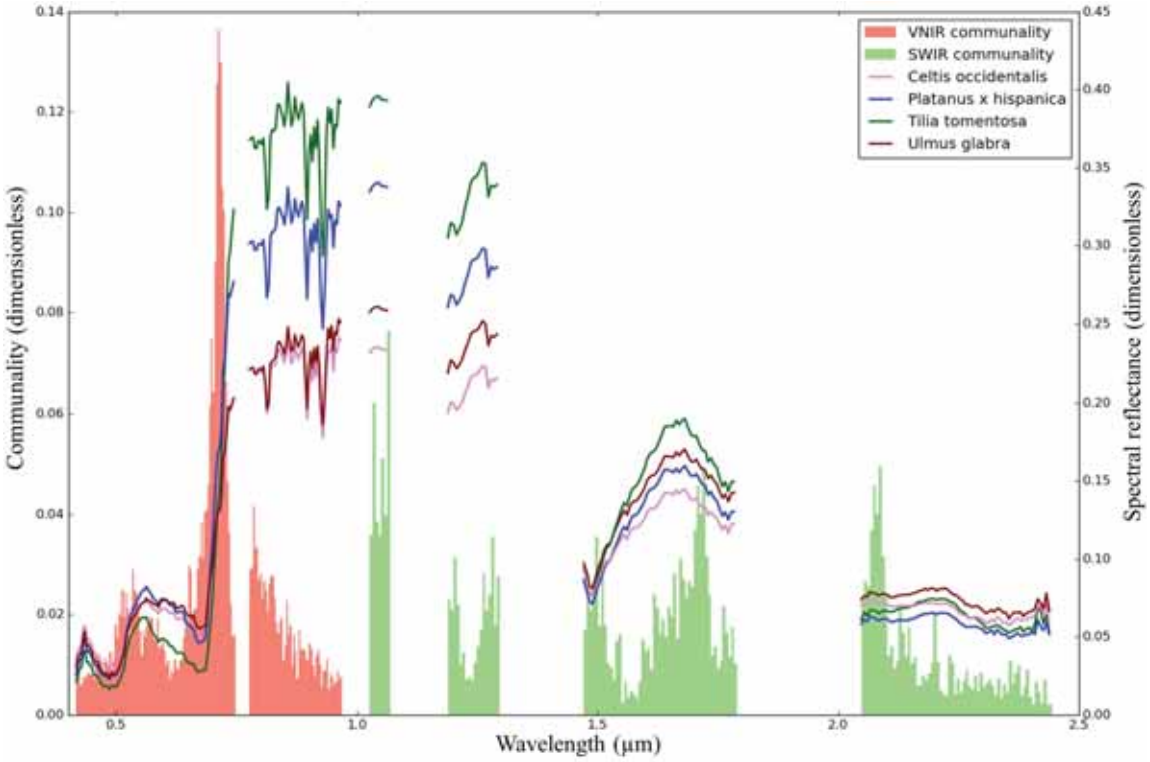

Figure I.9: Communalities of the spectral bands computed from the first 30 and 15 MNF components for the VNIR MNF and SWIR MNF sources. The spectral reflectances are computed as the average per species for the VNIR and SWIR data.

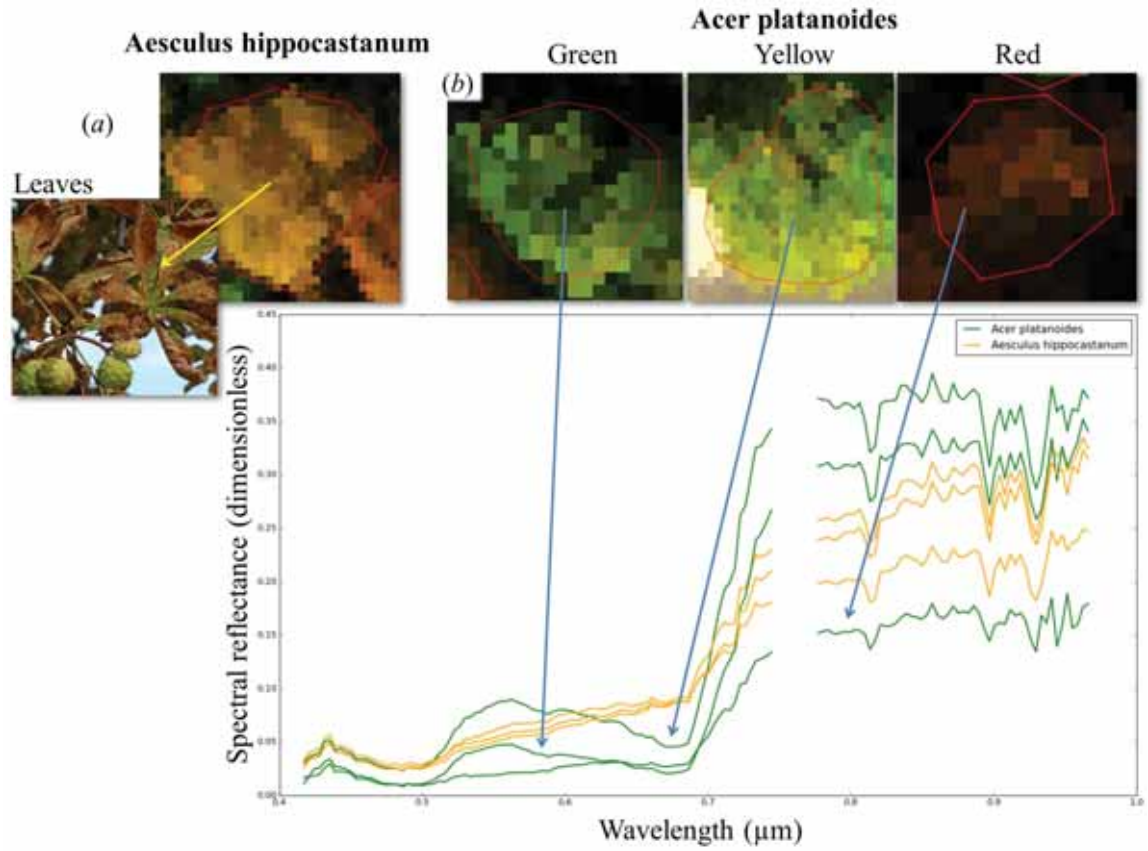

Figure I.10: Illustration of the *Aesculus hippocastanum* and the *Acer platanoides* VNIR data. Three trees are presented for each species. The spectral reflectances correspond to the average spectral reflectances per tree.

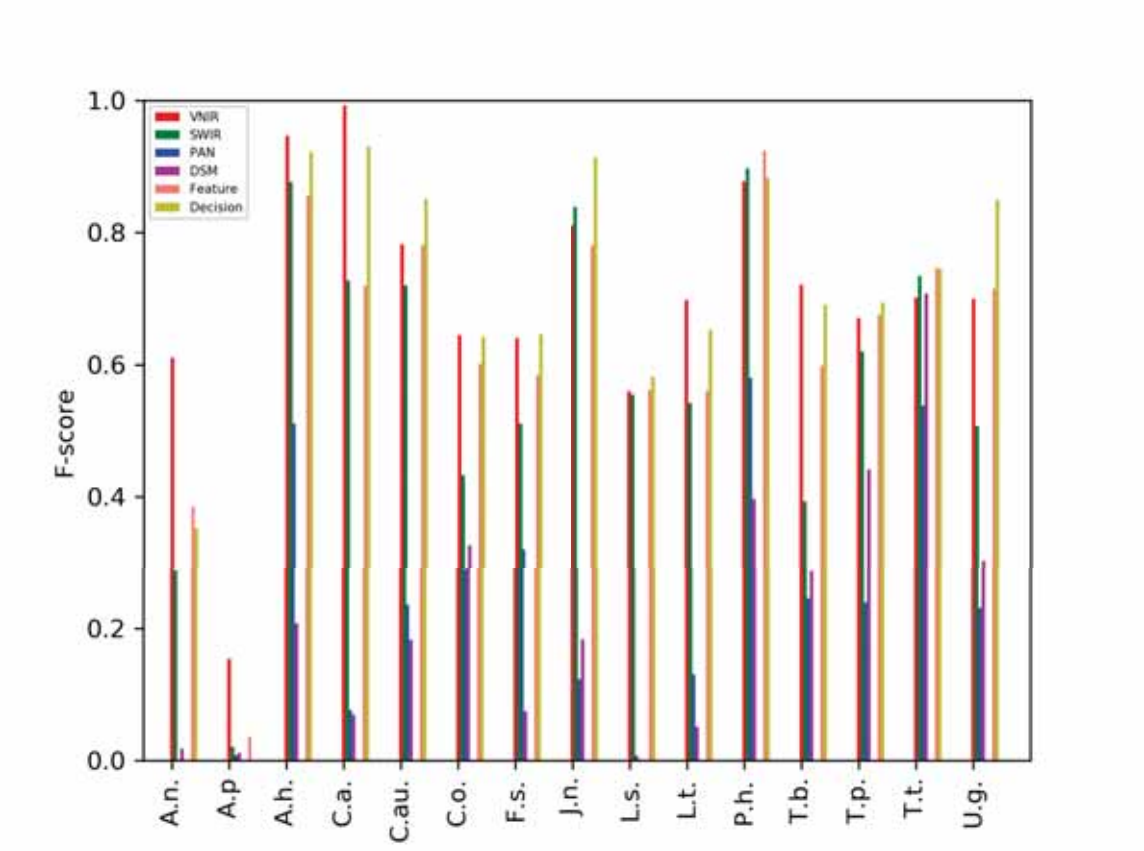

Figure I.11: F-score per species for each source and for feature and decision fusion strategies. The x-axis corresponds to the species code (Table I.3).

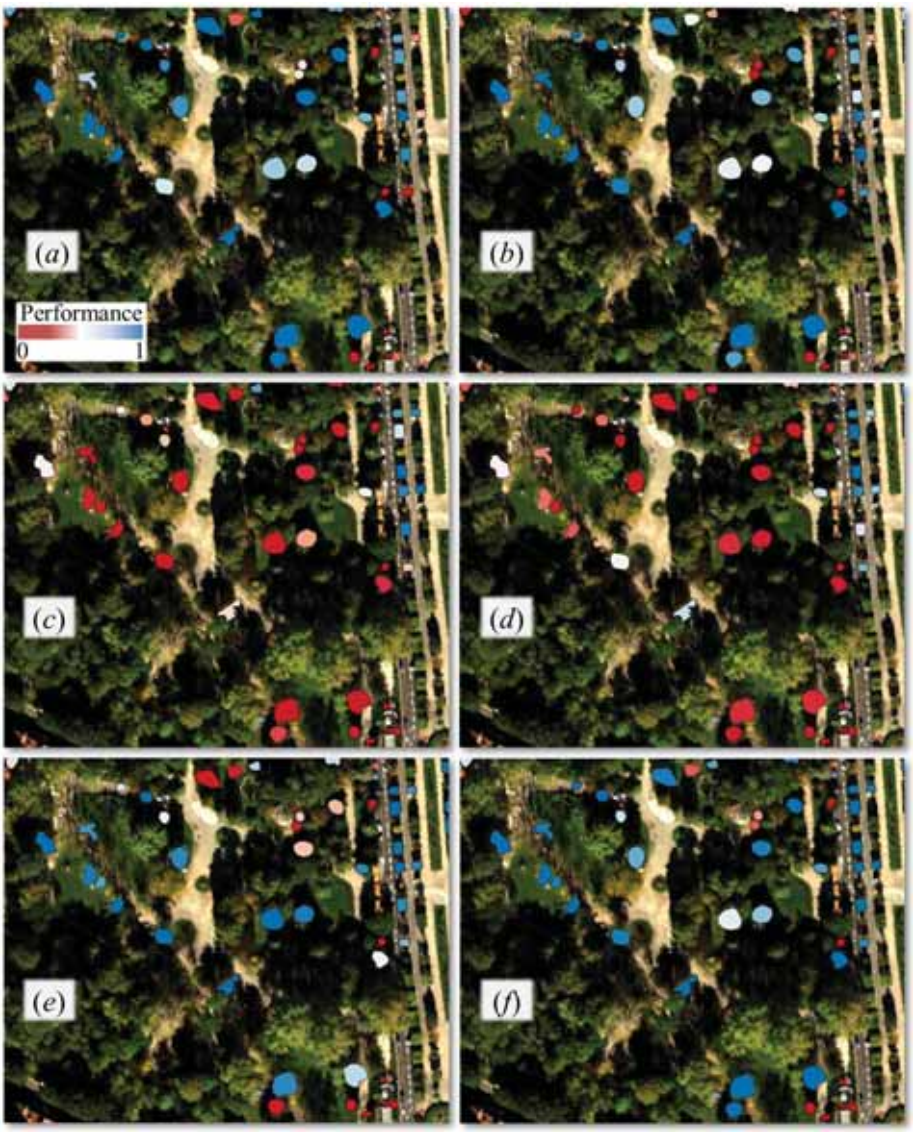

Figure I.12: Performance map from the sources toward the fusion. For each tree and each source, the performance is computed as the number of right species predictions for this tree, divided by the number of times this tree has been classified in the validation strategy (Section I.1.3.3). By way of example, a red, white or blue polygon indicates a tree which is never, one in two or always well classified, respectively. VNIR (a), SWIR (b), PAN (c) and nDSM (d) refer to the VNIR MNF associated to SVM, the SWIR MNF associated to SVM, the PAN associated to RF and the nDSM associated to RF, respectively. (e): feature level fusion. (f): decision level fusion.

Chapter II

Urban tree species classification from multiple spectral classifiers

Synthèse de l'article en français

L'objectif de cet article est de classer 5 espèces d'arbres dans un environnement urbain, Toulouse, France, en tirant profit de la richesse des données hyperspectrales à travers un ensemble de classifieurs fondé sur des mesures terrain.

Jeu de données

Le cadre de classification est similaire à celui utilisé dans le chapitre précédent. En effet, afin de sélectionner la meilleure stratégie de classification, un ensemble de classifieurs ainsi que deux méthodes de référence sont d'abord testées sur un site de référence où 5 espèces d'arbres sont préalablement identifiées et les couronnes d'arbres délinéées manuellement (94 arbres). Deuxièmement, l'approche retenue est introduite dans un processus automatique (délimitation des couronnes et classification des espèces) pour classer les espèces d'un site test, indépendant du site de référence utilisé pour l'apprentissage. Dans ce cas, la méthode est évaluée quantitativement pour l'espèce majoritaire (*Tilia tomentosa*), alors que la performance pour les autres espèces est analysée de manière qualitative en raison du manque terrain.

Comme les données hyperspectrales aéroportées VNIR ont été identifiées comme le principal moteur de la précision de la classification dans le chapitre précédent par rapport aux données SWIR, PAN et nDSM, seules les données VNIR sont considérées à travers les stratégies de classification suivantes. Le domaine spectral VNIR est particulièrement

TABLE II.1 : Principales caractéristiques du jeu de données. ”N” fait référence au nombre de bandes spectrales. GSD signifie Ground Sampling Distance. ASD correspond à Analytical Spectral Device.

	VNIR	**Field**
Source	HySpex VNIR-1600	spectroraiomètre ASD
Quantité	réflectance spectrale	réflectance spectrale
GSD	0.4 m	échelle feuille / canopée
Intervalle	0.4 - 1 μm	0.4 - 1 μm
N	160	160

pertinent car il donne des informations sur les pigments foliaires, la chlorophylle, les caroténoïdes, etc., et la structure du feuillage, Indice de Surface Foliaire (LAI), Distribution Angulaire des Feuilles (LAD), des caractéristiques biochimiques et biophysiques qui dépendent de l'espèce. Afin d'évaluer le potentiel de mesures spectrales terrain pour la classification des espèces au niveau aéroporté, des mesures ont été effectuées à l'échelle de la feuilles et de la canopée sur 7 arbres du site de référence. Les principales caractéristiques du jeu de données sont décrites dans la Table III.2, après que les prétraitements géométrique et radiométrique aient été effectués.

Méthode proposée

Une approche d'ensemble de classifieurs est proposée. Au moins un classifieur est dédié à la prédiction d'une espèce particulière. Chaque classifieur est constitué d'un SVM associé à un vecteur de caractéristiques composé de trois indices de végétation. L'apprentissage des modèles est fondée sur des données aéroportés ou terrain (feuille ou canopée). Les triplets d'indices sont choisis de manière à optimiser le F-score de chaque espèce sur l'image VNIR, à travers la prise en compte de pixels étiquetés dans l'image. Cela garantit que les indices extraits soient à la fois discriminants et invariants par rapport au changement d'échelle. À partir des votes des classifieurs, un score est calculé et permet de prédire l'espèce à l'aide d'une règle de décision. Les indices de végétation considérés dans cette étude sont ceux recensés par (Erudel et al., 2017) (plus de 100 indices). Deux méthodes de référence sont utilisées. La première est l'utilisation directe de la réflectance (appelée ”réflectance”) tandis que la seconde concatène tous les indices extraits dans un vecteur de caractéristiques (appelé ”concaténée”). La méthode proposée est appelée ”ensemble”.

Résultats

Les principaux résultats de cette étude sont donnés dans la Table II.2. L'ensemble de classi-

TABLE II.2 : OA (%) et κ (%) moyennés entre les différentes échelles des échantillons d'apprentissage (aéroporté, canopée, feuille) pour chaque méthode dans le cas du site de référence. Les scores en gras associés d'une étoile (*) font référence au meilleur score parmi les sources.

	Réflectance	**Concaténée**	**Ensemble**
OA (%)	36	58	**60***
κ **(%)**	23	44	**49***

fieurs proposé dépasse les approches concaténée et réflectance avec des valeurs moyennes d'OA et de κ de 60% (κ = 49%), 58% (κ = 44%) et 36% (κ = 23%), respectivement. La valeur de la métrique κ indique que l'approche d'ensemble réduit les confusions par rapport aux méthodes de référence. Ce comportement est particulièrement visible lorsque des échantillons d'apprentissage au niveau des feuilles sont utilisés. En effet, une valeur d'OA de 58% (κ = 46%) est obtenue dans le cas de l'ensemble de classifieurs, alors que des valeurs de 45% (κ = 27%) et 15% (κ = 2%) résultent respectivement de l'utilisation des méthodes concaténée et réflectance. Concernant l'effet du changement d'échelle, l'approche d'ensemble s'avère peu sensible. L'approche concaténée est sensible au changement d'échelle de la canopée à la feuille. Enfin, la méthode réflectance est très sensible au changement d'échelle. En conclusion, la méthode d'ensemble proposée est préférable, mais la méthode concaténée pourrait être utilisée pour des échantillons d'apprentissage aux niveaux canopée ou aéroporté.

Discussions

Ces résultats démontrent que l'ensemble de classifieurs proposé est meilleur que l'approche standard concaténée. Bien qu'il y ait peu de recherches sur l'utilisation d'une telle technique dans le contexte de la classification des espèces, ce comportement a déjà été noté dans d'autres applications (Ceamanos et al., 2010) et rend notre conclusion cohérente. Néanmoins, l'originalité de notre approche réside dans la prise en compte de classifieurs spécialisés pour chaque espèce. Ensuite, deux raisons principales peuvent être données pour expliquer pourquoi l'ensemble de classifieurs est meilleur que la méthode concaténée de l'état de l'art. Comme première explication, l'approche d'ensemble est

fondée sur la prise en compte de plusieurs classifieurs par espèce, lui conférant plus de robustesse. Deuxièmement, certains éléments du vecteur de caractéristiques peuvent être utiles pour la discrimination de certaines espèces, mais peuvent jouer le rôle de bruit pour d'autres espèces dans l'approche concaténée, provoquant potentiellement des erreurs. Au lieu d'avoir des classifieurs spécifiques aux espèces, des classifieurs dédiés à la discrimination de groupes de classes (feuillus, conifères, etc.) pourraient contribuer.

Les mesures spectrales au sol peuvent être utilisées pour la classification des espèces d'arbres dans les images. En particulier, notre recherche démontre que des résultats encourageants peuvent être obtenus en se fondant sur des mesures au niveau des feuilles, sachant que les approches de l'état de l'art considèrent généralement l'apprentissage avec des échantillons directement extraits des images (Fassnacht et al., 2016). Cela peut s'expliquer par l'utilisation d'indices de végétation qui peuvent être à la fois discriminants et invariants par rapport au changement d'échelle. De plus, l'approche retenue pour l'extraction des triplets d'indices suppose que des pixels des images étiquetés sont disponibles pour optimiser le F-score par espèce au niveau des images, sur la base d'échantillons d'apprentissage au niveau de la feuille ou de la canopée. Cela permet d'effectuer le transfert entre les deux échelles. Cette approche simple est également une limite de notre étude car il est difficile d'avoir des données étiquetées sur les images. Des techniques d'apprentissage par transfert (Tuia et al., 2016) pourraient être envisagées afin de rendre notre méthode plus opérationnelle, ainsi que des modèles de transfert radiatif (RT) (Gastellu-Etchegorry et al., 2004; Jacquemoud et al., 2009) pour simuler des échantillons d'apprentissage à partir de spectres foliaires.

Conclusions

L'objectif de cette étude est de tirer profit de la richesse des données hyperspectrales pour classer 5 espèces d'arbres dans un environnement urbain, Toulouse, France, avec une méthode d'ensemble à partir de mesures spectrales au sol. En particulier, les données hyperspectrales aéroportées VNIR sont considérées pour cette tâche, tandis que des mesures sur le terrain aux niveaux des feuilles et de la canopée acquis sur un site de référence sont destinées à être utilisées pour entraîner les modèles de classification supervisée. L'ensemble de classifieurs proposé est fondé sur la prise en compte d'au moins un classifieur par espèce, chacune étant constitué d'un SVM associé à trois indices spectraux. Une règle de décision est ensuite appliquée pour prédire l'espèce.

La conclusion principale est que la méthode proposée dépasse l'approche concaténée classique lorsque des échantillons d'entraînement au niveau de la feuille sont utilisés avec une

valeur d'OA de 58% (κ = 46%), au lieu de 45% (κ = 27%) et 15% (κ = 2%). Néanmoins, l'approche concaténée pourrait être utilisée avec des échantillons d'apprentissage au niveau de la canopée ou aéroporté. Ce travail montre que des mesures spectrales terrain peuvent être considérées pour la classification des espèces d'arbres. Ceci s'explique par les propriétés discriminantes et invariantes des indices, en plus de l'extraction automatique des triplets d'indices fondée sur l'optimisation de la précision au niveau des images.

Des travaux supplémentaires doivent être menés. Premièrement, la méthode d'ensemble doit être améliorée. En effet, la prise en compte d'autres classifieurs spécialisés dans la discrimination de certains groupes d'espèces pourrait aider. En outre, la prise en compte de techniques d'apprentissage par transfert plus avancées semble nécessaire pour améliorer les performances, ainsi que la simulation d'échantillons d'apprentissage en raison du manque de vérité terrain.

English part: Second article

The second paper is included in the next section[1].

II.1 An ensemble classifier approach for urban tree species classification from ground-based spectral references

Abstract: This study aims at identifying the best object-based classification strategy that takes advantage of the richness of hyperspectral data, for classifying 5 tree species in an urban area (Toulouse, France). Field spectral measurements at the leaf and canopy levels were carried out in a reference site, while airborne hyperspectral Visible Near-Infrared (160 spectral bands, spatial resolution of 0.4 m) data were acquired over Toulouse. We propose an ensemble classifier approach (at least one classifier per species) such as each classifier uses three vegetation indices, followed by Support Vector Machine supervised classification. Then, a decision rule based on the classifiers votes is applied to predict the species. The triplets of vegetation indices corresponding to each classifier are chosen in such way that they optimize the F-score of a given species, ensuring the complementarity of the classifiers. In this framework, the field data are intended to be used for learning (either airborne, canopy or leaf level), whereas the airborne data are used for testing, in order to assess the potential of field measurements for such classification task. Whatever the training samples level, two baseline approaches are used for comparison. A standard classification procedure using directly the spectral reflectance is chosen in order to evaluate the interest of using vegetation indices. A method which stacks all the selected indices in one feature vector is considered in order to assess the potential of the ensemble classifier. Regarding the results, the proposed method outperforms the baseline approaches in case of leaf level learning with an Overall Accuracy of 58% (κ = 46%), instead of 45% (κ = 27%) and 15% (κ = 2%) respectively. In particular, *Aesculus hippocastanum* trees are well classified because of their senescence, caused by the horse-chestnut leaf miner, and highlighted thanks to the vegetation indices. In conclusion, the proposed ensemble classifier approach improves the performance, and leaf and canopy levels learning give similar performance in comparison to the use of references from the images.

Keywords: tree species classification; urban remote sensing; hyperspectral; object-based; ensemble classifier; transfer; vegetation indices.

[1] **J. Aval**, S. Fabre, E. Zenou, D. Sheeren, M. Fauvel and X. Briottet. An ensemble classifier approach for urban tree species classification from ground-based spectral references, 2018.

II.1.1 Introduction

The tree species classification of natural forests, plantations and urban vegetation from remotely sensed data has been studied for more than four decades (Fassnacht et al., 2016; Sheeren et al., 2016). While the species knowledge is essential for the biodiversity monitoring in natural forests (Shang and Chisholm, 2014), it is a valuable information in the urban environment for effective urban planning and vegetation monitoring. Actually, the urban managers require this information since tree infrastructures can perform as freshness islands in dense and polluted cities during heatwave, and that the species composition is a crucial parameter of these islands (Doick et al., 2014). On the other hand, the spread of diseases, often species-specific ones, is made easier in urban areas due to the anthropogenic activity (Sebestyen et al., 2008). The cases of the canker stain of plane trees (*Platanus*) (Vigouroux, 2014) or the leaf miner of horse chestnut (*Aesculus hippocastanum*) (Percival et al., 2011) can be highlighted. In particular, the first one causes an incurable disease which wreak havoc in the south of France (Neu et al., 2014). The struggle against these diseases is based on several elements, including the monitoring, and requiring the species knowledge. The literature provides encouraging results in tree species classification, but it remains challenging in an urban context due to the tree diversity (species, life conditions, pruning, etc.) (Welch, 1982; Alonzo et al., 2013) with potentially a small number of trees per species (Aval, 2018).

Nowadays, airborne hyperspectral sensors (0.4 - 2.5 µm range with several hundred spectral bands) make it possible to measure the spectral reflectance of vegetation volumes with a spatial resolution of an order of magnitude of 1 m (Dalponte et al., 2009), much smaller than the size of tree crowns. This physical quantity is related to the foliar components, chlorophyll, carotenoids, etc. (Jacquemoud and Baret, 1990), and the foliage structure, Leaf Area Index (LAI), Leaf Angular Distribution (LAD), etc. (Verhoef, 1984). Being specific to each species, the spectral reflectance is a candidate of interest for object-based tree species classification as demonstrated by several studies (Clark et al., 2005; van Deventer et al., 2013). Even if the reflectance can be directly used for identifying the species (Ghiyamat et al., 2013), other works have investigated the interest of applying feature extraction techniques, in order to increase the inter-species variability while decreasing the intra-species variability. Whereas studies use Minimum Noise Fraction (MNF) components (Ghosh and Joshi, 2014), other ones select automatically the most discriminative bands (Fassnacht et al., 2014), by way of the consideration of spectral derivatives (Datt, 2000), continuum removal (Fassnacht et al., 2014) or vegetation indices (Clark and Roberts, 2012). These transformations improve the classification performance in general (from a few percent points to around 10pp). However, it seems that for such classifica-

tion task, these approaches do not allow the richness of the hyperspectral data to be fully exploited. Indeed, the features derivation is not based on an optimization process, for instance a process where the spectral features would be chosen to reach the highest level of species prediction accuracy.

Once the spectral features are extracted, the current tree species classification paradigm generally uses a supervised classification algorithm such as Support Vector Machine (SVM) or Random Forest (RF) (Féret and Asner, 2013; Sheeren et al., 2016). According to (Aval, 2018) in a data fusion background, a drawback of using one classifier, i.e. one feature vector and one classification algorithm (well known as feature level in the fusion context), is that the method is sensitive to the Hughes effect (Hughes, 1968) when the number of variables is much larger that the number of examples in the training set. Also, some elements of the feature vector can be useful for the discrimination of certain species, but can perform as noise for other species. In a fusion context, we highlight the potential of a decision level fusion approach instead of a standard feature level one, knowing that the decision fusion considers several classifiers, as much as there are sensors subject to fusion (Aval, 2018). To return to the subject, hyperspectral data can be viewed as several sub-spectral feature vectors that can be classified by as many classification algorithms, which would allow to deal with the previously mentioned issues. This type of approach can be designed through the ensemble classifier framework (Kuncheva, 2004; Engler et al., 2013). For example, at least one classifier could be dedicated to the prediction of a given species. On the other hand, the samples used for training these models are often directly extracted from the images, whereas for an operational purpose the use of field measurements could be investigated. Even if previous studies show that field and airborne spectral reflectances are often incomparable, especially in the case of leaf level ground measurements because of the variability of the canopy structure (Roberts et al., 2004), spectral features such as vegetation indices can be both discriminative and invariant to the change of scale (Cho et al., 2008). Moreover, the use of ground-based spectral references has already proven its potential in other contexts such as crop mapping (Nidamanuri and Zbell, 2011).

Summarizing the existing literature, there is minimal focus on an ensemble classifier approach for classifying tree species based on hyperspectral data. Moreover, ground-based spectral references are often not considered for such purpose. The objective of this study is then to classify 5 species in an urban environment, Toulouse, France, by taking advantage of the richness of hyperspectral data through an ensemble classifier approach based on field spectral references. In particular, the following issues are addressed:

1. Is the ensemble classifier approach the best classification strategy?

2. What is the potential of ground-based spectral references for tree species classification?

II.1.2 Materials

The materials used in this study are similar to those described in (Aval, 2018).

II.1.2.1 Study area

The study area is located in Toulouse, the fourth city in France with about 500,000 inhabitants (43.6 °N, 1.44 °E). According to vegetation managers, Toulouse should have approximately 140,000 trees. Two study sites are considered (Figure II.1, (a)): a reference site with a tree reference map (detailed in Section II.1.2.4) for comparing the ensemble classifier with other approaches and selecting the best one, and a test site for assessing the potential of that selected approach through an automatic process (Section II.1.2.5).

II.1.2.2 Airborne data

The airborne hyperspectral data were acquired on September 22, 2016 at 10:00 Universal Time (UT) during a campaign organized by the French Aerospace Lab (ONERA). The sun zenith angle was approximately 70 °. Further information about this campaign can be found in (Adeline et al., 2013). The HySpex Visible Near-Infrared (VNIR) (Köhler, 2016) system was installed on board an aircraft whose flight height was approximately 2,000 m over the study area. This system consists of an hyperspectral push broom camera with 160 spectral bands (0.4 μm - 1 μm). Focusing on the spatial resolution, the VNIR camera has a pixel Field Of View (FOV) of 0.18 mrad and 0.36 mrad across and along track. This results in spatial resolutions of 0.4 m and 0.8 m across and along track.

II.1.2.3 Preprocessing

The French Mapping Agency (IGN) provides us with a georeferenced Digital Surface Model (DSM) with a spatial resolution of 0.125 m. First, the VNIR hyperspectral image is registered on the DSM by defining Ground Control Points (GCP) using QGIS software and *gdalwarp* module from GDAL. Nearest neighbour resampling is chosen to preserve the original spectral data. Also, the Thin Plate Spline (TPS) transformation (Duchon, 1977) is applied for its ability to correct the deformations locally. Because the pixels have rectangular shapes with the longer side along track, a square grid with a spatial resolution

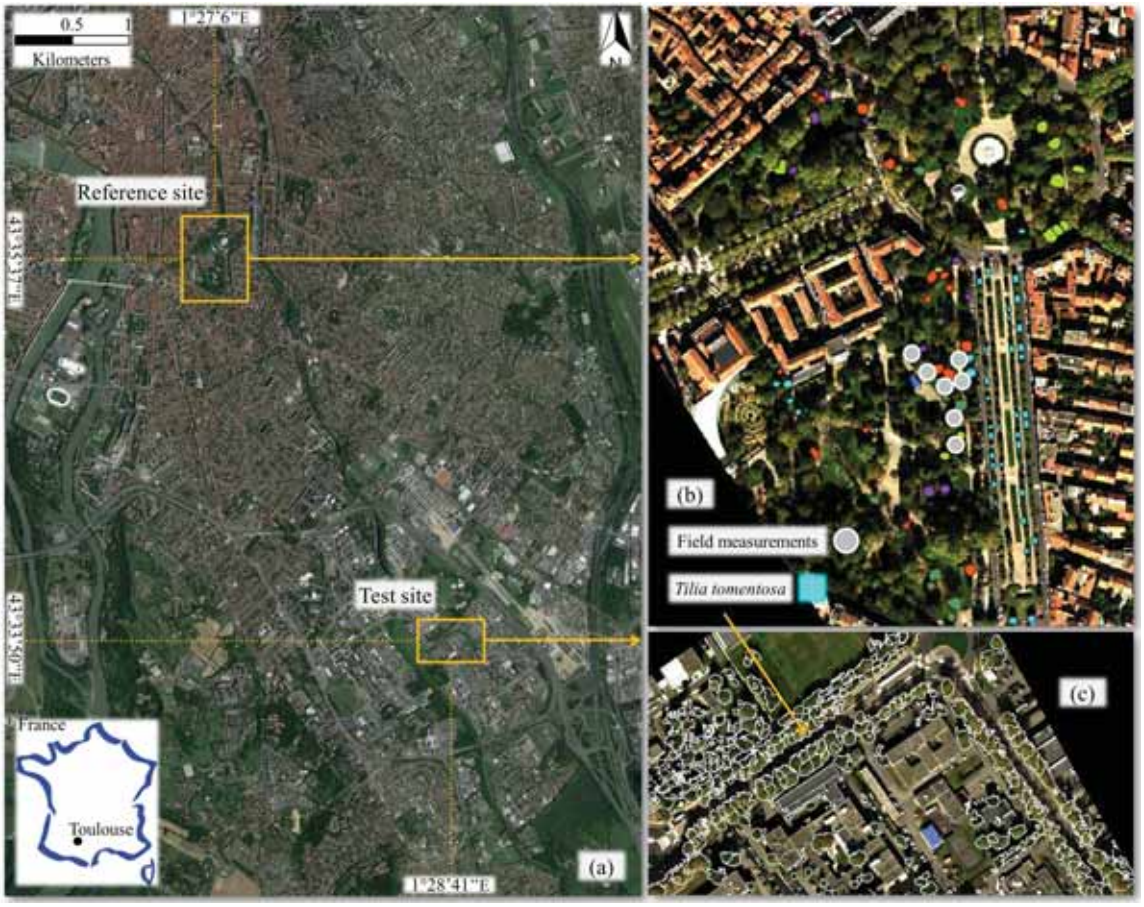

Figure II.1: (a) Overall view of the downtown part of Toulouse from Google Earth. The yellow rectangles indicate the two study areas: reference and test sites. (b) Reference site with park and alignment trees represented on the French Aerospace Lab airborne VNIR data. The coloured polygons indicate the delineations and the species of the inventoried trees in the reference site. Trees that were subject to field measurements are showed. (c) Test site with park and alignment trees, especially composed of *Tilia tomentosa* alignment trees, represented on the VNIR. The white outline polygons correspond to the automatic tree crown delineation results.

of 0.4 m (minimum between the rectangle sides) is chosen in order to preserve the original data. The error related to registration quality is less than a pixel according to visual assessment. Secondly, the hyperspectral data are atmospherically corrected to deal with spectral reflectances with the COCHISE platform (Poutier et al., 2002) based on MOD-

TRAN and assuming a flat scene. Finally, the spectral signatures have been smoothed with a Savitzky-Golay filter (Savitzky and Golay, 1964) for reducing noise.

II.1.2.4 Reference site and tree reference map

A tree reference map is built on the reference site from an existing inventory delivered by Toulouse city and from a field campaign (Figure II.1, (b)). There are 94 trees distributed among 5 species in this reference map, with an unbalanced sample which is representative of urban areas (Table II.3). The trees are delineated manually, ensuring that the selected pixels belong to trees. This avoids distorting the assessment of the classification because of delineation errors. Among these trees, 7 trees were subject to field measurements, 1 per species plus 1 more for each of the two majority species: *Aesculus hippocastanum* and *Tilia tomentosa*. An Analytical Spectral Device (ASD) spectroradiometer was used to measure the spectral reflectance on the reflective domain (0.4 μm - 2.5 μm). Whereas leaf level measurements were acquired the same day than the airborne campaign, canopy level measurements were acquired 6 days after (September 28) because the Toulouse city cherry picker employed for this acquisition was available only at this date. There has been no rain during this period. The field spectra are resampled to the VNIR hyperspectral image spectral resolution through a linear interpolation as the spectral sensitivities are comparable. The main characteristics of the trees in the reference map are given in Table II.3. The species code (Table I.3) will be used in Sections II.1.4 (results) and II.1.5 (discussions). Moreover, the Figure II.2 illustrates the spectral signatures, highlighting the differences between the different levels of measurement.

Table II.3: Main characteristics of the trees in the reference map. The training samples (last three columns) correspond to the trees that were subject to field measurements (Figure III.1, (b)).

Species scientific name (species code)	**Stem count**	**Canopy area (m^2)**	**VNIR pixel count**	**Airborne training pixels**	**Canopy training samples**	**Leaf training samples**
Aesculus hippocastanum (**A.h.**)	23	1433	7936	267	7	9
Celtis australis (**C.au.**)	10	1032	5856	66	5	4
Fagus sylvatica (**F.s.**)	10	529	2897	52	4	4
Juglans nigra (**J.n.**)	12	904	5021	213	5	4
Tilia tomentosa (**T.t.**)	39	1547	8397	219	10	4

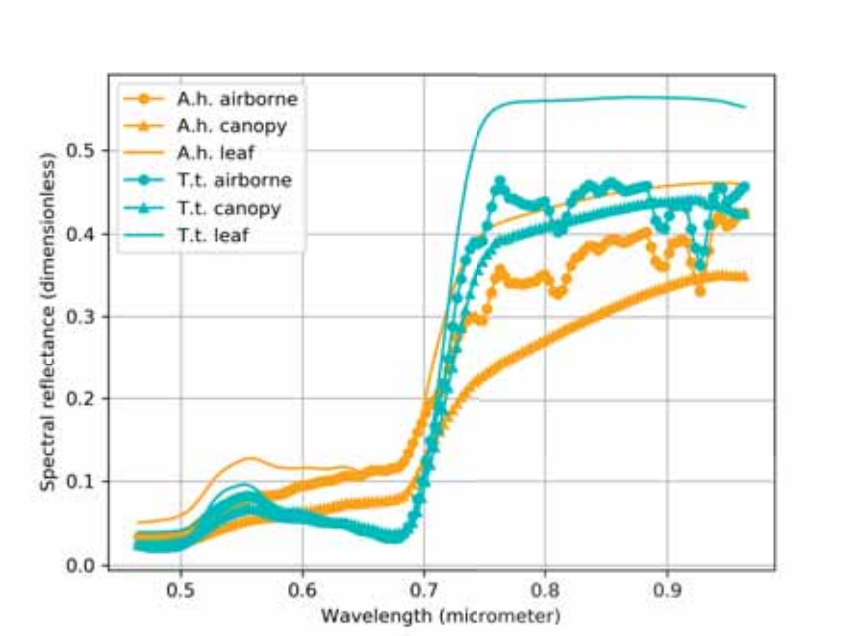

Figure II.2: Illustration of the *Aesculus hippocastanum* (A.h.) and *Tilia tomentosa* (T.t.) VNIR spectral signatures measured at three different levels: airborne, canopy, leaf. Each signature is computed as the mean of the spectral reflectances. The species code is presented in table II.3.

II.1.2.5 Test site and tree crown delineation

This test site is independent from the reference site and far from it in order to avoid spatial autocorrelation (Figure III.1, (a)). It is mainly composed of majority species alignment trees (*Tilia tomentosa*), easily identifiable by visual interpretation with the help of Google Street View (and checked by Toulouse city). This visual information is exploited to generate the reference classified product used to assess the performance of the proposed method (Figure III.1, (c)). In order to automate the processing chain, the trees are delineated automatically thanks to the method developed in (Adeline, 2014), based on (Iovan et al., 2008), and used in (Aval, 2018) (Figure III.1, (c)). Especially, the principle of the algorithm is to choose the highest pixel of the Canopy Height Model (CHM) as the first pixel of the first delineated tree. Then, the height is decremented and the corresponding pixel is either assigned to that first tree if it is at a distance less than 2 m here as in (Adeline, 2014), or assigned to a new tree, and so on. The produced delineation map allows localizing the trees for which species have to be defined. The performance is assessed quantitatively for

the majority species (*Tilia tomentosa*), while the other species are analysed in a qualitative way because of the lack of ground truth.

For both sites, the pixels in the shade are removed from the classification process to get suitable testing samples in comparison to the canopy training ones acquired in the sun (i.e. vegetation volumes directly illuminated by the sun). To do so, the area of the spectral reflectance (albedo) is computed for each pixel within each crown and all the pixels above the 75 percentile are kept, because the classification performance is stable beyond this value.

II.1.3 Methods

II.1.3.1 Ensemble classifier framework

The main steps of the proposed method are presented in Figure II.3. The term ”classifier” refers to both the feature vector and the associated supervised classification algorithm. In our context, there are different feature vectors (triplets of spectral indices), but the same algorithm is used for each triplet, the Support Vector Machine (SVM). Focusing on the Figure II.3, from the training data, the training step is carried out based on either airborne pixels, canopy samples or leaf samples. There is then a first testing step on a subsample of the hyperspectral image, in order to select the best classifiers, i.e. those that are robust to the change of scale. The resulting ensemble classifier is used for classifying the rest of the image. There are two possible cases for the definition of the subsamples. Regarding the reference site, ”1” and ”2” (Figure II.3) refer to two distinct subsamples of the reference site (50% / 50% split) by keeping the same percentage of the trees per species. This split is repeated 5 times to ensure a stable result. The Monte Carlo process is used for random selection of the sets (Dubitzky et al., 2007). Focusing on the real application, ”1” and ”2” correspond to the entire reference and test sites, respectively

II.1.3.2 Derivation of the classifiers

Many spectral indices used for characterizing biochemical components of vegetation such as chlorophyll, carotenoid, water, nitrogen, etc., have been reviewed by (Erudel et al., 2017) (more than 100 indices). We assume that these indicators are useful to discriminate the tree species considered in this study. From either the airborne training pixels, the canopy training samples or the leaf training samples, each possible triplet of indices (justified below), which can differ between the scales, is computed and allows as many SVM models to be trained, and applied to the subsample of the hyperspectral image. This ensures that the extracted indices are both discriminant and invariant to the change of

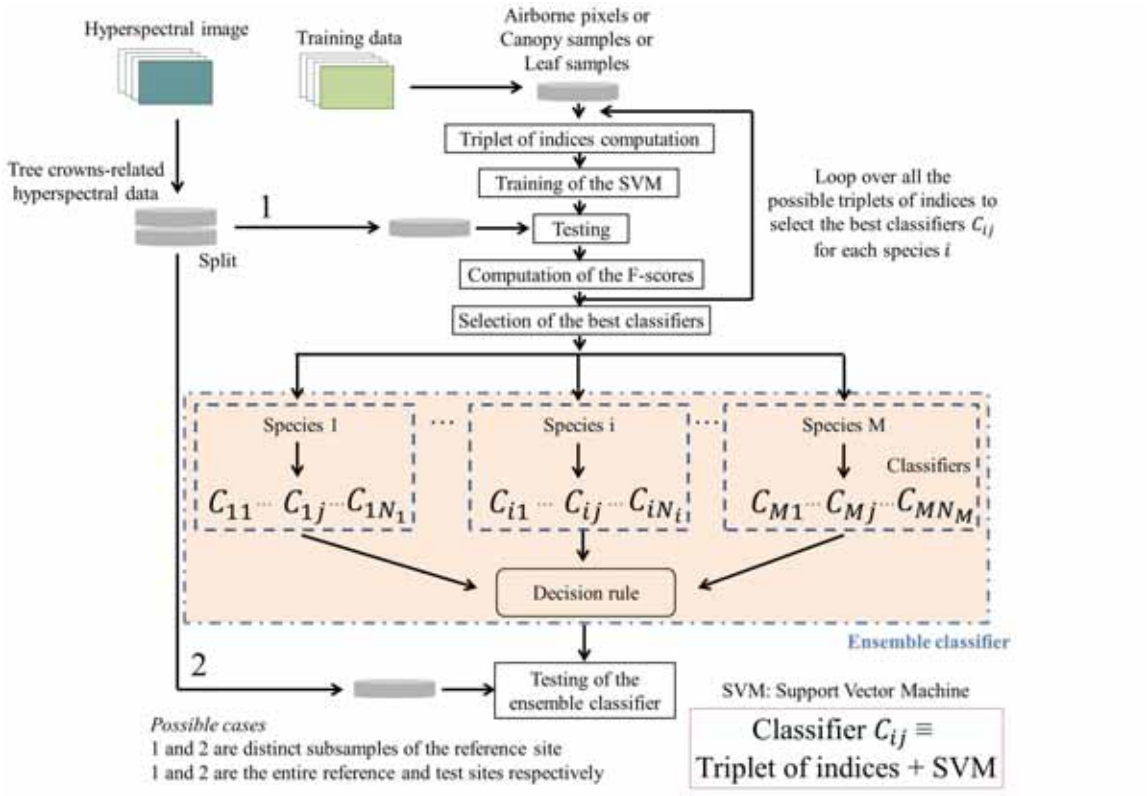

Figure II.3: Flowchart of the proposed ensemble classifier approach.

scale. Each of these classifiers is selected through the computation of the F-score, for each species (example of definition and use in (Stavrakoudis et al., 2014)) (loop in Figure II.3). The SVM supervised classification algorithm is chosen because it has demonstrated good performance in the literature (Féret and Asner, 2013), in particular when there are limited training samples, and because it is a non-parametric algorithm. The pixels within the crowns are classified and the decisions combined by majority vote, an efficient strategy according to (Alonzo et al., 2014). The Table II.4 summarizes the performance of the best classifiers, overall and per species. There is an optimal classifier in terms of Overall Accuracy (OA). But at the same time, for each species, it exists a significantly better classifier than the optimal one in terms of F-score. Thus, it is reasonable to assume that the combination of these species-specialized classifiers (i.e. an ensemble classifier approach) would improve the performance of the optimal one. Each feature vector is then composed of three spectral indices, as above this number of indices the F-score obtained for each

Table II.4: F-score (%) per classifier and per species in the case of canopy training samples and three vegetation indices. OA (%) and κ (%) per classifier. For example, $C^*_{F.s.}$ is the classifier that gives the best F-score for *Fagus sylvatica*, while C^*_{OA} is the classifier that gives the best *OA*. Bolded scores associated to the star (*) refer to the maximum F-score for this species. The species code is presented in Table II.3.

	C^*_{OA}	$C^*_{A.h.}$	$C^*_{C.au.}$	$C^*_{F.s.}$	$C^*_{J.n.}$	$C^*_{T.t.}$
OA (%)	**79***	70	70	47	60	51
κ (%)	**72***	54	61	36	46	31
A.h.	88	**100***	82	67	76	0
C.au.	67	0	**89***	40	46	0
F.s.	89	89	89	**100***	0	0
J.n.	71	0	48	31	**100***	62
T.t.	80	74	67	27	57	**86***

species-specific classifier are not better than the ones given in Table II.4. Indeed, F-score values of 100% are obtained with three indices for *Aesculus hippocastanum*, *Fagus sylvatica* and *Juglans nigra*. Focusing on *Celtis australis* and *Tilia tomentosa*, the F-score values showed in Table II.4 cannot be improved because of confusions for certain individuals of these species. Knowing that there is at least one classifier per species, the selection of the best classifiers is carried out based on two parameters: the minimum required classifier $F-score$, $F-score_{min}$ (0.1), and the maximum number of classifiers per species (5), both fixed after testing several values (between 0.1 and 0.5 in 0.1 steps for the first parameter, between 5 and 20 in 5 steps for the second one).

Decision rule We introduce for each species i, $F_i = \frac{\sum_{j=1}^{N_i} F-score_{i,j} \cdot w_j}{N_i}$, knowing that N_i refers to the number of classifiers for a species i, and $F-score_{i,j}$ is the F-score of the jth classifier of the species i. w_j is a boolean equal to 1 if the classifier under consideration votes for its species, else 0. The decision rule is then defined as follows:

- If the best F_i is higher than F_{min} (0.5) and higher than the best second one with a difference up to ΔF_{min} (0.2), both fixed after testing multiple values, we predict the species corresponding to $\arg\max_i F_i$.
- Else we predict the rejection class (called "rejected" in the next sections).

The proposed method is referred as "ensemble" in the next sections.

II.1.3.3 Baseline approaches

Two baselines are used for comparison. A first baseline consists in using directly the spectral reflectance in order to assess the interest of the vegetation indices (called "reflectance"). A second baseline is dedicated to assess the performance of the ensemble classifier, compared to an approach where the triplets of indices of each classifier are stacked in the same feature vector, and each redundant index removed (called "stacked"). Indeed, if an index appears several times among the triplets of indices, it is considered only once. Therefore, the dimension of the methods tested is not the same.

II.1.4 Results

Regarding the reference site, the Table II.5 gives the Overall Accuracy (OA) and κ of the approaches tested in this study. The Figure II.4 shows the tree species maps obtained for the test site with the stacked and ensemble classifier approaches based on leaf level measurements. In order to get a deeper analysis for the reference site, the confusion matrices of these approaches at the canopy level are given in Figure II.5.

Table II.5: OA (%) and κ (%) for the different methods depending on the training samples level when applied to the reference site. Bolded scores associated to the star (*) refer to the maximum score for a particular level.

	Reflectance	Stacked	Ensemble
Airborne	53	**62***	56
	40	**52***	46
Canopy	41	**68***	66
	27	53	**55***
Leaf	15	45	**58***
	2	27	**46***
Average	36	58	**60***
	23	44	**49***

Focusing on the leaf level training samples, the proposed ensemble classifier outperforms the stacked and reflectance methods with OA values of 58% (κ = 46%), 45% (κ = 27%) and 15% (κ = 2%) for the reference site, respectively. The significant improvement in terms of κ (+19pp and +44pp) demonstrates that the proposed approach reduces a lot

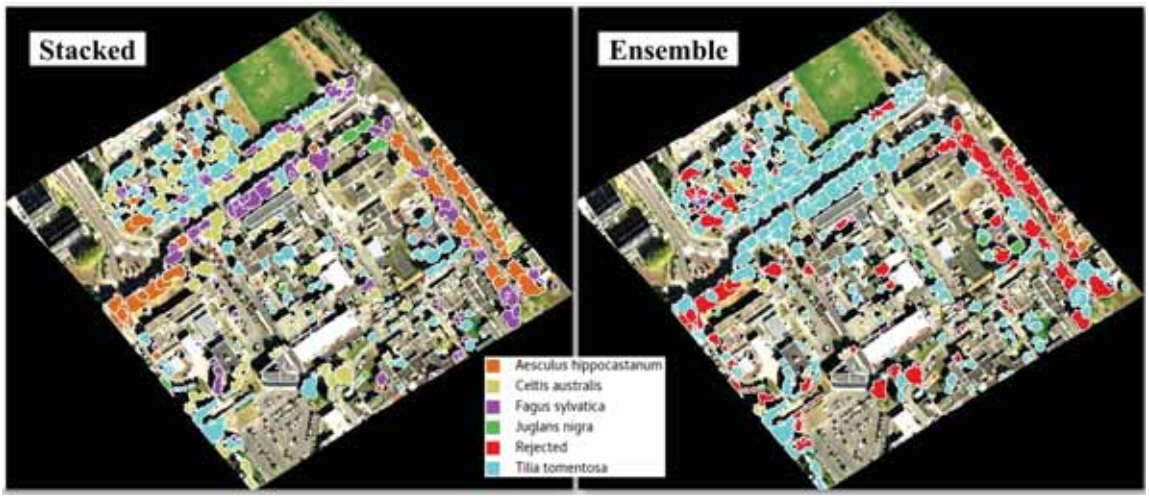

Figure II.4: Comparison maps between the stacked approach (left) and the ensemble classifier (right) based on leaf training samples when applied to the test site. The "Rejected" class concerns only the ensemble classifier approach.

the confusions between species. When assessing the ensemble classifier through an automatic process (on the test site), this behaviour is accentuated. While the stacked approach leads to the identification of 6% of the *Tilia tomentosa* alignment trees, the proposed one reaches 98% of correct detections. This is visible in Figure II.4 where the *Tilia tomentosa* alignment is well reconstructed thanks to the ensemble classifier, whereas the stacked approach leads to confusions with other species such as *Celtis australis* or *Fagus sylvatica*. Also, the proposed method predicts the rejected class for the majority of the alignment trees on the right of the scene (*Platanus x hispanica*), a consistent result as the supervised algorithms have not been trained for this species. However, the population of *Tilia tomentosa* is overestimated in the park at the top of the image. In conclusion, it is preferable to use the proposed approach in case of leaf training samples.

At the canopy level, the performance of the ensemble classifier is similar to that of the stacked approach with OA values of 66% ($\kappa = 55\%$) against 68% ($\kappa = 53\%$) respectively. These approaches remain significantly better than the reflectance method whose OA and κ values are equal to 41% and 27%. Focusing on the confusion matrices given in Figure II.5 in case of canopy training samples, the main phenomenon that can be observed is a confusion between *Tilia tomentosa* and the other species. For the stacked approach, 30% of *Juglans nigra*, 85% of *Fagus sylvatica*, 50% of *Celtis australis* and 15% of *Aesculus*

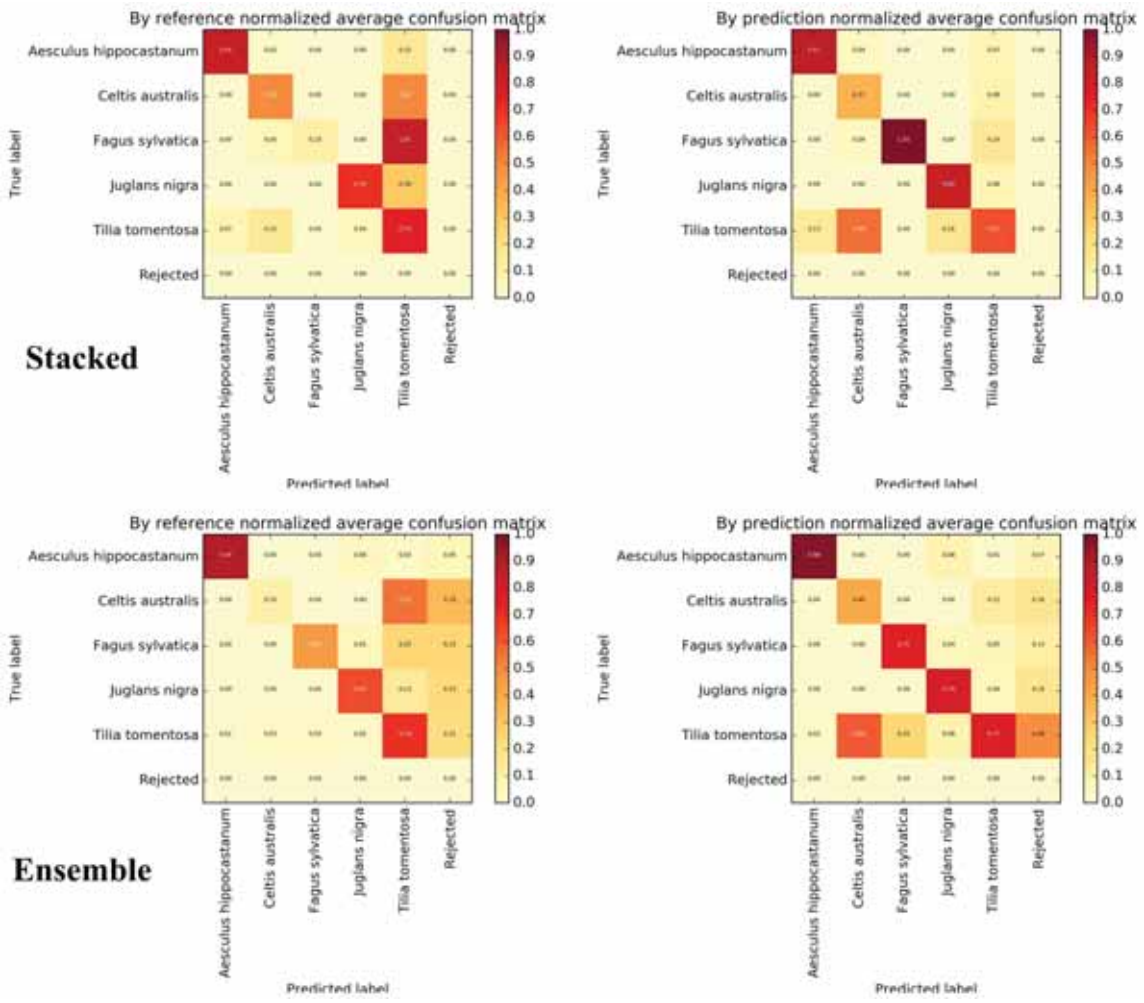

Figure II.5: Comparison confusion matrices between the stacked approach (top) and the ensemble classifier (bottom) based on canopy training samples when applied to the reference site. The matrices on the left are normalized by the reference (producer accuracy) while the matrices on the right are normalized by the prediction (user accuracy). The "Rejected" class concerns only the ensemble classifier approach.

hippocastanum are confused with *Tilia tomentosa* trees. Although the ensemble classifier approach gives similar results in terms of OA and κ, it causes less confusion between the species thanks to the consideration of the rejected class. Instead of being labelled as *Tilia tomentosa*, the trees that were subject to confusion are considered as the rejected class. For

that reason, the proposed ensemble classifier is the most appropriate choice when canopy training samples are used, but the stacked approach could be also considered for such task.

Regarding airborne training samples, the performance of the three approaches is quite similar with OA values of 53%, 62% and 56%, respectively (κ of 40%, 52% and 46%). The stacked approach is better than the ensemble classifier with improvements of OA and κ of 6pp. In this case, the stacked approach has to be preferred, although the ensemble classifier and the reflectance method could be also considered. On average among the training samples levels, the proposed ensemble classifier is slightly better than the stacked approach with an OA value of 60% (κ of 49%) in comparison to 58% (κ of 44%). The better improvement in terms of κ (+5pp) compared to that of the OA (+2pp) indicates that the ensemble method reduces more the confusions between the species. The performance of the reflectance approach is much lower in comparison to the ensemble (-24pp in terms of OA and -26pp of κ). Overall, the ensemble classifier approach is the most appropriate.

Regarding the effect of the change of scale, the ensemble classifier OA decreases down to 58% ($\kappa = 46\%$) compared to 66% (κ = 55%) from canopy to leaf level, while the OA is equal to 56% with airborne training samples. Regarding the stacked approach with leaf training samples, the OA and κ are declined to 45% and 27% in comparison to 68% and 53% at the canopy level, whereas these values are equal to 62% and 52% based on airborne training samples. Focusing on the use of the reflectance approach, a decrease of the performance is observed from the airborne to leaf level, by way of the canopy level with OA values of 53% ($\kappa = 40\%$), 41% ($\kappa = 27\%$) and 15% ($\kappa = 2\%$), respectively. In conclusion, the proposed ensemble classifier is not very sensitive to the change of scale. The stacked approach is more sensitive to this parameter, in particular from the airborne or canopy level to the leaf level. Finally, the reflectance method cannot be used for classifying tree species if canopy or leaf training samples are considered whereas airborne samples can be used.

II.1.5 Discussions

II.1.5.1 Best classification strategy

These results highlight the interest of using vegetation indices instead of the original spectral reflectance for the classification of tree species when ground-based spectral references are considered. While studies demonstrate the advantage of using such indicators in the context of vegetation species identification (Erudel et al., 2017), other works obtain similar or better results with the reflectance (Fassnacht et al., 2016). In our case, there are three reasons to explain why the spectral indices are more efficient than the spectral re-

flectance. With more than 100 indices, the richness of the parameters used in this research is substantial. Secondly, the leaf level spectral measurements are not directly comparable to the target samples which are pixels of the images. This is due to the canopy structure, especially the LAI and the LAD, which modifies the radiation through the vegetation volumes (Roberts et al., 2004). The comparison of the canopy and airborne samples is easier but the conditions of the acquisitions (spatial resolution, solar angle, atmospheric composition, etc.) can cause significant differences among these signals. Therefore, the field spectral measurements are not representative of the image pixels. Thirdly, the reflectance approach is much more sensitive to the Hughes effect (Hughes, 1968) as its dimension is much higher. Potentially discriminant and invariant, the spectral indices are candidates of interest when dealing with ground-based spectral references. These properties are probably also useful for airborne training samples as images acquired in different conditions are impacted similarly. Although the vegetation indices are powerful, the use of other techniques for automatic feature extraction could be investigated.

On the other hand, the ensemble classifier developed in this research is better than the commonly used stacked approach. In a fusion context involving hyperspectral, panchromatic and normalized Digital Surface Model Data (nDSM), it is shown that a decision level fusion (equivalent of the ensemble method) can be more appropriate than a feature level one (equivalent of the stacked approach) for a tree species classification problem (Aval, 2018), which is consistent with our finding. For other classification applications, the ensemble methods have already proven their efficiency in comparison to the stacked approach (Ceamanos et al., 2010). However, the ensemble classifier developed in this study is based on at least one species-specialized classifier, which allow the complementarity of the models subject to the decision rule to be optimized. Three reasons can explain why the ensemble method is the best. First, the ensemble approach is particularly robust because there are several classifiers per species. Secondly, stacking the triplets of indices within the same feature vector leads to features that are discriminant for certain species, but that perform as noise for other ones, which can cause errors. Finally, the Hughes effect, even if it is less significant than for the spectral reflectance (Hughes, 1968), is comparatively more important than in the ensemble method (only three indices per classifier). As an improvement of the proposed approach, other features could be considered as textural or structural ones. Also, instead of having classifiers dedicated to the predictions of specific species, additional classifiers could be designed for separating groups of classes (leafy, coniferous, etc.), in a hierarchical framework.

II.1.5.2 Potential of ground-based spectral references

The ground-based spectral references acquired in this study can be used for the classification of tree species at the airborne level. While the majority of the state-of-art approaches consider training samples directly from the images (Fassnacht et al., 2016), our works demonstrates that encouraging results can be obtained based on leaf level measurements. Canopy level samples lead also to encouraging accuracies. Still focusing on tree species classification, consistent results have been found in the thermal infrared context based on laboratory measurements (da Luz and Crowley, 2010), even if it is difficult to compare our findings. As explained above, the spectral indices are efficient for classifying tree species based on ground measurements because of their discriminant and invariant properties. In addition, the species-specific feature extraction carried out in the proposed method is an essential step in order to identify the best indices. Indeed, the triplets of indices have been chosen for optimizing the target accuracy, i.e. that at the airborne level, through a simple transfer approach. The transfer carried out in this study is however a limit of our approach. Indeed, it requires labelled airborne data, whereas such information is often not available. Transfer learning techniques could be considered for dealing with this issue (Tuia et al., 2016). Also, more training samples could improve the performance of the proposed approach. Radiative Transfer (RT) models such as (Gastellu-Etchegorry et al., 2004; Jacquemoud et al., 2009) could be considered to simulate representative airborne training samples based on leaf level measurements.

II.1.5.3 Link between the vegetation indices and the species

For certain species such as *Aesculus hippocastanum* or *Fagus sylvatica*, the link between the indices and the species is obvious as illustrated in Figure II.6. Focusing on *Aesculus hippocastanum*, the Plant Senescence Reflectance Index (PSRI) (Merzlyak et al., 1999) is particularly discriminant as the two clusters are well separated. The spectral reflectance of that species is consistent with this behaviour as the reflectance in the red region is high in comparison to the other species, which is highlighted by the PSRI. This is because all the trees of this species are affected by the horse-chestnut leaf miner. It necroses its foliage making it characteristic. The leaf miner attacks the parenchyma of the leaf, explaining the decrease of the reflectance in the near infrared (modification of the structure of the leaf). The drying up and leaf necrosis in the Visible increase the reflectance (there are less chlorophyll pigments thus less absorption). Regarding *Fagus sylvatica*, the Green NDVI (Gitelson et al., 1996) is discriminant and has to be related to the much lower reflectance around the green wavelengths in comparison to the other species. Being from *Fagus sylvatica 'Purpurea'* variety, the tree leaves have a significant anthocyanin content, causing a deep absorption in this spectral region. These links encourage the use of expert

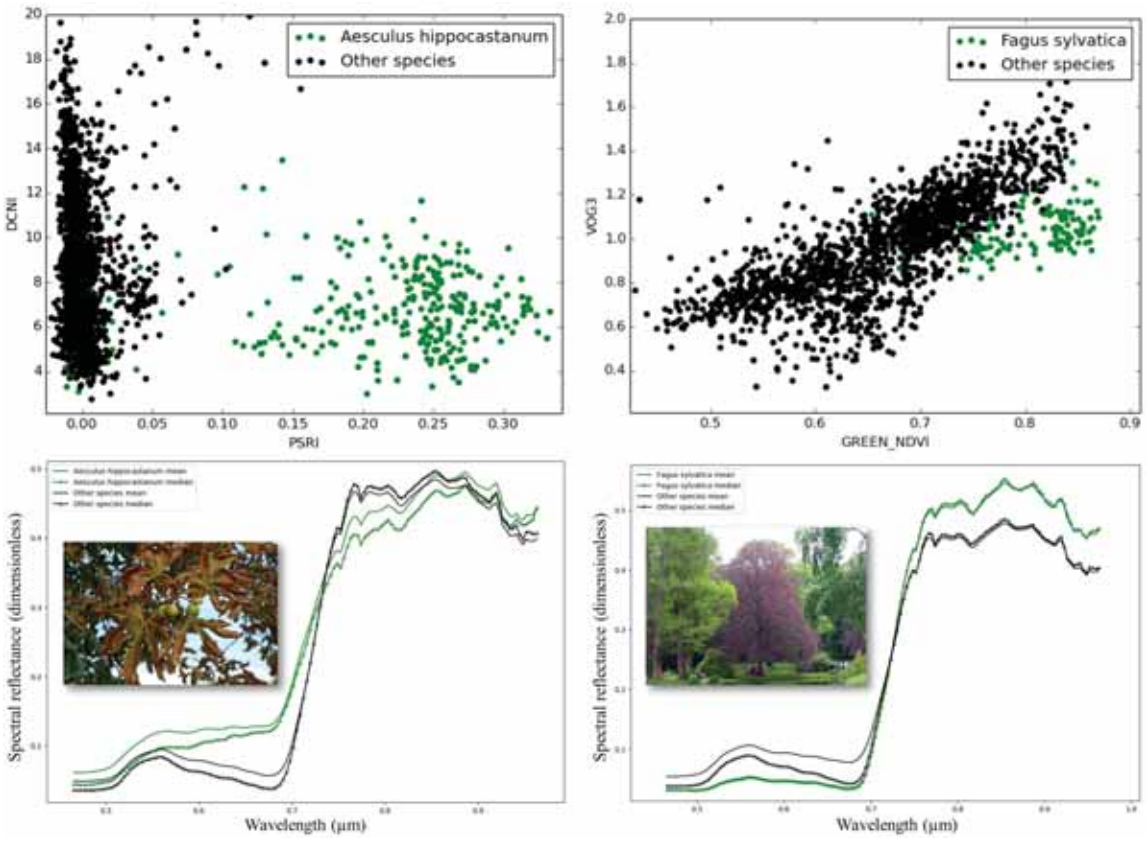

Figure II.6: Link between the spectral indices, the spectral signatures and the species for airborne level data. Left: *Aesculus hippocastanum*. Right: *Fagus sylvatica*. Top: Spectral indices space. Bottom: Spectral signatures space and an illustration of the species.

knowledge about the tree species. However for the other species, such explanation have not been found through our research.

II.1.6 Conclusions

The objective of this study is to classify 5 species in an urban environment, Toulouse, France, by taking advantage of the richness of hyperspectral data through an ensemble classifier approach based on field spectral references. Visible Near-Infrared airborne hyperspectral data are considered for such task. Focusing on the methodological framework, an ensemble classifier method is proposed where at least one classifier is dedicated to

the prediction of a particular species, ensuring the complementarity of the models. Each classifier consists of a Support Vector Machine classification algorithm associated to three vegetation indices. Three levels of training samples are considered: leaf, canopy and airborne levels. Whereas the use of the original spectral reflectance is considered as a first baseline (reflectance method), a feature vector composed of all the spectral indices is chosen as a second reference method (stacked approach).

The main conclusion of this study is that the proposed ensemble method outperforms the classical stacked and reflectance approaches in case of leaf level learning with an OA value of 58% ($\kappa = 46\%$), instead of 45% ($\kappa = 27\%$) and 15% ($\kappa = 2\%$), respectively. Similar behaviours are obtained for the other training levels, but it is more accentuated when leaf level measurements are considered. However, the stacked approach could be used in case of canopy or airborne level training samples. More, the proposed method reduces the misclassifications thanks to the rejected class. From another point of view, this study demonstrates that it is possible to use ground-based spectral references for tree species classification. It is explained by the discriminant and invariant properties of the vegetation indices, in addition to the automatic triplets of indices derivation based on the optimization of the accuracy at the airborne level. However, because the species considered in this study have a high LAI, this result can not be generalized to species with low values of LAI. Indeed, the effect of the soil background would be not negligible.

Further work is necessary, from the improvement of the ensemble method to a more efficient use of the field measurements. The consideration of other classifiers specialized for the discrimination of certain groups of species is of interest for our future research. Also, the consideration of transfer learning techniques seems to be necessary in order to improve the performance. For instance, a radiative transfer model could be used for simulating canopy level spectra based on leaf level ones, in order to get more training samples. Finally, the combination of radiative transfer models and deep learning approaches is a way of investigation in order to get automatically the most powerful features, and at the same time understanding what happens through the machine learning framework.

Acknowledgements

The authors would like to thank the French Aerospace Lab (ONERA) and the Région Occitanie for funding this research. Concerning the field data collection, thanks to the city of Toulouse and Jérôme Willm from the National Institute of Agricultural Research. Regarding airborne data, thanks to Philippe Deliot from ONERA and thanks to the French Mapping Agency (IGN) for their participation in the airborne campaign. Thanks to the

members of the Hyperspectral imagery for Environmental urban planning (HYEP) project, funded by the French National Research Agency (ANR). Thanks to Aurélie Michel and Karine Adeline for discussions about classification frameworks.

Author contributions

Josselin Aval and and co-authors conceived the study and wrote the paper; Josselin Aval performed field data collection, data preprocessing and implemented the data analysis.

Conflict of interests

The authors declare no conflict of interest.

Chapter III

Urban tree species classification from spectral and contextual features

Synthèse de l'article en français

Le but de cet article est de détecter les arbres qui appartiennent à un alignement de manière individuelle.

Jeu de données

Afin d'évaluer la robustesse de la méthode proposée pour différentes conditions, trois zones d'étude sont considérées. Chaque zone a une complexité particulière, détaillée dans Table III.1, avec une complexité croissante. En effet, la principale difficulté se manifeste pour les cas de connexion spatiale importante entre les arbres d'alignement et les autres populations d'arbres, une difficulté croissante du premier au troisième site.

Comme les arbres de rue peuvent être vus comme une végétation haute à proximité du réseau routier, des données hyperspectrales sont choisies pour la détection de la végétation, alors qu'un DSM est utilisé pour filtrer les objets hauts. Des données hyperspectrales étaient disponibles pour ce travail de doctorat, ce qui explique pourquoi ce type de données est utilisé pour cette tâche, mais des données multispectrales auraient pu être utilisées. Enfin, les données SIG sont utilisées pour détecter les pixels proches des routes. Les principales caractéristiques du jeu de données sont décrites dans le tableau III.2, après que les prétraitements géométrique et radiométrique aient été effectués.

TABLE III.1 : Principales caractéristiques des zones d'étude. "Contexte dans la ville" se réfère soit aux arbres d'alignement du centre de Toulouse, reliés spatialement à ceux des parc ("centre"), soit aux arbres d'alignement à proximité de propriétés privées, donc liés aux arbres privés ("privé"). "Taille" indique la forme de la délinéation des arbres d'alignement, résultant de la taille par les gestionnaires du milieu urbain. Le terme "Chevauchement" fait référence au niveau de chevauchement entre les arbres d'alignement. La "Connexion spatiale" met en évidence le niveau de connexion spatiale entre les arbres d'alignement et les autres populations d'arbres.

	Cas 1	**Cas 2**	**Cas 3**
Contexte dans la ville	centre	centre	privé
Genre	*Tilia*	*Platanus*	*Platanus*
Nombre d'arbres d'alignement	∼50	∼100	∼70
Nombre d'arbres total	∼100	∼400	∼700
Taille	rectangle	circulaire	circulaire
Chevauchement	Non	Intermediaire	Non
Connexion spatiale	Faible	Faible	Importante

TABLE III.2 : Principales caractéristiques du jeu de données. Le DSM est obtenu à partir d'acquisitions stéréoscopiques du système CAMV2. "N" représente le nombre de bandes spectrales. L'attribut principal du réseau routier est le type de route, c'est-à-dire pour chaque route s'il s'agit d'une voie primaire ou piétonne, une information utile pour détecter les arbres qui sont plantés le long des routes à circulation automobile.

	VNIR	**DSM**	**GIS**
Source	HySpex VNIR-1600	CAMV2	OSM
Type	raster	raster	vecteur
	Quantité		**Type de vecteur**
	réflectance spectrale	hauteur	ligne
	GSD		**Attribut principal**
	0.4 m	0.125 m	Type de route
Intervalle	0.4 - 1 μm		
N	160		

Méthode proposée

La méthode proposée comporte deux étapes principales : la détection de la végétation haute suivie de la délimitation individuelle des couronnes d'arbre d'alignement. En ce qui concerne la détection de la végétation haute, elle est le résultat de quatre masques dont l'intersection géométrique est calculée pour obtenir le masque final. En particulier, le NDVI est utilisé pour détecter la végétation tandis qu'un autre indice de la littérature est calculé pour supprimer les ombres. Enfin, la hauteur et la distance aux routes sont calculées pour détecter les objets hauts et les objets proches des routes, respectivement.

Le masque résultant est ensuite utilisé comme entrée de la délinéation des couronnes d'arbres d'alignement, à travers une approche MPP. Dans ce contexte, l'idée du MPP est de considérer que la carte des arbres d'alignement est une réalisation spécifique d'un processus ponctuel marqué $\mathbf{x}$, dans un espace d'état χ (espace de position associé à l'espace des marques). Chaque couronne (ou marque) d'arbre d'alignement est modélisée comme un cercle. Trouver la meilleure réalisation devient une minimisation d'énergie incluant deux termes d'énergie, l'énergie d'attache aux données $U_d(\mathbf{x})$, la modélisation d'arbres d'alignement individuels (niveau de l'arbre) et l'énergie d'interaction $U_i(\mathbf{x})$. Cette dernière modélise les caractéristiques contextuelles discriminantes des arbres d'alignement (niveau de l'alignement, hypothèses d'angle petit entre les arbres et hauteurs similaires).

La méthode proposée est comparée à deux méthodes de référence, la méthode de délinéation utilisée jusqu'ici (Adeline, 2014), sans et avec la prise en compte du réseau de routes. Cela permet de mettre en évidence la contribution de notre terme d'interaction $U_i(\mathbf{x})$. Le même masque que celui mentionné ci-dessus est utilisé. Une matrice de confusion est calculée afin d'évaluer la performance de ces méthodes. À partir de cette matrice de confusion, la précision du producteur (PA), la précision de l'utilisateur (UA) et le F-score sont utilisés pour la comparaison.

Résultats

La comparaison entre les approches de référence et la méthode proposée est présentée dans la Table III.3. La méthode proposée surpasse l'approche de référence ignorant le réseau routier avec des valeurs de F-score de 91%, 75% et 85% au lieu de 70%, 41% et 20% pour les trois cas, respectivement. Quand la méthode de référence utilise le réseau de routes, l'approche proposée permet d'améliorer la performance de 8pp et de 57pp dans les premier et troisième cas, respectivement. Cependant, une baisse de 7pp est observée pour le second cas. En conclusion, la méthode proposée est la plus appropriée pour la détection des arbres en alignement urbain pour les premier et troisième cas (pas de chevauchement mais

TABLE III.3 : Comparaison des méthodes de référence et proposée en termes de F-score (%). Les scores en gras associés à l'étoile (*) renvoient au score maximum.

	Méthode	Cas 1	Cas 2	Cas 3
Référence	**Sans GIS**	70	41	20
	Avec GIS	83	**82***	28
	Proposée	**91***	75	**85***

TABLE III.4 : Synthèse de la contribution de l'angle (θ), de la hauteur (h) et des données SIG en termes de score F (%). Les scores en caractères gras associés à l'étoile (*) renvoient au score maximum.

Combinaison	**Cas 1**	**Cas 2**	**Cas 3**
Avec GIS, sans θ et h	84	**79***	39
Avec θ et h	88	**79***	62
Avec GIS, θ et h	**91***	75	**85***

potentiellement une connexion spatiale importante), alors que la méthode de référence associée aux données SIG est meilleure pour la deuxième étude zone (c'est-à-dire en cas de chevauchement sans connexion spatiale significative).

La contribution de l'angle, de la hauteur et des données SIG dans la cartographie des arbres d'alignement est étudiée (avec ou sans chacune de ces caractéristiques) et résumée dans la Table III.4. Dans l'ensemble, les meilleures performances sont obtenues lorsque toutes les caractéristiques sont utilisées avec des valeurs de F-score de 91%, 75% et 85% pour les trois cas, par rapport à 76%, 58% et 26% sans aucune caractéristique. En particulier, l'utilisation des données SIG seule n'est pas suffisante comme le souligne le F-score pour le troisième cas (cas de connexion spatiale significative), égal à 39%. Cependant, utiliser seulement θ et h donne déjà de bons résultats avec des valeurs de F-score de 88%, 79% et 62%. En conclusion, les trois caractéristiques réunies constituent le meilleur ensemble de caractéristiques contextuelles discriminantes et doivent être utilisées pour cette tâche, en particulier dans le troisième cas (connexion spatiale significative). Mais pour les premier et deuxième cas (pas de connexion spatiale significative), la meilleure performance peut être atteinte sans données SIG.

Discussions

Tout d'abord, ces résultats démontrent que le cadre proposé permet de détecter les arbres d'alignement, individuellement, alors qu'aucune méthode n'a été dédiée à cette tâche jusqu'à présent. Comme élément de comparaison, le travail de (Wen et al., 2017) peut être cité. Ils utilisent une approche de détection des arbres d'alignement à l'échelle d'un amas d'arbres (pas de détection individuelle des arbres) et obtiennent une valeur de F-score de 89%, ce qui est cohérent avec notre approche bien que difficilement comparable. Mais comme indiqué plus haut, la méthode proposée cartographie les arbres individuellement, une tâche essentielle pour un suivi individuel de la santé des arbres d'alignement. En effet, la prévention de la chute des arbres malades ne peut être faite avec précision sans une connaissance individuelle du niveau de l'arbre (Fini et al., 2015).

Les résultats mettent également en évidence que la méthode proposée n'est pas la meilleure en cas de chevauchement, notamment parce qu'elle repose uniquement sur le masque de végétation haute. Ce problème bien connu de la littérature sur les méthodes de délimitation des couronnes d'arbres individuelles (Zhen et al., 2016) pourrait être traité en ajoutant une modélisation de la structure 3D des arbres dans l'énergie d'attache aux données, similaire à ce qui est fait dans l'approche de référence. D'autres modèles 3D plus précis pourraient bénéficier à la définition de l'énergie d'attache données dans l'approche MPP, notamment via l'utilisation de données LiDAR (Leckie et al., 2003; Zhen et al., 2015).

Comme mentionné précédemment, savoir si un arbre appartient à un alignement est intéressant pour les gestionnaires du milieu urbain, afin d'organiser un suivi spécifique, mais aussi afin d'améliorer les méthodes existantes, que ce soit pour les approches de délimitation des arbres ou pour la classification des espèces. En particulier, l'approche de classification des espèces fondée sur le VNIR testée dans le premier chapitre, est appliquée sur le troisième cas d'étude, donnant 74% des arbres *Platanus x hispanica* correctement identifiés, avant que la carte des espèces obtenue soit régularisée (100% des prédictions correctes) en utilisant l'appartenance à l'alignement, dérivée de la méthode proposée.

Conclusions

L'objectif est de cartographier les arbres d'alignement à partir de données aéroportées et d'informations contextuelles basées dans le cadre des MPP. Trois sites test sont considérés. Des données hyperspectrales aéroportées, des données DSM et GIS sont utilisées. Sur la base de ces données, les canopées proches des rues sont détectés grâce aux seuils de NDVI, d'indice d'ombre, de hauteur et de distance aux routes. Le masque de végétation haute obtenu constitue l'entrée d'une approche MPP. En particulier, les caractéristiques

contextuelles discriminantes des arbres d'alignement sont modélisées. Une approche de délimitation standard est considérée comme une référence.

En ce qui concerne les résultats, l'approche proposée surpasse la méthode de référence ignorant les données SIG avec une valeur F-score moyenne de 84% dans les trois zones d'étude, au lieu de 44%. Quand la méthode de référence exploite le réseau de routes, l'amélioration est moins significative (+20pp), la différence étant principalement pour le troisième cas (connexion spatiale significative). Pour une telle condition, la méthode proposée est la plus appropriée, mais elle fait quelques erreurs en cas de chevauchement important. Enfin, toutes les caractéristiques doivent être utilisées ensemble pour atteindre les meilleures performances, mais les données SIG ne sont pas nécessaires dans les cas où la connexion spatiale est faible.

D'autres travaux sont nécessaires. Nous voyons que la méthode proposée peut être améliorée en cas de chevauchement important. En outre, d'autres caractéristiques pourraient être utilisées pour modéliser les arbres de rue. Aussi, la méthode devrait être appliquée sur des cas plus difficiles, comme des zone avec des arbres d'alignement formant une canopée homogène. Dans ces cas, il est souvent impossible de distinguer les couronnes avec des données spectrales et un DSM, ce qui encourage l'utilisation d'une autre technologie. Par exemple, des données multitemporelles acquises durant l'hiver pourraient aider. L'espèce et l'état de santé des arbres seront également intéressantes à caractériser.

English part: Third article

The third paper is included in the next section[1].

III.1 Detection of individual trees in urban alignment from airborne data and contextual information: a marked point process approach

Abstract: With the current expansion of cities, urban trees have an important role for preserving the health of its inhabitants. With their evapotranspiration, they reduce the urban heat island phenomenon, by trapping CO_2 emission, improve air quality. In particular, street trees or alignment trees, create shade on the road network, are structuring elements of the cities and decorate the roads. Street trees are also subject to specific conditions as they have little space for growth, are pruned and can be affected by the spread of diseases in single-species plantations. Thus, their detection, identification and monitoring are necessary. In this study, an approach is proposed for mapping these trees that are characteristic of the urban environment. Three areas of the city of Toulouse in the south of France are studied. Airborne hyperspectral data and a Digital Surface Model (DSM) for high vegetation detection are used. Then, contextual information is used to identify the street trees. Indeed, Geographic Information System (GIS) data are considered to detect the vegetation canopies close to the streets. Afterwards, individual street tree crown delineation is carried out by modelling the discriminative contextual features of individual street trees (hypotheses of small angle between the trees and similar heights) based on Marked Point Process (MPP). Compared to a baseline individual tree crown delineation method based on region growing, our method logically provides the best results with F-score values of 91%, 75% and 85% against 70%, 41% and 20% for the three studied areas respectively. Our approach mainly succeeds in identifying the street trees. In addition, the contribution of the angle, the height and the GIS data in the street tree mapping has been studied. The results encourage the use of the angle, the height and the GIS data together. However, with only the angle and the height, the results are similar to those obtained with the inclusion of the GIS data for the first and the second study cases with F-score values of 88%, 79% and 62% against 91%, 75% and 85% for the three study cases respectively. Finally, it is shown that the GIS data only is not sufficient.

[1] **J. Aval**, J. Demuynck, E. Zenou, S. Fabre, D. Sheeren, M. Fauvel, K. Adeline and X. Briottet. Detection of individual trees in urban alignment from airborne data and contextual information: a marked point process approach. ISPRS Journal of Photogrammetry and Remote Sensing, 2018.

Keywords Street tree; Urban remote sensing; Airborne data; Geographic Information System; Individual tree crown delineation; Marked Point Process.

III.1.1 Introduction

The world urban population will increase to nearly 5 billions by 2030, and at the same time the urban land cover will increase by 1.2 millions km^2 (Seto et al., 2012). With this expansion of urban areas, urban canopies have an important role to play as they improve air quality (Yang et al., 2005), reduce heat islands (Doick et al., 2014), promote biodiversity and have a relaxing psychic action (Chiesura, 2004). Urban tree structures including street trees and park ones do not have necessarily the same functions / roles in the urban context (Bolund and Hunhammar, 1999). In addition to the properties mentioned above (Vailshery et al., 2013; Gillner et al., 2015), the street trees create shade, are structuring elements of the cities and decorate the roads (McPherson et al., 2016). They are also subject to specific conditions as they have little space for growth and are pruned, most often to be adapted to the constraints of the sites, and can be affected by the spread of diseases in single-species plantations (Sebestyen et al., 2008). As a case in point, a pruned lime tree (*Tilia*) has a life expectancy of 150 years against 800 years without constraint (Baraton, 2014; Fini et al., 2015). In order to highlight the crucial place of the street trees in the urban environment, the example of Paris, France can be cited with nearly 100,000 street trees (about half of the trees). These street trees cover around 700 km of roads and concern approximately 1600 roads out of 6000. Especially, the shadow produced by the street trees represents 3% of the area of Paris (Rol-Tanguy et al., 2010). The managers of the urban environment have to consider the distinctive characteristics of the street trees for a specific urban planning and a specific monitoring, and a first step is the individual street tree identification. Nowadays, this type of procedure is carried out manually, by field campaign or by photointerpretation (Pulighe and Lupia, 2016), and does not allow to cover large scales of continuous urban area with regular time basis.

Remote sensing opens the way to automate the individual street tree mapping. Indeed, airborne remote sensing sensors can cover entire cities with a spatial resolution of an order of magnitude of 1 m and with regular time basis (Alonzo et al., 2014). Airborne multispectral and hyperspectral sensors measure the spectral radiance and thus allow the vegetation to be detected (Xiao et al., 2004). Active sensors as Light Detection And Ranging (LiDAR) or passive sensors in stereoscopic configurations can be used to measure the height and makes it possible to characterize the vertical structure of the objects (MacFaden et al., 2012). With the association of these remote sensing technologies, the urban canopy considered as high vegetation can then be mapped (Ramdani, 2013). On the other hand,

Geographic Information System (GIS) data, especially vector data, constitute an important source of information and are often available at a city scale from the urban managers, but also more and more on a global scale from open databases such as OpenStreetMap (OSM). In a perspective of street tree identification, such vector data are of interest because an information such as the road network is often available and would allow the canopies close to the streets to be detected (Wen et al., 2017).

From these remote sensing data, individual tree mapping, conventionally termed Individual Tree Crown Delineation and Detection (ITCD), has been addressed for many years and several ITCD methods have been proposed (Zhen et al., 2016). Raster-based methods such as valley-following (Leckie et al., 2003), region-growing (Adeline, 2014), watershed segmentation (Chen et al., 2006) and template matching (Gomesa and Maillarda, 2014) have been developed. Point cloud-based and tree shape reconstruction approaches like K-means clustering technique (Gupta et al., 2010) and Hough transform (Van Leeuwen et al., 2010) have been explored respectively. Finally, there are methods combining raster, point, and a priori information such as Markov random fields (Ardila et al., 2011), Marked Point Process (MPP) approaches (Perrin et al., 2006) which can use a prior contextual information on the trees (Van Lieshout, 2000). Even if these methods have exhibited good performance in the literature, complicated urban and non-urban forests are still challenging (mainly in case of important overlaps) (Zhen et al., 2016). Focusing on the urban environment, of the 207 studies identified in the recent review of (Zhen et al., 2016) on the ITCD methods, only 18 have been applied in urban areas. The objective of these studies was to map the urban trees individually, and no distinction is made between the different structures of the trees in the urban context such as street trees and park trees.

However, these structures are of interest for the urban managers for a specific urban planning and a specific monitoring, with the example of the street trees highlighted previously. In addition, this information could be used in order to improve not only the individual tree mapping itself (by taking advantage of a prior contextual information knowledge about the urban trees depending on their structure), but also the tree species classification for example (by defining specific categories of urban trees depending on their structure because street trees have not necessarily the same spectral traits than park ones). To our knowledge, this consideration of the tree structures in the urban canopy mapping is the subject of only one study, (Wen et al., 2017) where an approach for classifying the urban canopies (patch-level classification) in three classes (park, roadside and residential-institutional canopies) has been proposed. GIS data and specific spectral, textural, shape and contextual features (such as the proximity to the road) are considered in order to characterize these classes. Shenzhen and Wuhan (China) constitute the study sites and the

method is based on WorldView-2 satellite imagery (spatial resolution of 2 m for the multi-spectral mode). F-score values of 76%, 89% and 87% are obtained for park, roadside and residential canopies respectively. In such a patch-level framework, there are confusions between the street trees and the other populations of trees because of the spatial connections between the canopies, which could be probably better handled with an individual detection approach.

Summarizing the existing literature, there is minimal consideration of the specific tree structures in the urban environment such as street trees and park trees. In particular, no individual tree mapping which takes into account the structure of the alignment trees has been proposed. To alleviate this issue, the aim of this paper is to map the trees which belong to an alignment individually. Airborne data and contextual information are used in an approach based on MPP, which allows a prior information to be modelled. For that purpose, the following issues are addressed:

1. What are the discriminative contextual features of the street trees?
2. How to model these features for individual street tree mapping?
3. Which features contribute the most in individual street tree mapping?

The paper is organized as follows. Section III.1.2 presents the study area and the data used for individual street tree mapping, followed by section III.1.3 with the description of the proposed method and a baseline ITCD method used for comparison. Afterwards, the results are showed in section III.1.4 and discussed in section III.1.5. Finally, main conclusions of the study are detailed and the perspectives of the work highlighted in the section III.1.6.

III.1.2 Materials

III.1.2.1 Study area

The study is carried out in Toulouse city located in the South West of France (43.6 °N, 1.44 °E). With about 500,000 inhabitants, Toulouse is the fourth city in France. The climate of Toulouse is temperate with oceanic, Mediterranean and continental characteristics. Concerning the urban vegetation, Toulouse would have approximately 140,000 trees with at least 20,000 street trees according to urban managers. Three areas in Toulouse downtown are selected in this study (figure III.1).

The three study cases are presented in figure III.2. In all cases, the street trees form lines

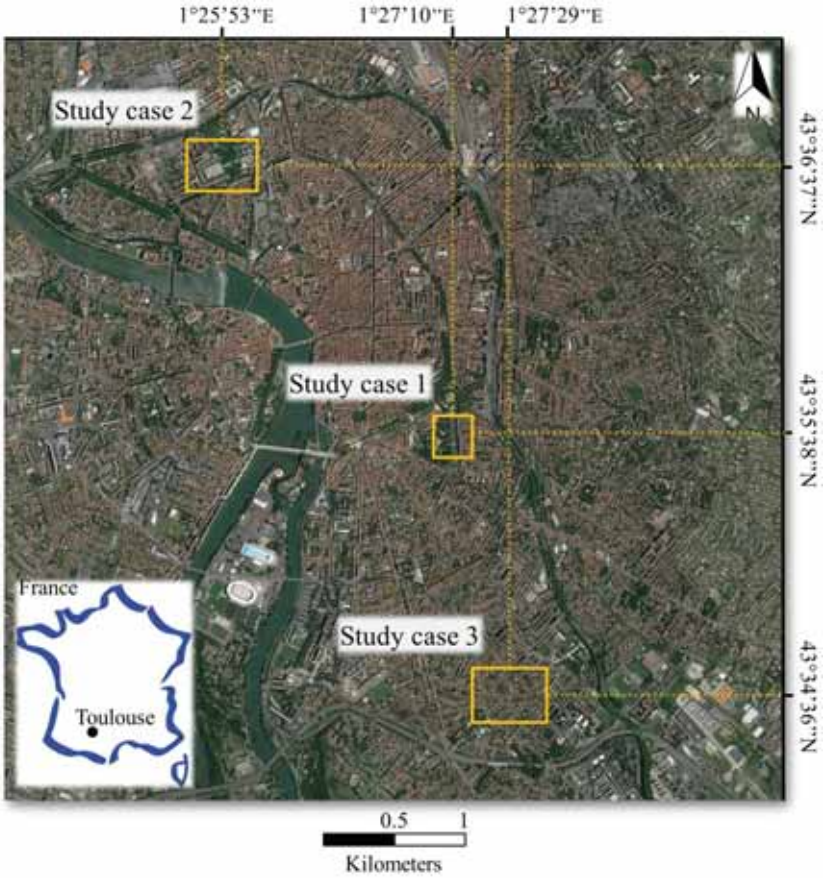

Figure III.1: Study area with the three study cases represented on a Google Earth image. The yellow rectangles correspond to each study case.

Figure III.2: Description of the study cases. At the top, field view is showed for each case. At the bottom, Google Earth images illustrate each case.

along roads and are pruned as it will be highlighted in section III.1.3.1. This results in small angle between the street tree trunks and similar tree heights. The first study site is located in the center of Toulouse and includes street and park trees. The street trees do not overlap and are silver linden trees (*Tilia tomentosa*). The second site is also located in the center of Toulouse and also includes street and park trees. However, this case is more challenging than the first one because the number of trees is higher and the street trees overlap more and are organized in two adjacent lines. In this case, the street trees are plane trees (*Platanus x hispanica*). The third site is situated in a quarter of private properties. For this site, the number of trees is high with a complex spatial organization because of the presence of many garden trees. The majority of the trees are not aligned, and spread over a great extent. The street trees do not overlap and are plane trees (*Platanus x hispanica*).

III.1.2.2 Airborne and GIS data

Airborne data were acquired on October 24, 2012 at 11:00 UT (Universal Time) during the UMBRA campaign (Adeline et al., 2013) organized by the French Aerospace Lab (ONERA) and the French Mapping Agency (IGN). The flight height was approximately

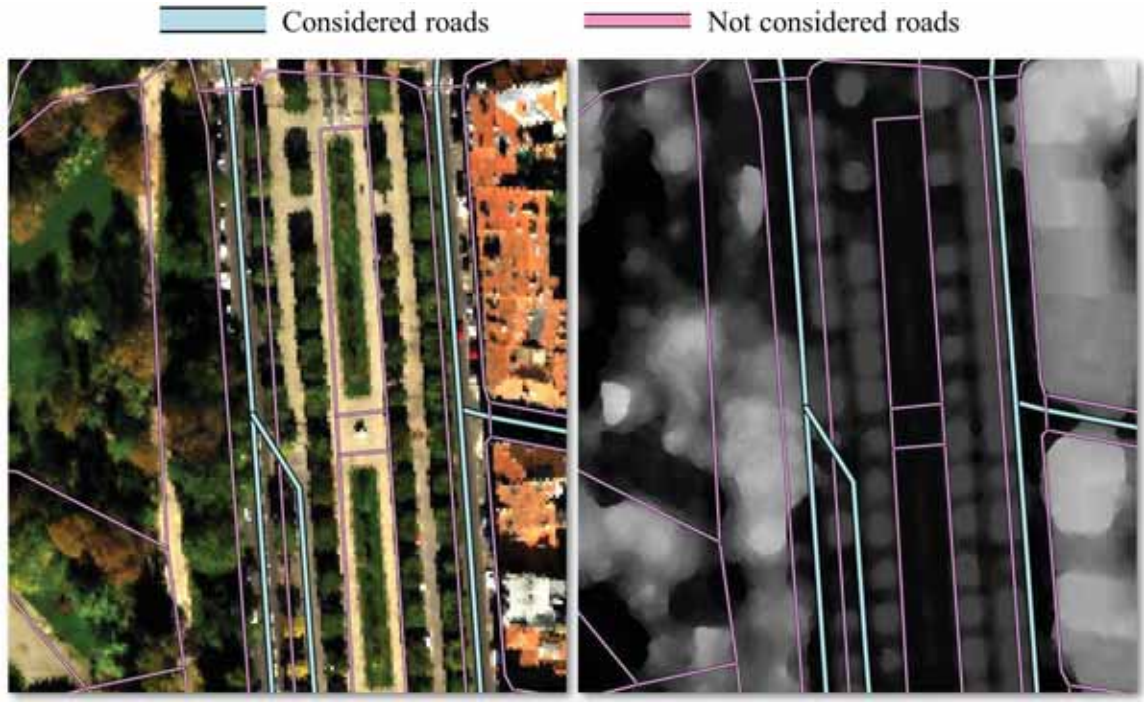

Figure III.3: Illustration of the data used in this study for the first site. Left: VNIR reflectance image. Right: DSM. The GIS data is represented by blue and violet lines indicated over the airborne data.

2,000 m over the study area. The HySpex Visible Near-Infrared (VNIR) system (Köhler, 2016) was used and consists of an hyperspectral push broom camera with 160 spectral bands (0.4 µm - 1 µm). About the spatial resolution, the VNIR camera data are acquired with a spatial resolution of 0.4 m and 0.8 m across and along tracks, respectively. In order to build a DSM, the French Mapping Agency CAMv2 system was used (Souchon et al., 2010) for performing stereoscopic acquisitions with an overlap of 80%. A vector layer of roads derived from the OSM database is used and identified as "GIS data" in the next sections. In OSM, each road of the road network is characterized by the attribute *type* which describes the type of road (motorway, primary, path, etc.). Only the primary, secondary, tertiary, residential and service roads are considered because we assume that the street trees are only planted along roads with motor vehicle traffic (figure III.3).

III.1.2.3 Preprocessing

Geometric and radiometric preprocessing is carried out. From the stereoscopic measurements, the French Mapping Agency provides us a georeferenced DSM with a spatial resolution of 12.5 cm. Then, the VNIR image is registered on the DSM by defining Ground Control Points (GCP) on QGIS and by using the function *gdalwarp* from GDAL. Nearest neighbor resampling is applied in order to preserve the original spectral data. Also, the Thin Plate Spline (TPS) transformation (Duchon, 1977) is applied for its ability to correct the deformations locally. Because the VNIR pixels have rectangular shapes with the longer side along track, a square grid with a spatial resolution of 0.4 m (minimum between the rectangle sides) is chosen to preserve the original data. Visual assessment suggests that the error is less than a pixel for the whole data set. Furthermore, the hyperspectral data are atmospherically corrected to deal with spectral reflectances with COCHISE (Poutier et al., 2002) based on MODTRAN and assuming a flat scene. The DSM is resampled to the VNIR image resolution (0.4 m x 0.4 m) with the nearest neighbor resampling. In order to get a normalized DSM (nDSM) and assuming a flat ground, the ground altitude is estimated as the altitude corresponding to the maximum of the DSM histogram and we make the difference between the DSM and the estimated altitude (Adeline, 2014). The size of the bins of the histogram is 1 m.

III.1.3 Methods

The description of the proposed method is carried out in section III.1.3.1, followed by section III.1.3.2 with the description of a baseline ITCD method used for comparison.

III.1.3.1 Proposed street tree mapping

The figure III.4 presents the proposed street tree mapping scheme. First, the high vegetation close to the streets is detected (section III.1.3.1). Secondly, the street tree crowns are delineated based on MPP which allow a prior contextual information to be modelled via an interaction term (section III.1.3.1). In this paper, we assume that the street trees can be characterized by the following discriminative contextual features:

- A street tree is close to a road.
- A street tree is aligned with its neighbors.
- A street tree is the same height as its neighbors.

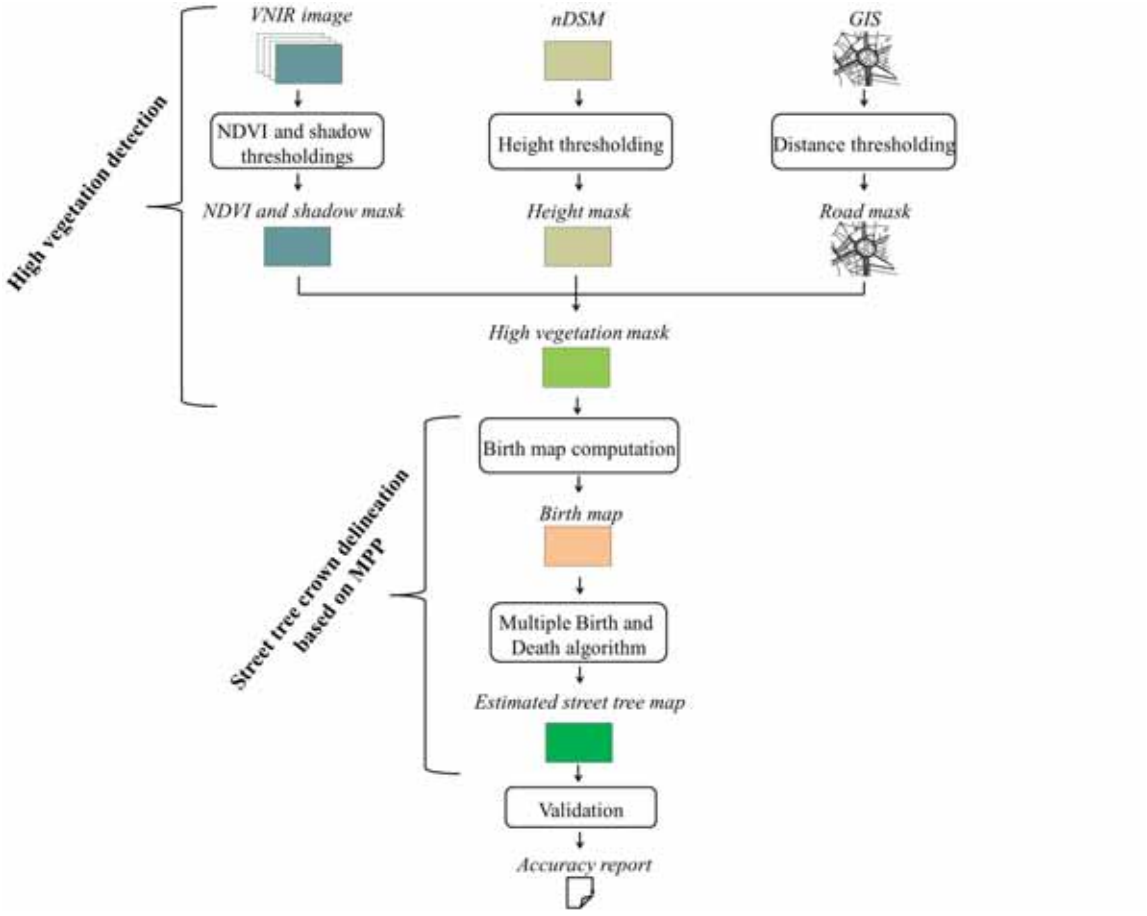

Figure III.4: Graphic representation of the street tree mapping method.

High vegetation detection Four masks (vegetation, shadow, height and optionally distance) are combined (geometric intersection, i.e. logical operator AND) in order to generate a high vegetation mask. This last one is then used for computing the data energy $U_d(\mathbf{x})$ of the MPP, defined in section III.1.3.1 (figure III.4). The mask is the MPP input data that allows to reduce the computational time, by restricting the search space P (introduced in section III.1.3.1). For the vegetation mask, the NDVI (Normalized Difference Vegetation Index) index (Rouse Jr et al., 1974) is computed for each pixel of the VNIR image from a red and an infrared hyperspectral bands. Above a threshold determined automatically with the Otsu method (Otsu, 1975), the pixels are kept. About the shadow mask, the spectral reflectance cannot be retrieved in shadows, as the atmospheric correction method is based on a flat scene hypothesis (figure III.3 near trees and buildings). To avoid errors from these shadow regions, a literature index defined as: $I = 1/6 * (2 * R + G + B + 2 * NIR)$ (Nagao et al., 1979) is used for its efficiency and simplicity in the same way as the NDVI. Regarding the height mask, all the pixels with a nDSM value higher than 5 m are filtered (the minimum height value of the alignment trees in Toulouse according to urban managers). Finally, a distance to the roads mask is optionally used to assess the contribution of the GIS data to the street tree mapping. Below a distance threshold of 20 m, the pixels are retained. Indeed, from the width of the roads and the rules for planting the trees along the roads in France, street tree trunks can be at more than 10 m from the middle of the roads. A margin of 10 m is taken in order to consider all the pixels of the crowns. For building the high vegetation mask, the NDVI and the shadow index could be computed from multispectral data. For next sections, the use of the GIS data is referred as "with GIS", "without GIS" otherwise (with or without road network information).

Street tree crown delineation based on MPP The street tree map can be viewed as a space where positions and attributes of street trees are a specific realization of a marked point process noted $\mathbf{x}$ (Van Lieshout, 2000). The proposed method assumes that the street tree crowns can be represented as disks. In this context, a state space χ in which $\mathbf{x}$ is a realization can be defined such as:

$$\chi = P \times M = [1, X_M] \times [1, Y_M] \times [r_m, r_M] \tag{III.1}$$

where P and M correspond to the position space and the space of the marks, respectively. Regarding the positions, X_M and Y_M are the column and line numbers of the VNIR image. About the marks, r_m and r_M are the minimum and maximum radius of the disks (2 m and 8 m respectively because the street trees are pruned and have their radius included in this range according to urban managers). To find the realization of $\mathbf{x}$ which corresponds to the street tree map, the issue becomes an energy minimization including two energy terms called the data energy $U_d(\mathbf{x})$ and the interaction energy $U_i(\mathbf{x})$ (Perrin et al., 2006). In the

context of the street tree mapping, the data term models individual street trees (tree level) while the interaction term models the discriminative contextual features of the street trees (alignment level with hypotheses of small angle between the trees and similar heights).

The data energy $U_d(\mathbf{x})$ is the sum of the individual data energies $U_d(x_i)$ of each street tree x_i (x_i is defined by its position and radius). The computation of $U_d(x_i)$ is taken from (Zhou et al., 2010). Instead of computing a grey level radiometric distance between the pixels in the disk and the pixels in the concentric annulus around the disk (i.e. outside the possible crown) corresponding to x_i, a simple difference between the proportion of high vegetation pixels (from the high vegetation mask) in the disk and the proportion of high vegetation pixels in the concentric annulus is computed. We consider that x_i corresponds to a street tree if this distance exceeds a certain threshold d_0 (0.2 fixed after testing multiple values between 0 and 1) (equation III.2).

$$\begin{aligned} d(x_i) = \frac{N_{x_i}}{N_{t,x_i}} - \frac{N_{a_{x_i}}}{N_{t,a_{x_i}}} \quad if \quad \frac{N_{x_i}}{N_{t,x_i}} - \frac{N_{a_{x_i}}}{N_{t,a_{x_i}}} > 0 \quad else \quad d(x_i) = 0 \\ then \quad U_d(x_i) = 1 - \frac{d(x_i)}{d_0} \quad if \quad d(x_i) < d_0 \\ and \quad U_d(x_i) = \exp[-(d(x_i) - d_0)] - 1 \quad if \quad d(x_i) \geq d_0 \end{aligned} \tag{III.2}$$

with N_{x_i} and N_{t,x_i} the number of high vegetation pixels and the total number of pixels in the disk (i.e. $\frac{N_{x_i}}{N_{t,x_i}}$ is a proportion of high vegetation pixels in the disk). Similarly, $N_{a_{x_i}}$ and $N_{t,a_{x_i}}$ are the number of high vegetation pixels and the total number of pixels in the concentric annulus whose radius is fixed to 1 m in order to include pixels all along the annulus.

The interaction energy $U_i(\mathbf{x})$ is the sum of an energy $U_{i_s}(\mathbf{x})$ that ensures the stability of the process and the street tree feature energy $U_{i_f}(\mathbf{x})$ that models the features of the street trees. As in (Perrin et al., 2005), the energy $U_{i_s}(x_i)$ for a street tree x_i is defined according to the intersected areas between the street tree crowns and avoids an excessive overlap of the trees (for example trees located almost in the same place) (equation III.3).

$$U_{i_s}(x_i) = \sum_{j \neq i} \frac{A_{x_i} \cap A_{x_j}}{\min(A_{x_i}, A_{x_j})} \tag{III.3}$$

where A_{x_i} and A_{x_j} refer to the areas of x_i and x_j. Concerning the street tree feature energy $U_{i_f}(\mathbf{x})$, it is defined by considering the features of the street trees in the urban environment illustrated in figure III.5. Whereas the not street trees have no particular spatial organization and different heights, the street trees form lines and are pruned in the same way, most

often to adapt the trees to the constraints of the sites. This results in a small angle between the trees and similar shapes (here we only consider the height to model the shape). In order to model these features, we define the street tree feature energy $U_{i_f}(x_i)$ for a street tree x_i based on its features and the features of two of its neighbouring street trees, i.e. an alignment is modelled from three street trees (figure III.5 and equation III.4). We choose three trees because such a model is more flexible in case of curved roads for example.

$$\begin{aligned} U_{i_f}[x_i,(x_j,x_k)] &= \frac{1}{a+b}\cdot(a\cdot\frac{\theta_{ijk}}{\pi/2}+b\cdot\frac{|h_i-h_j|+|h_i-h_k|}{2\cdot\max(h_i,h_j,h_k)})\in[0,1] \\ then \quad U_{i_f}(x_i) &= \min_{(j,k),i\neq j\neq k} U_{i_f}[x_i,(x_j,x_k)] \quad for \quad (x_j,x_k) \quad in \quad V_{x_i} \\ & \quad if \quad V_{x_i}\neq\emptyset \quad else \quad U_{i_f}(x_i)=1 \\ where \quad \theta_{ijk} &= \pi-\varphi_{ijk} \quad if \quad \pi/2<\varphi_{ijk}\leq\pi \\ and \quad \theta_{ijk} &= \varphi_{ijk} \quad if \quad 0\leq\varphi_{ijk}\leq\pi/2 \end{aligned} \tag{III.4}$$

and the boolean a and b are used to study the contribution of the angle between the trees (first term) and the heights (second term) in the street tree mapping. For next sections, $a=1$ is referred as "with θ" whereas $a=0$ is referred as "without θ" (with or without angle information). Similarly, $b=1$ is referred as "with h" whereas $b=0$ is referred as "without h" (with or without height information). φ_{ijk} is the angle between the two segments joining the center of x_i with the centres of x_j and x_k. In particular, $U_{i_f}[x_i,(x_j,x_k)]$ is computed for all pairs of neighbours x_j and x_k in the neighbourhood V_{x_i} of x_i (radius of 25 m because the distance between two trees in an alignment does not exceed 10 m according to urban managers and the first and last trees have to be considered). The minimum $U_{i_f}[x_i,(x_j,x_k)]$ is then retained as $U_{i_f}(x_i)$. If there is no pair of neighbours, $U_{i_f}(x_i)$ is equal to 1 (penalized), because an alignment tree is never isolated. When there is a small angle between the street tree x_i and its neighbouring street trees of similar heights, $U_{i_f}(x_i)$ is close to 0. Other configurations are penalized as they result in higher values of $U_{i_f}(x_i)$.

For the energy minimization, the simulated annealing Multiple Births and Deaths process (MBD) (Descombes et al., 2009) is chosen as the optimization algorithm because it has proven good performance in the literature when applied in combination with MPP for mapping tree plantations in a rural environment (Perrin et al., 2005). The principle of this algorithm is to alternate phases of "birth" (proposal of street trees) and phases of "death" (removal of the street trees that are not relevant in the sense of the defined energy). A temperature term that decreases during the process is used to explore different tree configurations. This is necessary in order to reach the global minimum of the energy and not to stop at a local minimum. First, we initialize the temperature T (fixed to 0.01) and the birth rate δ (equal to 200 which corresponds to the order of magnitude of the number of trees

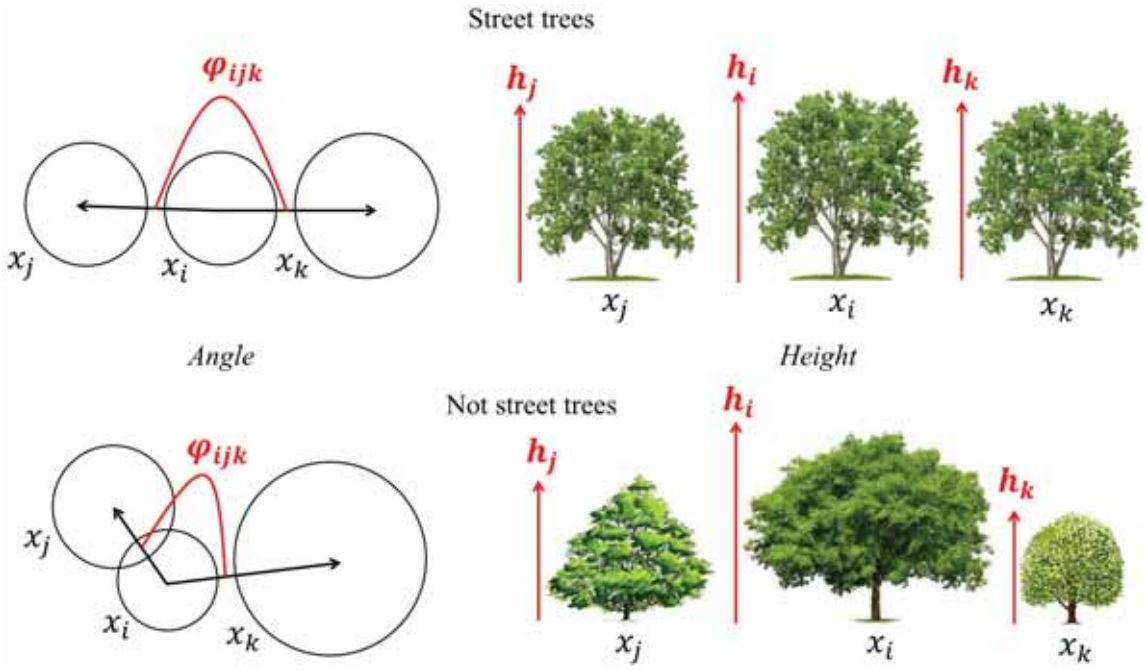

Figure III.5: Illustration of the contextual features used to compute the interaction energy. At the top, the case of street trees is described (similar height, tree species and alignment). At the bottom, the case of trees which are not street trees is presented (various height, different tree species and no alignment).

in the scenes). Concerning the value of the initial temperature which is similar to the one used in (Descamps et al., 2009), the minimized energy distributions resulting from higher initial temperature values (tests from 0.01 to 0.05) are similar. The lowest is then kept to reduce computational time. Then, the algorithm is defined as follows:

1. **Birth of the street trees:** For each pixel s of the VNIR image, if there is not already a street tree at this position, we place a street tree with the probability $\delta * B(s,r)$ at this position. $B(s,r)$ is proportional to the data energy $U_d(x_i)$ corresponding to a disk placed at the pixel s with a radius r and is used to reduce computational time as in (Descombes et al., 2009). Otherwise, the street trees would be randomly positioned uniformly.

2. **Sorting of the street trees according to their energy:** We compute the data energy $U_d(x_i)$ for each street tree x_i in the current street tree map. Then, the street trees are sorted according to decreasing data energy.

3. **Death of the irrelevant street trees:** For each street tree x_i taken in this order, we compute the death rate as follows:

$$\begin{aligned} d(x_i) &= \frac{\delta \cdot e(\mathbf{x}\backslash x_i)}{1+\delta \cdot e(\mathbf{x}\backslash x_i)} \\ where \quad e(\mathbf{x}\backslash x_i) &= \exp\left(\frac{\alpha \cdot U_d(x_i) + \beta \cdot U_{i_s}(x_i) + \gamma \cdot U_{i_f}(x_i)}{T}\right) \end{aligned} \tag{III.5}$$

   with α, β and γ corresponding to the weights of the different energies (fixed to 1, 1 and 3.5 respectively after testing multiple values). $\gamma = 0$ refers to "without θ and without h" (without angle information and without height information).

These three steps are repeated 1000 times as the street tree map does not change from one iteration to another at this stage of the process. In other words, the convergence is reached at this stage. In order the reduce computational time, γ is set to 0 for the first 600 iterations to map the trees without distinction between the street trees and the not street trees. From the 600th iteration, γ is set to 3.5 to map the street trees. At the end of each iteration, T and δ are multiplied by a factor of 0.997 (similar to that used in (Descombes et al., 2009)).

III.1.3.2 Baseline tree crown delineation used for comparison

To our knowledge, no individual street tree mapping has been proposed. Thus, as a baseline for comparison, we chose a standard tree crown delineation that is today the only available solution for the purpose of that paper, even if the objectives are not exactly the

Table III.5: Description of the confusion matrix. "True" refers to the real street trees and not street trees in the scene while "Predicted" refers to the predicted street trees and not street trees by the method under consideration. The "-" symbol signifies that we do not take into account the "True Negative" trees (well predicted not street trees). To consider the "True Negative" trees, it would be necessary to have the number of park trees in the first and second study cases, and the number of trees in private properties for the third study case, an information that is not available.

		Predicted		
		Street tree ($\widehat{st}$)	**Not street tree** ($\widehat{\overline{st}}$)	**Total**
True	**Street tree** (st)	True Positive (TP)	False Negative (FN)	TP+FN
	Not street tree ($\overline{st}$)	False Positive (FP)	-	
	Total	TP+FP		

same. The baseline tree crown delineation considered as the reference method is a region growing method developed in (Adeline, 2014) and inspired from the work of (Iovan et al., 2008). This type of approach is chosen as a baseline because it is commonly used in the literature (Zhen et al., 2016). In particular, a Canopy Height Model (CHM) is derived from a high vegetation mask obtained similarly to the one generated with the proposed method and the DSM. The CHM is smoothed with a Gaussian filter whose standard deviation s_{Gauss} is equal to 2 as in (Adeline, 2014). This allows the irregularities at the surface of the trees to be removed. Indeed, because of the foliage structure at the top of the trees, there can be multiple local maximums that do not correspond to multiple trees. The high vegetation mask is then treated such that the smallest regions are removed. This is done according to a parameter N_{min}^{tree} which defines the minimum number of pixel per tree (here equal to 5 as in (Adeline, 2014)). From this step, every pixel of the CHM is assigned to a particular tree by decreasing height. As an initialization step, the highest pixel of the CHM is chosen as the first pixel of the first delineated tree. Then the height is decremented and the corresponding pixel is either assigned to that first tree if it is at a distance d_{adj} less than 2 m here as in (Adeline, 2014), or assigned to a new tree, and so on.

III.1.3.3 Accuracy assessment

In order to assess the results of the methods and compare their performances, a confusion matrix is built by visual interpretation (table III.5).

From the confusion matrix, the Producer Accuracy (PA), the User Accuracy (UA) and the F-score are used to assess the performance specifically for the street trees: $PA(\%) = 100 \cdot \frac{TP}{TP+FN}$, $UA(\%) = 100 \cdot \frac{TP}{TP+FP}$ and $F-score(\%) = 100 \cdot \frac{2 \cdot PA \cdot UA}{PA+UA}$

III.1.4 Results

In this section, two main results are presented. First, the proposed method with all the features (with GIS, θ and h, section III.1.3.1, equation III.4) is compared to the baseline method for the three study sites. The F-score values and the produced street tree maps are described for both methods as well as the TP, FP and FN confusion matrix terms in order to get an exhaustive comparison of the two methods (overall and specifically for the street trees). Secondly, the contribution of the angle, the height and the GIS data in the street tree mapping is studied (with or without each feature). The eight possible combinations of these features (2 raised to the power 3) are analysed in order to identify the set of the most discriminative contextual features for street tree mapping.

III.1.4.1 Comparison between the proposed method and the baseline method

The figure III.6 presents the results of the proposed and reference methods overall. As expected, the proposed method outperforms the reference method with F-score values of 91%, 75% and 85% against 70%, 41% and 20% for the three study cases respectively. In addition, from the first to the second and third cases (more challenging cases), the F-score values of the proposed method remain stable whereas the F-score values of the baseline is decreasing. As another point of comparison, the baseline gives F-score values of 83%, 82% and 28% for the three cases when aware of the GIS data.

In order to explain that results, the figure III.7 shows the confusion matrices of the two approaches. In particular, there are a lot of "False Positive" trees (table III.5) for the reference method with 38, 290 and 613 trees (UA values of 55%, 26% and 11%) instead of 0, 0 and 21 trees (UA values of 100%, 100% and 77%) for the proposed method in the three cases respectively. This type of error is obviously expected because the reference method is not dedicated to map the trees with a differentiation between the street trees and the not street trees. Thus, while the proposed approach maps the street trees correctly in the three cases, the baseline method logically confuses the street and park trees in the first and second cases, and the street and private trees in the third case. This difference is logically less significant when the baseline uses the GIS with 16, 11 and 368 "False Positive" trees (UA values of 75%, 88% and 19%). Moreover, the F-score's decrease of the reference method from the first to the second and third cases is explained by a larger number of trees in the

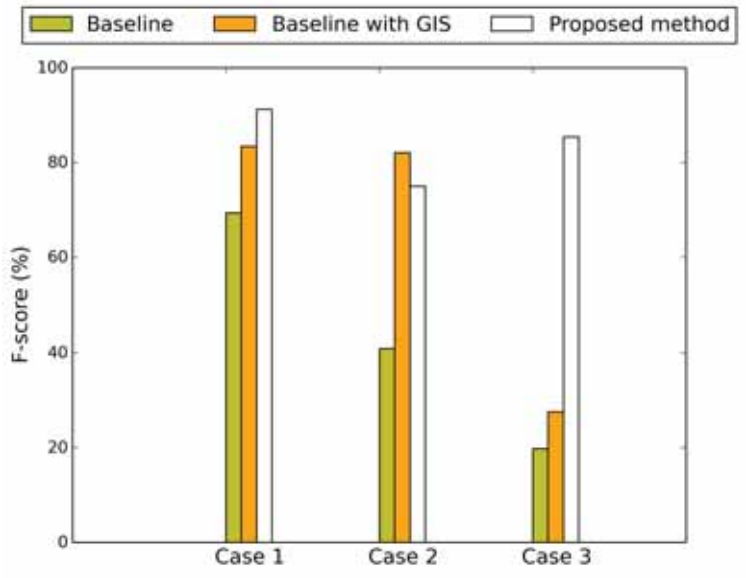

Figure III.6: Comparison of F-score between the proposed method and the baseline method for each study case.

third case than in the second case and a larger number of trees in the second case than in the first case. These statements are illustrated in figure III.8 where the produced maps of the baseline and proposed schemes are compared for the third case. We can see that the two methods map the actual street trees correctly but that the trees of the private properties are obviously identified as street trees by the reference method. Focusing on the baseline, this figure demonstrates also that a lot of not street trees can be filtered just using the GIS data.

In order to assess the performance of the two methods when we focus on the real street trees, the number of "False negative" trees (table III.5) is compared between the two approaches. In particular, there are 8, 43 and 3 "False Negative" trees (PA values of 84%, 60% and 96%) for the proposed method instead of 3, 4 and 0 trees (PA values of 94%, 96% and 100%) for the reference method. The proposed method tends to underestimate the number of street trees (especially for the second area as explained in the next section). In comparison, the reference method performs better in that it produces fewer "False Negative" trees, but over the 290 "False Positive" trees obtained for the second study site, 276 are in the park and 14 among the street trees lines (13% of the street trees). Thus many street trees are oversegmented by the reference method. The important overlaps of the

	Case 1		Case 2		Case 3	
			Baseline method (■)			
	$\widehat{st}$	$\widehat{\overline{st}}$		$\widehat{st}$	$\widehat{\overline{st}}$		$\widehat{st}$	$\widehat{\overline{st}}$
st	47	3	st	103	4	st	72	0
$\overline{st}$	38	-	$\overline{st}$	290	-	$\overline{st}$	613	-
			Baseline method with GIS (■)			
	$\widehat{st}$	$\widehat{\overline{st}}$		$\widehat{st}$	$\widehat{\overline{st}}$		$\widehat{st}$	$\widehat{\overline{st}}$
st	47	3	st	82	25	st	72	0
$\overline{st}$	16	-	$\overline{st}$	11	-	$\overline{st}$	368	-
			Proposed method (□)			
	$\widehat{st}$	$\widehat{\overline{st}}$		$\widehat{st}$	$\widehat{\overline{st}}$		$\widehat{st}$	$\widehat{\overline{st}}$
st	42	8	st	64	43	st	69	3
$\overline{st}$	0	-	$\overline{st}$	0	-	$\overline{st}$	21	-

Figure III.7: Comparison between the confusion matrices of the baseline method and the confusion matrices of the proposed method for the three study cases. Top: Confusion matrices for the baseline method. Bottom: Confusion matrices for the proposed method. The colors refer to the colors used in figure III.6. The prediction is per column.

trees in the second case is a reason of these errors. The reference method is sensitive to the irregularities at the surface of the canopies. On the other hand, the data energy defined in the proposed method is only based on the high vegetation mask which results in FN trees when there are overlaps of high vegetation.

III.1.4.2 Contributions of the angle, the height and the GIS data in the street tree mapping

The figure III.9 presents the results of the proposed method for different configurations of features in order to identify the best set of the discriminative contextual features of the street trees. With all the features (with GIS, θ and h, section III.1.3.1, equation III.4), the F-score values are 91%, 75% and 85% instead of 76%, 58% and 26% without any feature (without GIS, θ and h) for the three cases respectively. Without GIS but with θ and h, the F-score values become 88%, 79% and 62%. With GIS but without θ and h, the F-score

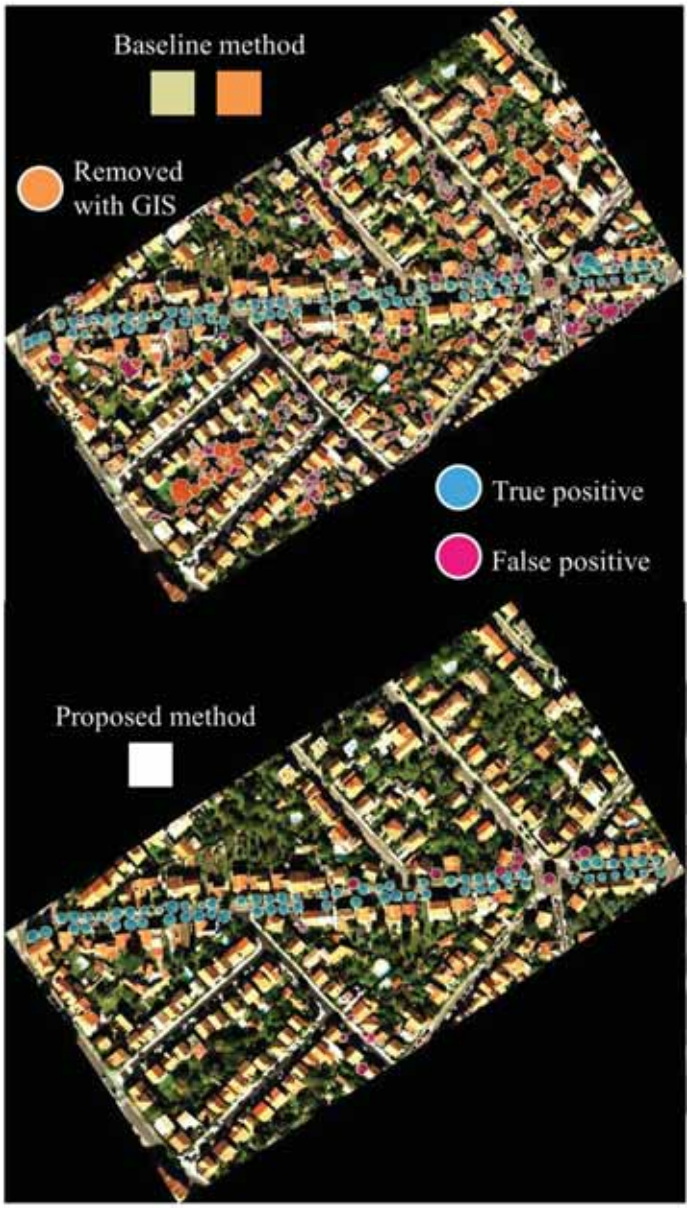

Figure III.8: Comparison of the produced maps for the baseline method and the proposed method for the third case. The colors under ”Baseline method” and ”Proposed method” refer to the colors used in figure III.6.

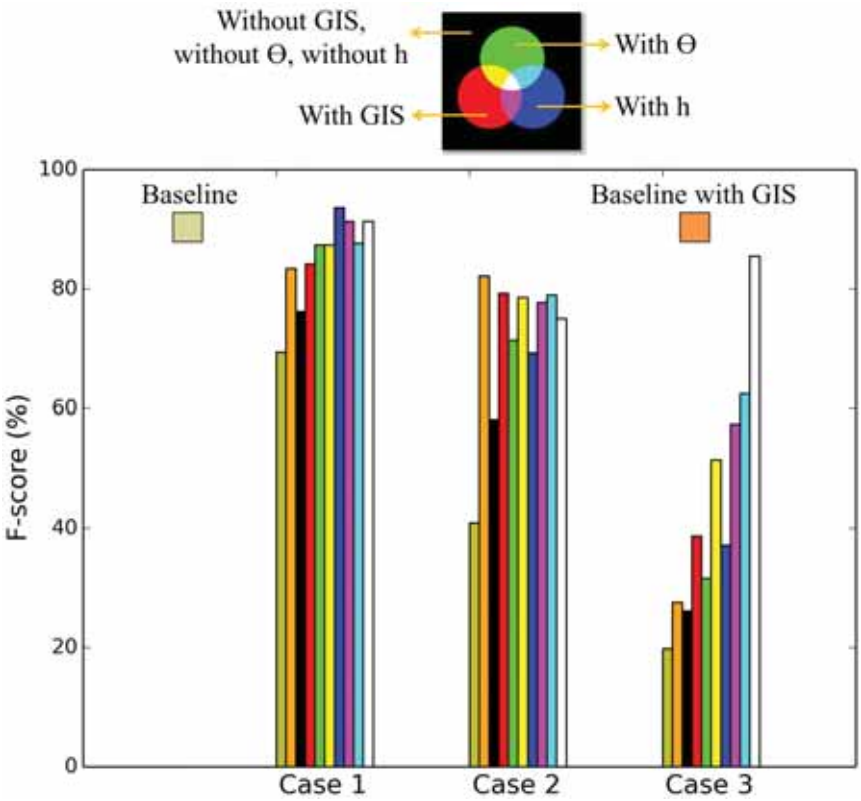

Figure III.9: Contribution of the angle, the height and the GIS data in terms of F-score for each study case. Each color corresponds to a combination of red (if GIS used), green (if θ used) and blue (if h used) colors. The color under "Baseline method" and the white color refer to the colors used in figure III.6.

values are equal to 84%, 79% and 39%. These results demonstrate that there is a benefit to exploit together the GIS data, the angle between the trees and the heights (F-score improvements of 15pp, 17pp and 59pp compared to the case where no feature is used). However, using only θ and h gives already good results. On the other hand, using only the GIS data is not appropriate and needs the integration of θ and h. Focusing on θ and h, they have to be used together. The figure III.10 illustrates the contribution of the angle between the trees and the heights. With the integration of these features, the majority of the street trees are mapped correctly. This result is expected because the street trees form lines and are mostly the same height which is highlighted in the interaction energy defined in the proposed method.

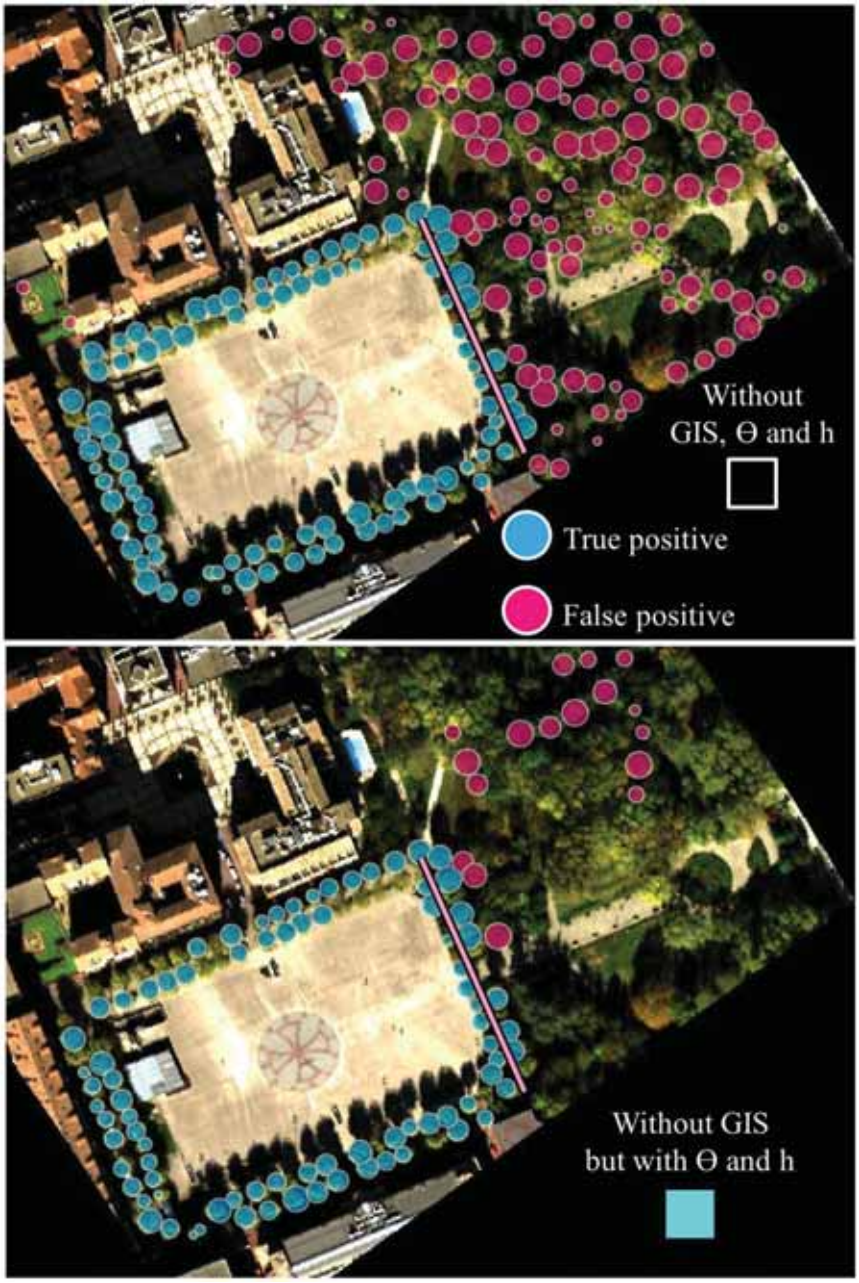

Figure III.10: Comparison of the produced maps with and without the use of θ and h for the second case. The colors under "Without GIS but with θ and h" and "Without GIS, θ and h" refer to the colors used in figure III.9. The pink road is a not considered pathway.

In order to go further in the analysis, the figure III.11 shows the confusion matrices of the three main configurations of features. The F-score improvements obtained with the integration of the street tree features is mainly explained by a decrease of the number of "True Negative" trees. Nevertheless, the evolution of the number of "False Negative" among the configurations of features shows that the "False Negative" tree number is increasing with the integration of θ and h. Especially, there are 3 further "False Negative" trees (PA decreases of 6pp and 4pp) with the integration of these features for the first and second cases when using the GIS data. Even if these errors are marginal, this trend is observed among the set of simulations. In fact, the street trees do not form perfectly straight lines and that they do not have strictly the same heights (hypothesis not always verified). Thus a too strong integration of these features (via the parameter γ in the proposed method, section III.1.3.1, equation III.5) can result to consider some "non perfect" street trees as not street trees. In the second case, the number of "False Negative" trees is particularly high when using GIS data. Indeed, in the GIS data, there is no considered road at the right side of the square as it is a pathway in the park (figure III.10, pink road). As a consequence, all the street trees along this pathway are filtered at the high vegetation detection step when using GIS data as they are too far from the closest roads (other sides of the square).

III.1.5 Discussions

III.1.5.1 Individual tree detection in its context

These results demonstrate the ability of the proposed method to detect the street trees in three different circumstances, while a standard tree crown delineation obviously does not allow the specific urban tree structures to be identified. This performance is consistent with that of (Wen et al., 2017) who obtained a F-score of 89% when mapping roadside canopies with a patch-level approach. However, our scheme maps the trees individually, which is essential for an individual health monitoring of the street trees. Indeed, the prevention of the fall of sick trees cannot be carried out if the trees are not mapped individually. In the urban environment, the alignment trees are subject to specific conditions as they have little space for growth, are pruned and can be affected by the spread of diseases in single-species plantations (Fini et al., 2015; Sebestyen et al., 2008).

The proposed study also highlights the interest of considering the tree in its context, i.e. considering tree structures. In addition to their usefulness for urban managers, the tree structures could be used in order to improve not only the individual tree mapping itself (by taking advantage of a prior contextual information knowledge about the urban trees depending on their structure), but also the tree species classification for example (by defining

Case 1 Case 2 Case 3

With GIS but without θ and h (■)

	$\widehat{st}$	$\widehat{\overline{st}}$
st	45	5
$\overline{st}$	12	-

	$\widehat{st}$	$\widehat{\overline{st}}$
st	74	33
$\overline{st}$	6	-

	$\widehat{st}$	$\widehat{\overline{st}}$
st	72	0
$\overline{st}$	229	-

Without GIS but with θ and h (■)

	$\widehat{st}$	$\widehat{\overline{st}}$
st	39	11
$\overline{st}$	0	-

	$\widehat{st}$	$\widehat{\overline{st}}$
st	82	25
$\overline{st}$	19	-

	$\widehat{st}$	$\widehat{\overline{st}}$
st	70	2
$\overline{st}$	83	-

With GIS, with θ and h (□)

	$\widehat{st}$	$\widehat{\overline{st}}$
st	42	8
$\overline{st}$	0	-

	$\widehat{st}$	$\widehat{\overline{st}}$
st	64	43
$\overline{st}$	0	-

	$\widehat{st}$	$\widehat{\overline{st}}$
st	69	3
$\overline{st}$	21	-

Figure III.11: Comparison between the confusion matrices depending of the contextual features of the street trees used for the three study cases. Top: with GIS but without θ and h. Middle: without GIS but with θ and h. Bottom: with GIS, with θ and h. The colors refer to the colors used in figure III.9. The prediction is per column.

specific categories of urban trees depending on their structure because street trees have not necessarily the same spectral traits than park ones). This type of contextual approach is not only applicable in the urban environment, but also in the rural environment with the case of plantations. Regarding natural forests, the vegetation grows in a specific way, depending on sunshine, temperature, moisture, soil, species and other neighbourhood properties (Alvarez-Uria and Körner, 2007). As an example, it is well known that deciduous trees can not grow beyond a certain altitude (Miyajima and Takahashi, 2007). Such contextual elements could be studied and used as prior information in order to better understand the urban and non-urban tree ecosystems, and improve the existing mapping algorithms.

Nevertheless, modelling the context is not sufficient. Indeed, the results of this study highlight that the proposed method is overall successful in mapping the street trees but makes some errors in cases of significant overlap. This is a known issue in the literature of the individual tree crown delineation methods (Zhen et al., 2016). As an improvement and similarly to the employed reference method, the height could be used in addition to the high vegetation mask in the data energy. Also, the more accurate individual tree modelling of the other standard tree crown delineation algorithms (Leckie et al., 2003; Zhen et al., 2015; Chen et al., 2006) could be of benefit to the definition of the data energy in the MPP approach. On the other hand, the mask extraction is based on simple thresholding procedures that produce hard masks, which are combined using intersection operator. They may be prone to noise and may not recover from any artefacts in any of the data sources, which can cause errors in cases of significant overlap. The figure 12 shows the behaviour of the proposed method with all the features (with GIS, θ and h, section 3.1.2, equation 4), from the NDVI threshold estimated with the OTSU algorithm, to lower (-0.1) and higher (+0.1) values of the NDVI threshold. Even if the number of high vegetation pixels decreases from the lowest to the highest NDVI threshold value, especially around the trees, the street tree maps are comparable. Only, few differences in terms of radius are observed, but it seems more related to the optimization process. Although the use of hard masks is appropriate here (not very sensitive to the value of the threshold), mainly because of the object-based approach, softer density images of birth probability could be used instead to improve the performance, through another definition of the data energy.

III.1.5.2 Discriminative contextual features of the street trees

This research aims at highlighting the discriminative features of the street trees in order to map them accurately. In comparison to the work of (Wen et al., 2017) who used a series of patch-level metrics that describe the spatial patterns of the roadside canopies, the proposed method is mainly based on the angle between the trees and the heights, which

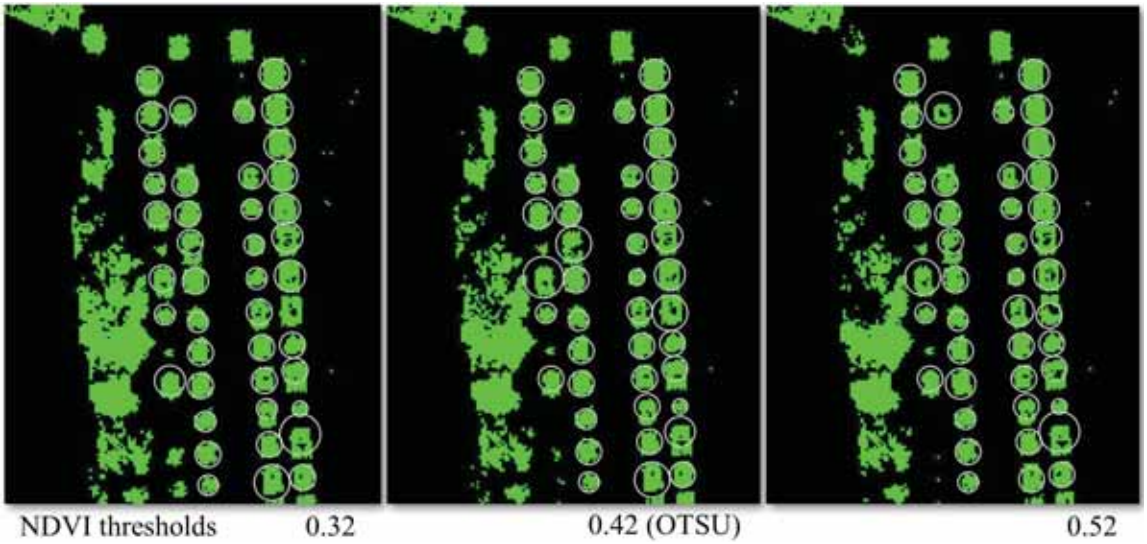

Figure III.12: Street tree maps obtained with the proposed method for different values of NDVI thresholds in the first study case. All the features (with GIS, θ and h, section III.1.3.1, equation III.4) are used. Values of shadow, height and distance thresholds are fixed.

are simple characteristics derived from field observations. Especially, the results show that the angle and the height are essential parameters in order to get a correct street tree map. When adding the GIS data, the performance of the baseline and proposed approaches is improved but using the GIS data only is not sufficient. This can be understood by the spatial connection between the street trees and other populations of trees such as park, private trees, etc. (particularly shown in the third study area). Also, the GIS data can lead to confusion as shown in the second study site and highlights that the proximity to the roads is probably not an intrinsic feature of the street trees. The street trees often have the highlighted properties in the cities around the world but other features could be used for improving the performance of the proposed method such as the distance between two consecutive trees in an alignment, or spectral similarities within an alignment. This type of discriminative feature identification could be carried out for mapping other types of vegetation such as hedges which form lines and have similar heights. The same statement can be made regarding the detection of the vineyards.

However, using more and more features requires logically the definition of as many parameters. In this study, the MPP parameters are the same for the three cases which demonstrates that the proposed scheme is robust from a case to another, but that the set of parameters could be better estimated for each case. The parameter estimation is an important step which effectively impacts the performance of the methods based on MPP (Chatelain et al., 2009; Hadj et al., 2010). On the other hand, the MBD optimization algorithm has been chosen, in particular for its very good speed of convergence and its simplicity of implementation. Instead, other algorithms such as Reversible Jump MCMC (RJMCM) or Multiple Birth and Cut (MBC) could have been used. While the MBD outperforms the RJMCMC in terms of speed of convergence, the MBC reaches a lower energy level than the MBD but in longer time (Descombes et al., 2009; Gamal-Eldin et al., 2010). Thus the MBC could be considered as another optimization algorithm in order to get a more accurate street tree map.

III.1.5.3 Applicability of the proposed method in other cities

The proposed framework has been applied to three study areas and the observed trends are the same among these different conditions. But of course, the applicability of the method in other cities is not assessed here. In order to get an idea of that applicability, we have computed the street tree feature energy (section III.1.3.1, equation III.4) of all the inventoried street trees in the tree database of Paris (figure III.13). This database contains, among other information, the location of the street tree trunks and the heights.

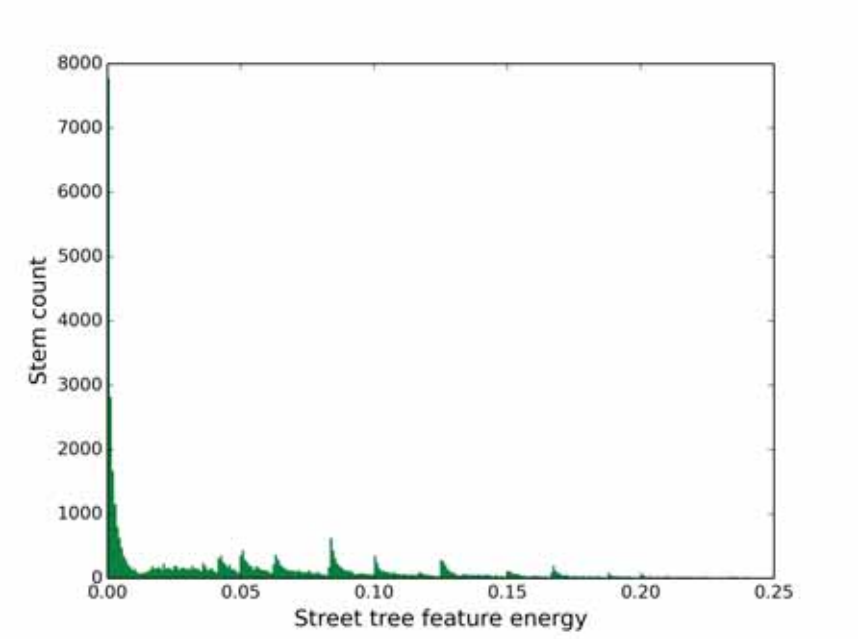

Figure III.13: Histogram of the street tree feature energy (section III.1.3.1, equation III.4) computed from 43168 street trees of the tree database of Paris. Only the main species are taken into account: plane tree, horse chestnut, pagoda tree and lime tree (*Platanus x hispanica*, *Aesculus hippocastanum*, *Sophora japonica* and *Tilia tomentosa*).

Table III.6: Comparison of the baseline and proposed methods in terms of computational burden.

Framework		Case 1	Case 2	Case 3
Without GIS	**Baseline**	∼10s	∼10min	∼30min
	Proposed	∼20min	∼50min	∼3h
With GIS	**Baseline**	∼10s	∼1min	∼20min
	Proposed	∼20min	∼30min	∼1.5h

This histogram shows that the street trees of Paris have overall the discriminative contextual features highlighted in this paper. In particular, the peak near 0 corresponds to around 8000 street trees that are perfectly aligned with their neighbours and have exactly the same heights than their neighbours. And knowing that the height resolution of the Paris tree database is 1 m, the other local peaks highlight recurrent differences in height. For example, the peak around 0.08 corresponds to perfectly aligned street trees with a height of 12 m and two neighbours with heights of 11 m, or a neighbour with a height of 11 m and the other with a height of 13 m, or a neighbour with a height of 14 m and the other with a height of 12 m, etc. On the other hand, the flexibility of the alignment model based on three trees should allow to deal with curved roads. This is encouraging with a view to doing the street tree map of Paris and this attests the potential of the proposed method for other cities.

Another key element when talking about the applicability of the proposed method in other cities is the computational burden of the method. The table III.6 highlights the duration of the baseline and proposed approaches, knowing that the baseline is written in Interactive Data Language (IDL) and the proposed one in python. The baseline method is the best in terms of computational burden with execution times of approximately 10s, 10min and 30min (10s, 1min and 20min with the GIS data) for the three study cases, instead of 20min, 50min and 3h (20min, 30min and 1.5h with the GIS data) for the proposed method. Logically, the execution time decreases when using the GIS data because the high vegetation mask covers a smaller area. This table highlights also that the behaviour of the baseline and proposed approaches are different according to the increasing surface covered by the high vegetation mask from the first to the third case. Indeed, using the GIS data, the baseline approach is 120 times faster than the proposed one for the first study case, whereas the baseline becomes 30 and 4.5 times faster for the second and third cases, respectively. From a small scene to a larger scene, the number of pixels increases more than the number

of trees. The baseline approach suffers from its pixel-to-pixel approach which implies the consideration of each pixel during the delineation process. As the data energy can be computed in advance for the proposed method, it can benefit from its object oriented principle. Finally, even if the computational burden is not an issue for a city-wide application (because it does not need very frequent updating), and that the baseline gives better execution times for the cases under consideration in this study, the proposed framework may be more appropriate for scenes larger than those considered here.

III.1.5.4 Improving tree species classification in urban alignment

For several applications such as state of health monitoring, tree species information is essential (Fassnacht et al., 2016). Remote sensing gives encouraging results in tree species classification, especially thanks to hyperspectral data (Alonzo et al., 2014), but in urban environment it remains a challenging task because of the large tree diversity (species, age, life conditions, pruning, etc.) (Welch, 1982). As a case in point, we propose to map the species of the third study area, whose alignment is composed of *Platanus x hispanica* trees). In particular, the tree crowns are first estimated thanks to the proposed method with all the features (with GIS, θ and h, section III.1.3.1, equation III.4). Then, the tree species are classified thanks to an object-based approach similar to that used in (Alonzo et al., 2014). In fact, the VNIR pixels within the tree crowns are classified, followed by a majority vote for each crown. The learning step is carried out from pixels of a reference site situated near the first study area, and the Minimum Noise Fraction (MNF) components are used as feature vector. Focusing on the figure III.14 (top: before regularization), there are mainly errors in the right part of the alignment (18 trees on the basis of 69 detected are misclassified). From that baseline, the proposed method can be used in order to regularize the species estimation within the alignment. Indeed, the different tree triplets identified thanks to the street tree feature energy $U_{i_f}(\mathbf{x})$ (section III.1.3.1, equation III.4) can easily be linked to form networks. The majority species of each network can then be assigned to the corresponding trees. With 100% of correct predictions (over the 69 detected trees) for this study case, the figure III.14 (bottom: after regularization) demonstrates the potential of the proposed method for improving tree species classification within urban alignment. Even if a limit of that approach occurs for cases of alignments with multiple species, the proposed method could be modified in order to handle these cases, only if the trees are planted in a specific way (bispecific alignment whose species is alternated one in two for example).

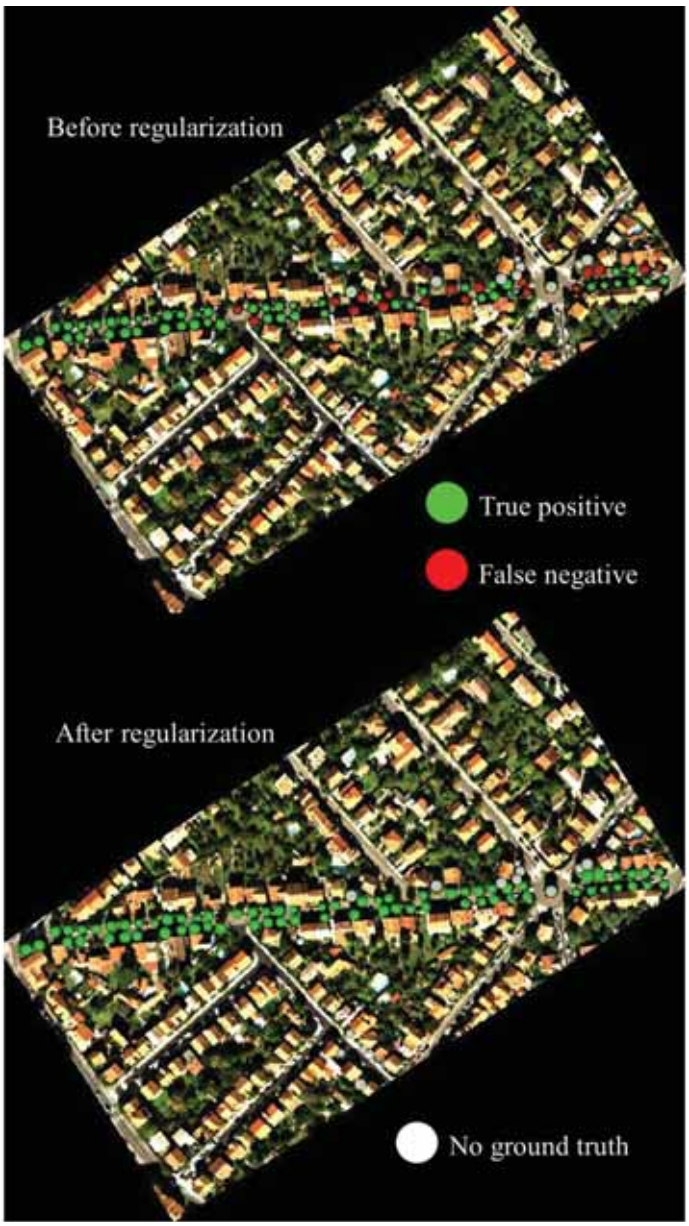

Figure III.14: Improvement of tree species classification in urban alignment thanks to the proposed method for the third case (whose alignment is mainly composed of *Platnaus x hispanica* tress).

III.1.6 Conclusions

The objective of this study is to map the street trees using airborne data and contextual information based on MPP method. Three test sites are considered for assessing the performance of the proposed method under different conditions. Airborne hyperspectral data, a DSM and GIS data which included the roads are used, but multispectral data could be used instead of the hyperspectral data. From these data, the vegetation canopies close to the streets are detected thanks to simple thresholds of NDVI, shadow index, height and distance to the streets respectively. The obtained high vegetation mask is then treated through a scheme based on MPP. In particular, the discriminative contextual features of the street trees (hypotheses of small angle between the trees and similar heights) are modeled in the interaction energy of the MPP. As a baseline for comparisons, a standard region growing crown delineation approach is considered.

Regarding the results, the proposed method logically outperforms the reference method with overall F-score values on the three study sites of 85% with all the features against 44% with differences of 15pp, 38pp and 65pp respectively in favor of the proposed method. It demonstrates the ability of the proposed method to map the street trees in three different circumstances. Focusing on the contributions of the discriminative contextual features in the individual street tree mapping, the F-score values are 91%, 75% and 85% with all the features (with GIS, θ and h). Without GIS but with θ and h, the F-score values become 88%, 79% and 62% (-3pp, +4pp and -23pp). It is thus more appropriate to exploit together the GIS data, the angle between the trees and the heights. Nevertheless, using only θ and h gives already good results (76% overall on the three study sites). Finally, the GIS data alone is not sufficient.

Further work is necessary and there are many perspectives. In this paper, we see that the proposed method can be improved in the case of significant overlap. In addition, the same set of parameters have been used for the three cases. Compared to the obtained results, it shows that the proposed approach is robust but highlights the importance of the parameter estimation step. Also, other features could be used in order to model the street trees. Moreover, the method will be applied on more difficult cases, like area with street trees forming a homogeneous canopy. In theses cases, it is often impossible to distinguish the crowns from the spectral data and from the DSM, which encourages the use of another technology. For example, multitemporal data acquired during the winter could help. In the long term, the proposed approach could be improved to map the other populations of trees such as park trees. The species and the state of health of the trees will be also of interest.

Acknowledgments

The authors would like to thank the French Aerospace Lab and the Occitanie region for funding this research. Regarding airborne data, thanks to Karine Adeline, Philippe Deliot and the French Mapping Agency for their participation in the UMBRA campaign. Thanks to the members of the Hyperspectral imagery for Environmental urban planning (HYEP) project, funded by the French National Research Agency (ANR). Thanks to Bertrand Peyce, Etienne Ducasse and Manon Daumur for discussions about the features of the urban environment.

General discussion

Synthèse en français

L'objectif de cette thèse est la cartographie des arbres en milieu urbain, en particulier l'identification des espèces, fondées sur plusieurs sources de données aéroportées, hyperspectrales, PAN et nDSM, en tirant profit de leur complémentarité. Pour ce faire, une première partie est dédiée à l'évaluation de cette complémentarité, et à la sélection de la meilleure stratégie entre la fusion au niveau des caractéristiques et la fusion au niveau de la décision. Ensuite, la nature des résultats obtenus nous a conduit à optimiser la complémentarité, en explorant en profondeur la richesse des données hyperspectrales à travers un ensemble de classifieurs. Enfin, dans la troisième partie, une méthode de détection des arbres d'alignement de manière individuelle est proposée, la première partie soulignant l'intérêt de discriminer les différentes structures d'arbres en milieu urbain (alignements, parcs, etc.) pour la classification des espèces.

Classification des espèces d'arbres urbains à partir de multiples sources de données aéroportées

Les résultats du premier chapitre montrent que les données hyperspectrales sont le principal moteur de la précision de classification, tandis que les caractéristiques fondées sur les données PAN et nDSM contribuent marginalement. Ceci est cohérent avec les travaux précédents (Fassnacht et al., 2016). Deuxièmement, notre travail renforce l'intérêt d'appliquer des algorithmes de réduction de dimension tels que le MNF pour traiter des données de grande dimension telles que les données hyperspectrales (Ghosh et al., 2014a; Zhang and Xie, 2012). Alors que certaines espèces avec des signatures spectrales uniques sont bien identifiées, par exemple *Aesculus hippocastanum*, *Cedrus deodara*, *Platanus x hispanica*, la procédure de classification proposée échoue à correctement classer d'autres espèces telles que *Acer platanoides* ou *Acer negundo*. Ces erreurs de classification sont dus à une dynamique phénologique élevée au sein des individus de ces espèces. Cette

limite suggère l'utilisation de données acquises au cours d'une autre saison, en particulier l'été. Des données multitemporelles sont également intéressantes pour caractériser la phénologie de chaque espèce, utile pour la discrimination des espèces d'arbres (Sheeren et al., 2016). Pour la fusion, la légère amélioration obtenue démontre que le VNIR seul est suffisant. Cette constatation est cohérente avec les travaux précédents (Dalponte et al., 2008; Jones et al., 2010), mais d'autres obtiennent des améliorations plus significatives (Dalponte et al., 2012; Alonzo et al., 2014). Même si les méthodes de fusion proposées dans cette thèse ne permettent pas d'augmenter significativement la précision de classification, une analyse de complémentarité des sources a été réalisée et explique clairement pourquoi la fusion n'est pas bénéfique. Le résultat principal de cette analyse est que les sources ne sont que très faiblement complémentaires. Alors que la stratégie de fusion au niveau de la décision proposée permet de tirer profit de ces rares cas, elle échoue à préserver la performance originale du VNIR. Cette thèse montre donc que si les sources sujettes à la fusion ne sont pas très complémentaires, les performances ne peuvent être améliorées, même avec une stratégie de fusion efficace. Les approches d'ensemble de classifieurs où les classifieurs sont très complémentaires pourraient être considérées.

Classification des espèces d'arbres urbains à partir de multiples classifieurs spectraux

Les résultats du second chapitre démontrent que l'ensemble de classifieurs proposé est meilleur qu'une approche standard qui concatène toutes les caractéristiques dans le même vecteur de caractéristiques. Ceci est cohérent avec les résultats antérieurs (Ceamanos et al., 2010; Engler et al., 2013). Cela peut s'expliquer par trois raisons principales. L'approche d'ensemble est particulièrement robuste car il existe plusieurs classifieurs par espèce. Deuxièmement, la concaténation des triplets d'indices dans le même vecteur de caractéristiques conduit à des caractéristiques qui sont discriminantes pour certaines espèces, mais qui jouent le rôle de bruit pour d'autres, ce qui peut provoquer des erreurs. Enfin, l'effet de Hughes, même s'il est moins significatif que pour la réflectance spectrale (Hughes, 1968), est comparativement plus important que dans la méthode d'ensemble (seulement trois indices par classificateur). Plus le nombre de caractéristiques est important, plus ces deux derniers phénomènes sont présents. Afin d'améliorer la méthode proposée, des classifieurs supplémentaires pourraient être conçus pour séparer des groupes de classes (feuillu, conifère, genre, etc.).

Classification des espèces d'arbres urbains à partir de caractéristiques spectrales et contextuelles

Les résultats du troisième chapitre montrent que les arbres d'alignement peuvent être cartographiés avec une grande précision grâce à l'approche MPP proposée. Cependant, bien que la méthode proposée réussisse à identifier les arbres d'alignement qui ne se chevauchent pas, des déclins sont observés dans les cas de chevauchement important, comme dans la deuxième zone d'étude. Ce comportement correspond à un problème bien connu (Zhen et al., 2016; Alonzo et al., 2014), mais notre modélisation simple des arbres individuels (fondée uniquement sur les valeurs binaires du masque) est particulièrement sensible au chevauchement. Il y a donc intérêt à modéliser plus précisément les arbres, en utilisant par exemple leur structure 3D. De nos jours, les approches efficaces de délimitation de la couronne arborescente sont basées sur les données LiDAR (Gupta et al., 2010), l'énergie des données du MPP pourrait donc être inspirée de ces méthodes. Un terme d'attache aux données fondé sur le nDSM pourrait déjà être efficace pour améliorer les performances dans de tels cas de chevauchement.

Conclusions

L'objectif de cette thèse est d'améliorer les approches de classification des espèces d'arbres. La contribution principale du travail de thèse est le développement d'approches de cartographie des espèces d'arbres urbains qui exploitent la complémentarité de plusieurs sources d'informations, de différents capteurs aéroportés à de multiples classifieurs spectraux, en passant par l'information contextuelle. Toulouse, France, est la zone d'étude et des données hyperspectrales VNIR et SWIR, PAN et nDSM sont considérées. La conclusion principale résultant du premier chapitre est que le VNIR seul est suffisant dans notre contexte (OA value of 75%). En se concentrant sur la deuxième partie du travail, le résultat principal est que l'ensemble de classifieurs proposé conduit à de meilleurs résultats que l'approche concaténée standard en classant 5 espèces, avec une précision globale moyenne de 60% contre 58%. En ce qui concerne le troisième chapitre, la détection des arbres d'alignement proposée est fonctionnelle pour les couronnes ne se chevauchant pas, avec un taux de détection moyen de 85% dans les trois zones d'étude considérées.

Perspectives

A la fin de cette thèse, il reste diverses limites. En se concentrant sur l'objectif principal du travail, la classification des espèces d'arbres, la performance et l'efficacité des approches doivent être améliorées. Un premier levier est la considération de plus d'échantillons

d'apprentissage, générés via un modèle de transfert radiatif par exemple. De plus, comme les erreurs de délinéation sont critiques (en particulier pour les arbres des parcs), le paradigme actuel de la cartographie des arbres urbains, à savoir la délimitation de la couronne suivie de la classification des espèces, pourrait être remis en question en considérant une approche unifiée. Parce que les données hyperspectrales sont rares, une réflexion sur des systèmes d'acquisition complémentaires est nécessaire. D'un autre côté, plus d'informations contextuelles pourraient être utilisées, pour notre application mais aussi pour d'autres. Enfin, l'état de la santé constitue le but ultime de ce travail et cela pourrait être étudié à partir de données multitemporelles.

English part

The objective of this thesis is the urban tree mapping, in particular the species identification, based on several airborne data sources, hyperspectral, PAN, nDSM, and contextual information by taking advantage of their complementarity. To do so, a first part is dedicated to assess this complementarity, and select the best fusion framework between the feature and decision level ones. Then, the nature of the obtained results lead us to optimize the complementarity, by exploring deeply the richness of the hyperspectral data through an ensemble classifier approach. Finally, in the third part, a method for detecting individual street trees is proposed, as the first part highlighted the interest of discriminating the different tree structures in the urban environment (alignment, park, etc.) for the species classification task.

Urban tree species classification from multiple airborne data sources

The first chapter of this thesis is dedicated to identify the best object-based fusion strategy taking advantage of the complementarity of several heterogeneous airborne data sources for improving the classification of 15 tree species in an urban area (Toulouse, France). Hyperspectral VNIR and SWIR, PAN and nDSM data are considered for that purpose. A decision level approach is proposed and compared to a standard feature level fusion. Whereas a first step is the extraction of spectral (MNF components), textural (Haralick parameters) and structural (height ratios) features in order to characterize the tree species, a second step aims at classifying the resulting feature vectors within the tree crowns. Then, a decision profile is computed for each crown and subject to a decision rule, leading the species predictions.

The results of this chapter raise that the hyperspectral data are the main driver of the classification accuracy, whereas the PAN and nDSM-based features contribute marginally. This is consistent with previous works (Fassnacht et al., 2016). In particular, the VNIR sources are better than the SWIR ones, a result that may be attributed among others to the better spatial resolution of the VNIR data in comparison to the SWIR, resulting in an lower number of examples in the SWIR training set. However, other works related to tree species classification demonstrate that the VNIR is more powerful than the SWIR owing to wavelengths sensitive to pigments and red edge in this spectral domain, independently of the spatial resolution (Fassnacht et al., 2016). Secondly, our work reinforces the interest of applying feature reduction algorithms such as the MNF to deal with high dimensional

data such as hyperspectral data, especially when the number of training samples is small in comparison to the dimension of the data (Ghosh et al., 2014a; Zhang and Xie, 2012). The analysis of the spectral signatures based on the contribution of each spectral band in the MNF components reveals that the red edge region seems to be particularly relevant for the classification task. This finding is consistent with previous studies but accentuated here, probably because of the period of acquisition (autumn). Whereas certain species with unique spectral signatures are well identified, e.g. *Aesculus hippocastanum*, *Cedrus deodara*, *Platanus x hispanica*, the proposed classification procedure fails in correctly classifying other ones such as *Acer platanoides* or *Acer negundo*. These poor performance cases are due to high phenological dynamics among the individual of these species. This limit suggests the use of data acquired during another season, especially summer. Multi-temporal data are also of interest in order to characterize the phenology of each species, useful for tree species discrimination (Sheeren et al., 2016). Other species with intermediary accuracies, e.g. *Liquidambar styraciflua*, *Celtis occidentalis*, are slightly confused with other ones, because of the inherent tree diversity in the urban environment (life conditions, pruning, age, etc.), with potentially a small number of samples. To deal with this limit, it seems that the richness of the hyperspectral data needs to be explored more deeply, in order to highlight the intrinsic spectral properties of each species. It can be carried out through the use of specific spectral indicators such as vegetation indices, even if the reviewed studies in the recent work of (Fassnacht et al., 2016) do not demonstrate a clear advantage of their use.

On the other hand, the textural and structural features used in this thesis contribute marginally in the classification of tree species. While previous studies aiming at quantifying the contribution of textural and structural features when included with multispectral data find a significant improvement (Zhang and Hu, 2012; Franklin et al., 2000), that contribution is slight when hyperspectral data are considered (Alonzo et al., 2014). This behaviour can be easily understood as by their nature, hyperspectral data are more powerful than multispectral data for tree species classification, especially when several species are targeted. The textural and structural features have therefore to be more accurately modelled. Based on the 14 cm PAN images available for this thesis, the extracted textural parameters reflect the specific spatial arrangement of the foliage for species with large leaves such as *Platanus x hispanica* or *Tilia tomentosa*, which may explain why these two species are well classified with PAN-based features. However, this spatial resolution is not appropriate for the majority of the species which have leaves size smaller than decimetric resolution. This suggests the use of an acquisition system with a better spatial resolution, for example an Unmanned Aerial Vehicle (UAV) based system. Moreover, sub-crown derived features depending on radiometric levels (sun / shadow for example) can help in

improving the performance based on textural features. Focusing on structural features, their potential appears to be limited if only nDSM-based features are computed. Indeed, the characteristics that can be extracted from surface models are only related to the 3D shape of the trees, which varies significantly among the individuals of a particular species due to pruning. For example, only the pruned *Tilia tomentosa* street trees are well identified in the reference site, not the park ones, poorly represented and having various shapes. Dense point clouds derived from LiDARs would allow intra-crown structural features to be considered, features which seem to be more intrinsic among the species. The results of this chapter highlight also the interest of discriminating the different tree structures in the urban environment (street trees, park trees, etc.), as the performance of the proposed classification procedure is particularly high for the alignment trees in comparison to the one obtained for the park individuals. Based on such information, a specific processing could be applied for the classification of park trees. A regularization of the species prediction within an alignment could be carried out, as the street trees often belong to mono-specific species alignments. However, there may also be bi-specific species ones.

Focusing on the fusion, the slight improvement obtained demonstrates that there is little interest in fusing hyperspectral VNIR and SWIR, PAN and nDSM data in our context, when compared to the performance resulting from the VNIR source alone. This finding is consistent with previous works (Dalponte et al., 2008; Jones et al., 2010), but other ones obtain more significant improvements (Dalponte et al., 2012; Alonzo et al., 2014). Even if the fusion framework proposed in this thesis does not allow the classification accuracy to be strongly increased, a complementarity analysis of the sources has been carried out and explain clearly what happens through the fusion. In particular, its objective is to quantify the contribution of each data type in the classification task and then, to check if the proposed fusion success in taking advantage of this complementarity. The main result of this analysis raises that there are indeed few cases of complementarity. Whereas the proposed decision level fusion success in taking advantage of these cases, the fusion fails in preserving the original performance of the VNIR, the best individual source. This thesis therefore shows that if the sources subject to fusion are not highly complementary, the performance can not be improved, even with an efficient fusion model. Ensemble classifier approaches where the classifiers are highly complementary could be considered. On the other hand, this chapter shows that a decision level fusion is better than a feature level fusion in our context of low complementarity. First, this is explained by the Hughes effect (Hughes, 1968), because the feature level implies the use of feature vectors with an higher number of variables (spectral, textural and structural together). Secondly, as the textural or structural features are mainly useful for few species, they disturb the classification for the other species. Finally, the residual error resulting from the registration step may cause

declines. Once again, ensemble classifier could be interest to deal with these issues. The decision fusion developed in this chapter has been validated for a test site, independent of the reference site used for learning. Although a more significant improvement is obtained in comparison to the reference area, mainly because we focus on the two majority species, *Platanus x hispanica* and *Tilia tomentosa*, the slight cost-benefit ratio of the fusion encourages the use of the VNIR only, with a deeper investigation of the richness of the hyperspectral-based features through an ensemble classifier approach.

Urban tree species classification from multiple spectral classifiers

Following the results of the first chapter, the second chapter of this thesis is dedicated to develop an ensemble classifier approach based on ground references for classifying 5 species in Toulouse city, France. Airborne VNIR hyperspectral data are considered while leaf and canopy level spectral signature measurements are used in order to train the supervised classification algorithms. In particular, each species is associated to a specific classifier, before a decision rule is applied to predict the species. Each classifier consists of three vegetation indices followed by SVM classification, and the triplets of indices are chosen to optimize the F-score of each species. Two baselines are used for comparison: the direct use of the spectral reflectance and the use of a feature vector composed of all the triplets of indices.

The results of this chapter demonstrates that the proposed ensemble approach is better than a standard one which stacks all the features is the same feature vector. This is consistent with previous findings (Ceamanos et al., 2010; Engler et al., 2013). This can be explained by three main reasons. The ensemble approach is particularly robust because there are several classifiers per species. In such way, the overfitting phenomenon which can happen through the derivation of the best three spectral indices, is alleviated. More there are classifiers per species, more the performance will be better, as long as each classifier reaches an high F-score for each species. Secondly, stacking the triplets of indices within the same feature vector leads to features that are discriminant for certain species, but that perform as noise for other ones, which can cause errors. Finally, the Hughes effect, even if it is less significant than for the spectral reflectance (Hughes, 1968), is comparatively more important than in the ensemble method (only three indices per classifier). More the number of features is large, more these last two phenomena are present. In order to improve the proposed method, additional classifiers could be designed for separating groups of classes (leafy, coniferous, genus, etc.). On the one hand, it is assumed in this

study that there is at least one efficient classifier per species, which is partially verified as the F-score obtained of each species in not equal to 100%. It is likely that more there will be species, more this assumption may be invalid. Let's focus on a classification problem involving four species, s_1, s_2, s_3 and s_4. If there exists an efficient classifier only for s_1 and s_2, and another which is able to predict if the species is either s_1 or s_4, or s_2 or s_3, it is sufficient to predict the species unambiguously. On the other hand, even if there exists an efficient classifier for each species, additional ones are of interest for the robustness of the approach. For example, if VNIR-based features are found to be able to discriminate only two or three groups of spectral-similar species, a classifier dedicated to the identification of such group could be derived. In both cases, a more complex decision rule would have to be defined through a logical reasoning.

The field measurements considered in this study can be used for the classification of tree species within airborne images. While the majority of the state-of-art approaches consider training samples directly from the images (Fassnacht et al., 2016), our research demonstrates that encouraging results can be obtained based on leaf level spectroscopic measurements. Canopy level samples lead also to encouraging accuracies. Still focusing on tree species classification, consistent results have been found in the thermal infrared context based on laboratory measurements (da Luz and Crowley, 2010), even if it is difficult to compare our findings. There are three reasons to explain why the spectral indices are more efficient than the spectral reflectance. With more than 100 indices, the richness of the parameters used in this research is substantial. Secondly, the leaf level spectral measurements are not directly comparable to the target samples which are pixels of the images. This is due to the canopy structure, especially the LAI and the LAD, which modifies the radiation through the vegetation volumes (Roberts et al., 2004). The comparison of the canopy and airborne samples is easier but the conditions of the acquisitions (spatial resolution, solar angle, atmospheric composition, etc.) can cause significant differences among these signals. Therefore, the field spectral measurements at the canopy level are not representative of the image pixels. Thirdly, the reflectance is much more sensitive to the Hughes effect (Hughes, 1968) as its dimension is much higher. In addition, the species-specific feature extraction carried out in the proposed method is an essential step in order to identify the best indices. Indeed, the triplets of indices have been chosen for optimizing the target accuracy, i.e. that at the airborne level, through a simple transfer approach. The transfer carried out in this study is however a limit of our approach. Indeed, it requires labelled airborne data, whereas such information is often not available. Transfer learning techniques could be considered for dealing with this issue (Tuia et al., 2016). Also, more training samples could improve the performance of the proposed approach. Radiative Transfer (RT) models such as (Gastellu-Etchegorry et al., 2004; Jacquemoud et al., 2009) could be considered

to simulate representative airborne training samples based on leaf level measurements.

Urban tree species classification from spectral and contextual features

Based on the results of the first chapter, the third chapter of this thesis aims at detecting the trees that belong to urban alignments. Three study areas of the city of Toulouse, France, are considered. While hyperspectral and DSM data are used in order to detect high vegetation, the road network available from GIS data allows the pixels close to the streets to be identified. The resulting mask (vegetation, height and distance) is then processed through a MPP approach. The interaction energy of the process models the contextual features of the street trees, i.e. a small angle between the individuals and similar heights. An energy minimization is carried out with a Multiple Birth and Death (MBD) scheme, and allows street tree maps to be estimated for the three study cases. The second case is challenging because of the significant overlap between the street trees crowns, while the third one is particularly difficult owing to the high spatial connection between the alignment trees and the other ones.

The results of this chapter show that the street trees can be mapped with high accuracies thanks to the proposed MPP approach, going further than the recent method of (Wen et al., 2017) which aims at classifying the urban canopies (patch-level classification) in three classes (park, roadside and residential-institutional canopies). Indeed, we propose an individual detection which is robust to the spatial connection between the street trees and the other populations of trees. However, while the proposed method success in identifying non overlapping street trees, declines are observed for cases of significant overlap such as in the second study area. This behaviour corresponds to a well known issue (Zhen et al., 2016; Alonzo et al., 2014), but our simple modelling of the individual trees (only based on the binary values of the mask) is particularly sensitive to overlap. There is thus an interest in more accurately modelling the trees, by using their 3D structure for example. Nowadays, efficient tree crown delineation approaches are based on LiDAR data (Gupta et al., 2010), the data energy of the MPP could therefore be inspired from these methods. A DSM-based data term could be already efficient for improving the performance in such cases of overlap. Winter acquisitions could also be of interest. On the other hand, the contextual features used in this thesis for modelling the street trees are simple, the angle between the trees (assumed to be small) and the heights of the trees (assumed to be similar), and allow to reach high performance. The GIS data are essential only for the third study area. Nevertheless, some errors remain, in particular for the second and third study

areas. This could be improved thanks to a more comprehensive modelling of the trees that belong to alignments. For example, the street trees are often equidistant, a characteristic that can be included in the interaction term of the MPP. Moreover, they have the same shape due to pruning, thus a constraint on the radius is another possible solution. In addition to be close to the roads, the street trees are parallel to these roads, thus a dot product between the street vectors and the tree triplet vectors could be computed and used in the MPP energy. The species can also be used as the alignments are often mono-specific or bi-specific.

This thesis demonstrates that the contextual modelling is obviously interesting for the extraction of the street trees, but also in order to improve the performance of existing methods (tree species classification as well as tree crown delineation approaches in particular). Indeed, the results of this chapter show that the post-regularization of the species predictions among the individuals of a mono-specific alignment is possible thanks to the alignment membership knowledge, accessible thanks to the street tree mapping proposed in this thesis. Still focusing on the species classification, the pruned trees of an alignment have specific spectral responses as they are particularly stressed. Knowing the alignment membership, it is possible to include this information in the classification framework in order to apply a specific processing for the identification of the street trees as well as for the park ones. Because the urban tree alignments are often mono-specific, this characteristic could also be used to enrich the species training sets (existing databases). In natural forests, other information could be used such as the fact that certain species can not live above a certain altitude, grows in a specific way, depending on sunshine, temperature, moisture, soil and other neighbourhood properties. On the other hand, the crown delineation approaches could benefit from a contextual modelling. The street trees can overlap as in the second study area, potentially in a more important way forming homogeneous canopies that cause the trees to be not discernible. A prior information about the position of the trees, through an equidistance assumption, is a solution for improving the tree detection in these conditions. Moreover, a constraint on the crown delineation shape could be used in order to avoid erroneous estimations such as certain obtained with the region growing approach used in this thesis. Also, the pixels within the tree crowns belong to the same species, thus the corresponding spectral variability is necessary low in comparison to the between-trees variability (except specific cases of diseases), which is another possible *a priori* information.

Conclusions

The objective of this thesis is to improve the current urban tree species classification approaches. The main contribution of the PhD work is the development of urban tree species mapping approaches that take advantage of the complementarity of several sources of information, from different airborne sensors to multiple spectral classifiers, by way of contextual information. Toulouse, France, is the study area and hyperspectral VNIR and SWIR, PAN and nDSM data are considered for that purpose. While the first part of the thesis is dedicated to the development of a fusion method, the second and third parts concern the design of an ensemble classifier approach and the detection of the trees that belong to urban alignments, respectively. The main conclusion resulting from the first chapter is that the VNIR hyperspectral data, with an overall accuracy of 75% for the reference site, are the main driver of the performance to classify 15 species, in comparison to the SWIR (69%), PAN (43%) and nDSM (35%) data. The best fusion strategy (decision level) slightly improves the performance of the VNIR with an overall accuracy value of 77% (+2pp). We conclude that, because there is little interest of including SWIR, PAN and nDSM data, the VNIR alone is sufficient in our context. Another conclusion is that the complementarity analysis of the sources subject to combination is essential for a good understanding of the behaviour of the fusion. Focusing on the second part of the work, the main result is that the proposed ensemble classifier leads to better results than the standard stacked approach when classifying 5 species, with an average overall accuracy of 60% (κ = 49%) instead of 58% (κ = 44%) among the three training samples levels. The resulting conclusion is that the proposed approach is not very sensitive to the change of scale. Regarding the third chapter, we concluded that the proposed street tree mapping method is functional for non overlapping crowns with an average detection rate of 85% among the three study areas under consideration. Another conclusion is that the angle and the height are discriminative features of the street trees, while the GIS data is only necessary for the third case.

Perspectives

At the end of this PhD thesis, there are still many limitations. In particular, some concern the input data of the methods, the methods, while others are about the context in which they are used.

Our methods suffer from the limited number of training samples available in this PhD study. This impacts the performance of the tree species classification in the study areas considered in this work. The proposed approaches can not be operational for other cities as the training samples used in our context are not representative of all the urban environ-

ments, owing to the great variety of species in a city changing according to geographical locations. Two approaches are possible to remove this constraint. One strategy is the modelling of representative training samples. The idea is to create a model able to simulate the spectral reflectance of each tree species, including its intra-class variability, for any configuration of acquisition. This kind of model exists for simulating canopy reflectance based on leaf level measurements and canopy structure parameters (LAI, LAD, etc.), Discrete Anisotropic Radiative Transfer (DART) (Gastellu-Etchegorry et al., 2004) for example. There are also models for simulating leaf reflectance based on foliar components contents, PROSPECT for instance (Jacquemoud and Baret, 1990). However, it is challenging to get the distribution of these parameters (LAI, chlorophyll, etc.) for each species. On the other hand, through a collaborative event (e.g. Mapathon), based on Google Earth and Google Street View for example, the second one aims at taking advantage of the power of people in order to get quickly many training samples. The difficulty to discriminate certain species by photointerpretation is the main drawback of that approach, but many trees could be delineated. Then, the combination of the two approaches may be of interest. For example, training samples from a Mapathon could be used to estimate the distribution of these parameters for each species based on a radiative transfer model, in order to generate more training samples in a second phase.

A second perspective of this thesis is the reflection on the design of the sensors. Indeed, without a suitable remote sensing system for extracting the information of interest, no information can be extracted, whatever the processing applied. It is a whole processing chain that has to be considered. Focusing on the tree species classification of this thesis, the usefulness of the textural features derived from PAN images (14 cm spatial resolution) is disappointing, as well as the interest of the structural ones based on a nDSM (spatial resolution of 12.5 cm). Whereas a finer spatial resolution seems to be more appropriate for the extraction of textural parameters, smaller than the size of the leaves (accessible thanks to UAV-based systems), LiDAR point cloud would give information about the within tree crown structure, more species-specific than the crown surface. The intensity of LiDAR is probably an interesting information, especially for characterizing the structure of the foliage. In addition, the community focuses mainly on nadir view acquisitions. Since the trees are long vertically, there is probably an interest in using oblique views, in particular in the urban environment where the tree stands are significantly distributed. Recent studies are focusing on the consideration of multispectral LiDAR for vegetation applications. It can be both interesting for tree crown delineation and species classification. Last but not least, hyperspectral data at high spatial resolution are rare, there is thus an interest in combining for instance airborne hyperspectral data (at one date) and multitemporal satellite data such as Sentinel-2, SPOT 6/7, etc. However, the spatial resolution of Sentinel-2 (10

m for the R, G, B and IR canals) does not allow the trees to be discerned. In the future, if several airborne hyperspectral images are available per year, there will probably be an interest in using them to improve species classification. UAV-based hyperspectral systems could be also of interest (spatial resolution around 10 cm). The proposed methods have to evolve in order to be practical for time series.

Another outlook is questioning of the current urban tree species mapping paradigm. The methods developed are parts of the standard urban tree species mapping paradigm: tree crown delineation followed by tree species classification. To our mind, this paradigm has not been thought of and results from the research of two communities. Actually, this has been natural to gather the existing delineation and classification algorithms for mapping the tree species. However, the urban tree species maps derived from this framework are still not perfect, due to the errors made through the delineation step as well as other ones occurring during the classification, either intrinsically or as a consequence of the imperfections of the delineation. In order to improve the urban tree species mapping approaches, one way is to refine both procedures separately as realized in the literature. It is also possible to reverse the two procedures and combine with the results of the traditional way. Another strategy would be to adopt an unified framework, similarly to what has been done in the detection of street trees, in particular because the estimation of the tree crown benefits from the species knowledge, but also as the shape of the crowns is related to the species. While hyperspectral data are often used for tree species classification, LiDAR technology is particularly considered for tree crown extraction. Why not use hyperspectral data and thus the species knowledge for delineation? The MPP framework is a candidate of interest for developing such method.

The comprehensive modelling of the context is another subject of interest. Pixel-based at the origin, the current paradigm of the tree species mapping in the remote sensing community tends to be more and more object-based, because of the spatial resolution getting finer, but essentially because the trees are intrinsically objects with particular features that can be considered for improving the standard pixel-based approaches. While it is natural to consider objects, it is *a priori* not appropriate to develop pixel-based approaches. The pixel is only a sensor-based characteristic. The development of object-based frameworks is then not questionable, either focusing on tree species mapping or other applications. However, the object is still an approximation. Whereas the objects are considered as isolated entities in the majority of the cases, which is reasonable if we talk about anomaly for example, the trees are often connected each others, either in the urban environment with the alignment individuals, but also in forests. The objects live in a particular context, trees, crops, buildings, animals, humans, etc. Expert knowledge about this context is underuti-

lized, whereas it could improve the existing schemes via an understanding of the main phenomena involved. After pixel-based and object-based frameworks, object structure-based ones, e.g. considering a group of trees, have to be explored as it will give a more accurate modelling of the scenes.

The ultimate goal remains the health monitoring of the trees. Once the trees are delineated and their species estimated, they can be monitored in order to detect potential diseases or weaknesses due to the environment, and predict the eventual spread in the case of diseases. In addition to the case of diseases, the urban trees can be subject to hard conditions, in particular during heatwave or atmospheric pollution conditions. In a first time, the idea can be to focus on a particular case, for example the Canker stain of *Platanus* which is the consequence of a fungus: *Ceratocystis platani*. Whereas training samples can be obtained for tree species classification if enough effort is made, it is more difficult to get examples of sick trees, knowing that the state of health can vary a lot among the individuals affected by the disease. A first strategy consists in acquiring a strong expert knowledge about the effect of the disease on the remote sensing signals, and include it in a method. Such knowledge is probably difficult to learn, but allows a good understanding of what happens from the data to the extraction of information. Another approach is the consideration of deep learning techniques, by considering that we have many examples of health cases, and diseases ones can be detected as anomalies. Secondly, the prediction of the evolution of the problematic state of health is of particular interest, in order to prevent damages, based on multitemporal data. A satellite mission with a very rapid-revisit could of interest. Is it possible to prevent the problematic case before it happens?

Appendices

Appendices of chapter I

Tree reference map

The tree reference map is a key component of this thesis. It has been designed thanks to the QGIS software as illustrated in Figure 3.15. QGIS is a powerful tool which allows each tree to be inventoried through its delineation (polygon) and attributes, in a *shapefile* easily handled with the GDAL python module. In the example of Figure 3.15, the attributes concern for example the species of the trees (*species*), the potential field measurements that have been carried out (*volume*, *leaf*), the accuracy of the species classification method (*eSpecies*), or the values of specific vegetation indices (*index1*). Focusing on Figure 3.15, right, the selected yellow tree is dark brown, which indicates that its value of PSRI is particularly high. The attribute table tells that this is an *Aesculus hippocastanum* tree and ground truth shows that this tree is attacked by the horse chestnut leaf miner, which is consistent.

Confusion matrices

The confusion matrix is a powerful complementary tool for assessing the performance of classification methods. The Figures 3.16, 3.17 and 3.18 show the results of the mono-source and multi-source classifications for the reference area. These confusion matrices are consistent with the Table I.5 in the sense that the hyperspectral data are the main driver of the classification accuracy while the textural anbd structural features contribute marginally. The additional information that can be extracted from the confusion matrices is, as indicated by its name, the confusion between species. For example, the confusion matrices of the VNIR (Figure 3.16, top) reveals that there is a significant confusion between *Acer platanoides* and *Liquidambar styraciflua*, which can be explained by similar phenological behaviours during autumn. Moreover, there is a confusion between *Taxus baccata* and *Cedrus deodara* by the SWIR (Figure 3.16, bottom), which can be expected

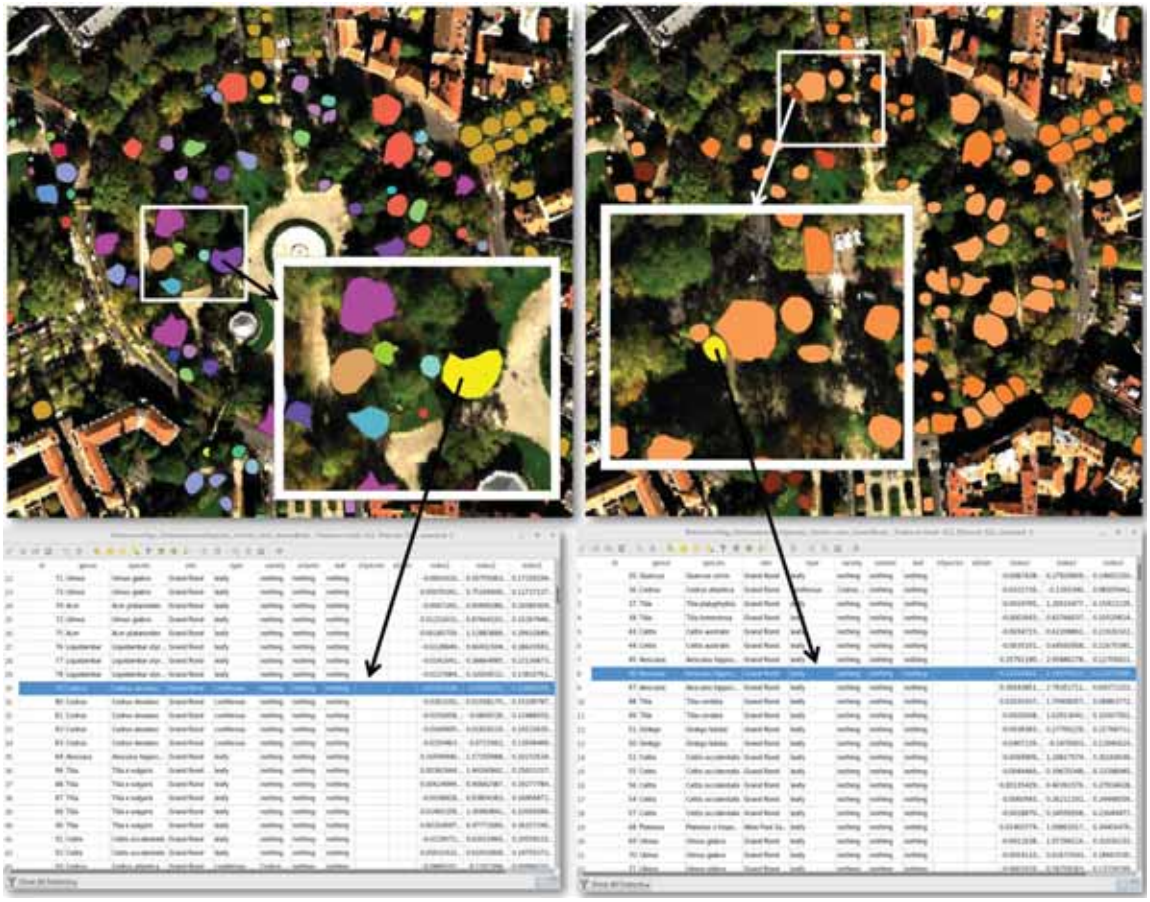

Figure 3.15: Illustration of the tree reference map designed with QGIS. The polygons refer to the delineations of the inventoried trees in the reference site. Top left: each color corresponds to a particular species. Top right: each color indicates a specific level of PSRI vegetation index (*index1*). Whereas light brown trees have a low value of PSRI, dark brown ones have an high value. Bottom: corresponding attribute tables with the yellow selected tree highlighted in blue.

as these trees are two coniferous. On the other hand, the confusion matrix of the decision level fusion (Figure 3.18, top) show that the fusion improves the confusion with *Tilia tomentosa* as each source confuses this species. This result is consistent with the Figure I.8 (application on the test site).

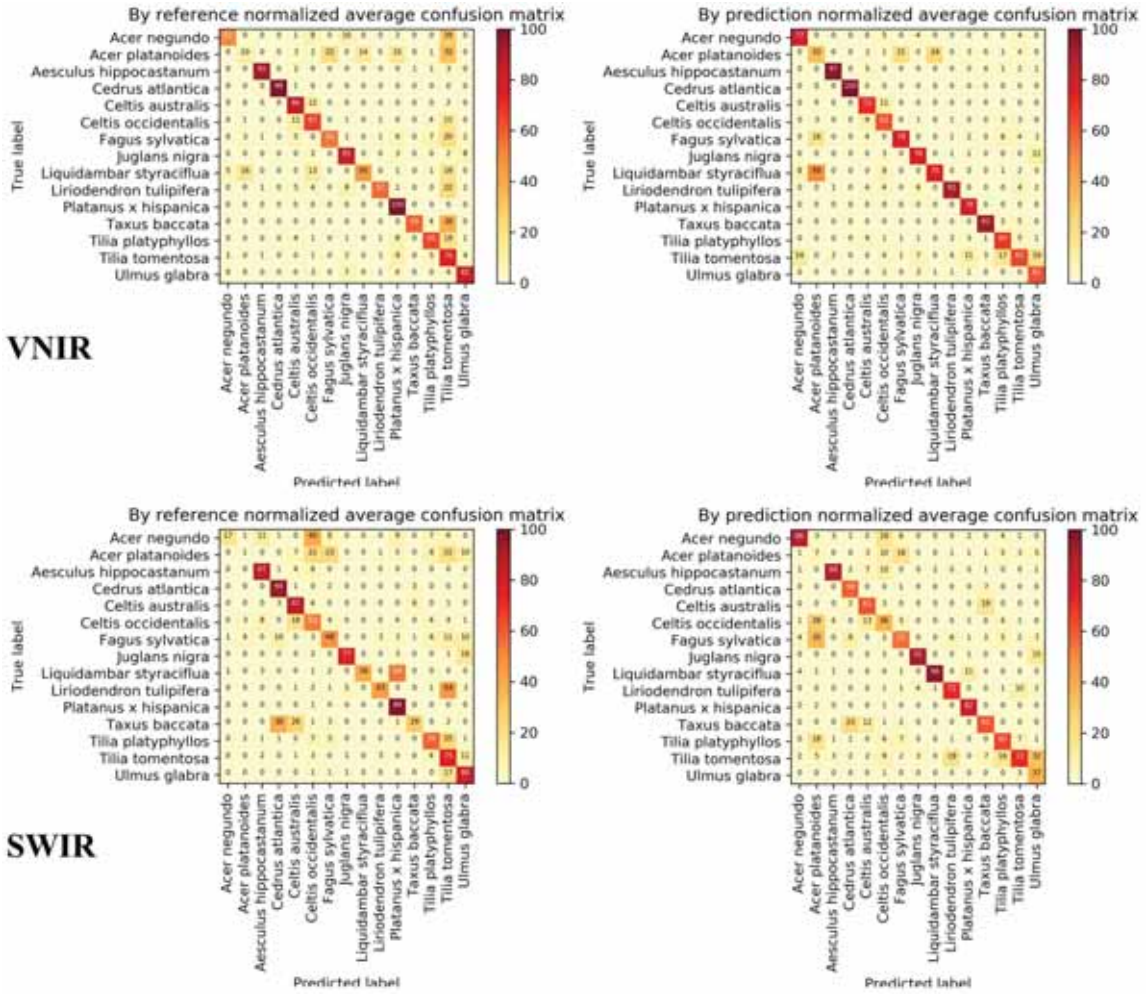

Figure 3.16: Reference site confusion matrices for the VNIR and SWIR MNF sources (%). The prediction is per column.

Appendix of chapter II

Field campaign

The trees that were subject to field measurements are showed in Figure 5. They are planted in the ”Jardin des Plantes” (botanic garden) in Toulouse. In particular, the necrotic leaves

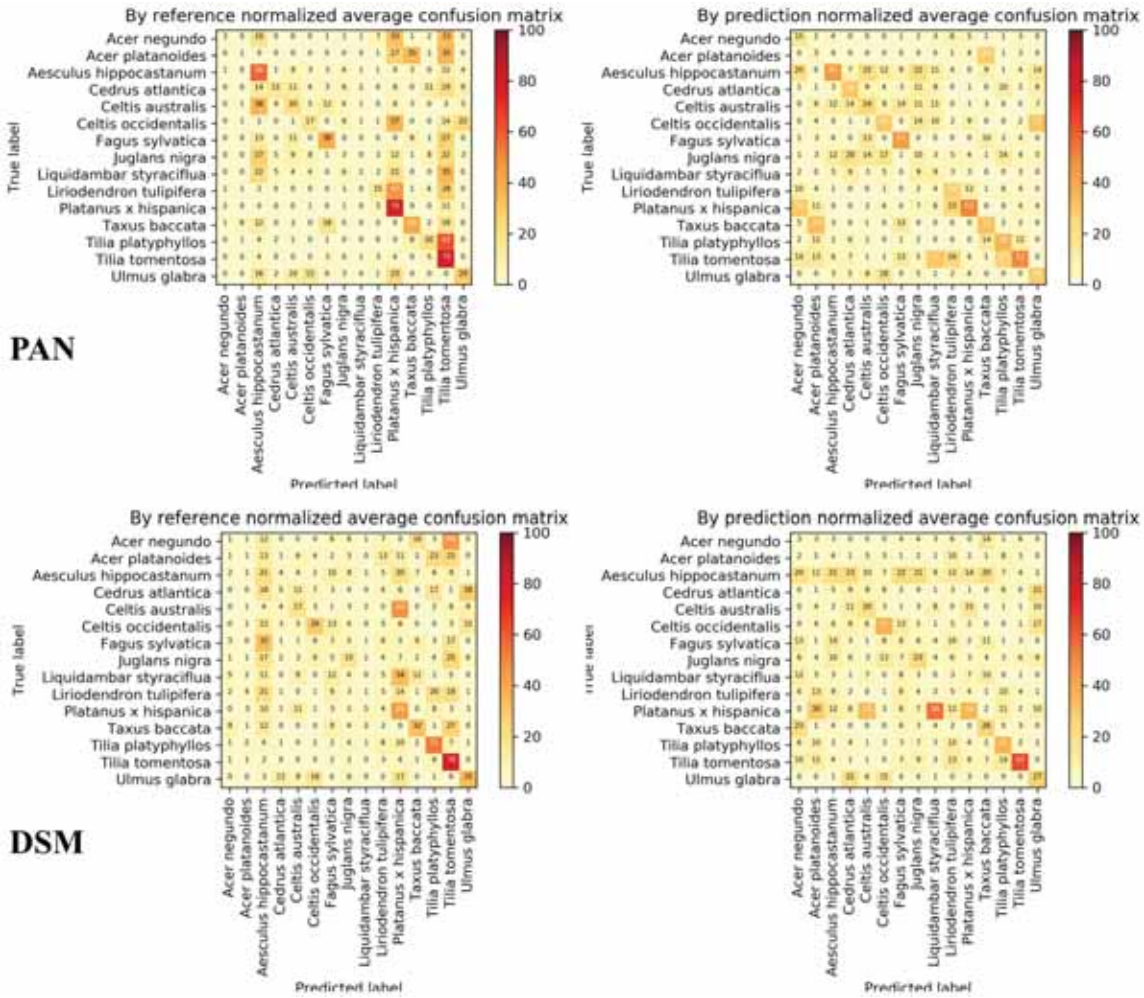

Figure 3.17: Reference site confusion matrices for the PAN and nDSM sources with four sub-objects and object scale, respectively (%). The prediction is per column.

of *Aesculus hippocastanum* are easily identifiable. However, the leaves of *Fagus sylvatica* are not purple as the content of anthocyanin is very low in that individual. The Figure 3.20 illustrates the acquisition of the canopy level measurements thanks to a cherry picker of Toulouse.

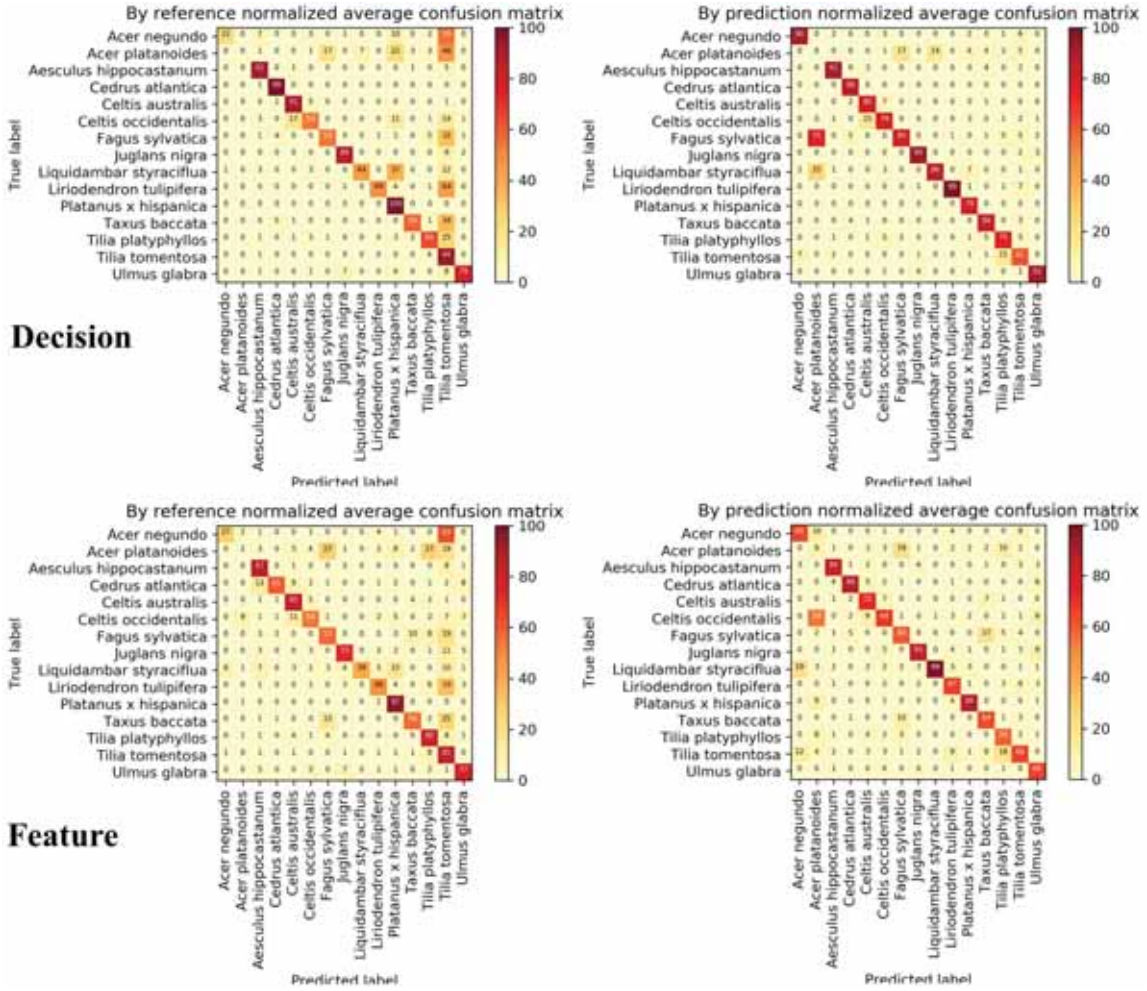

Figure 3.18: Reference site confusion matrices for the decision and feature level fusions (%). The prediction is per column.

Additional results

Figure 3.21 compares the confusion matrices of the ensemble and stack approaches in the case of leaf level training samples. In particular, the proposed ensemble method allows *Tilia tomentosa* to be less confused thanks to the rejected class. For the other species, the results are similar. In order to fuel the discussion, the Figure 3.22 illustrates the values of the spectral indices of *Juglans nigra*, as well as the associated spectral reflectances. In

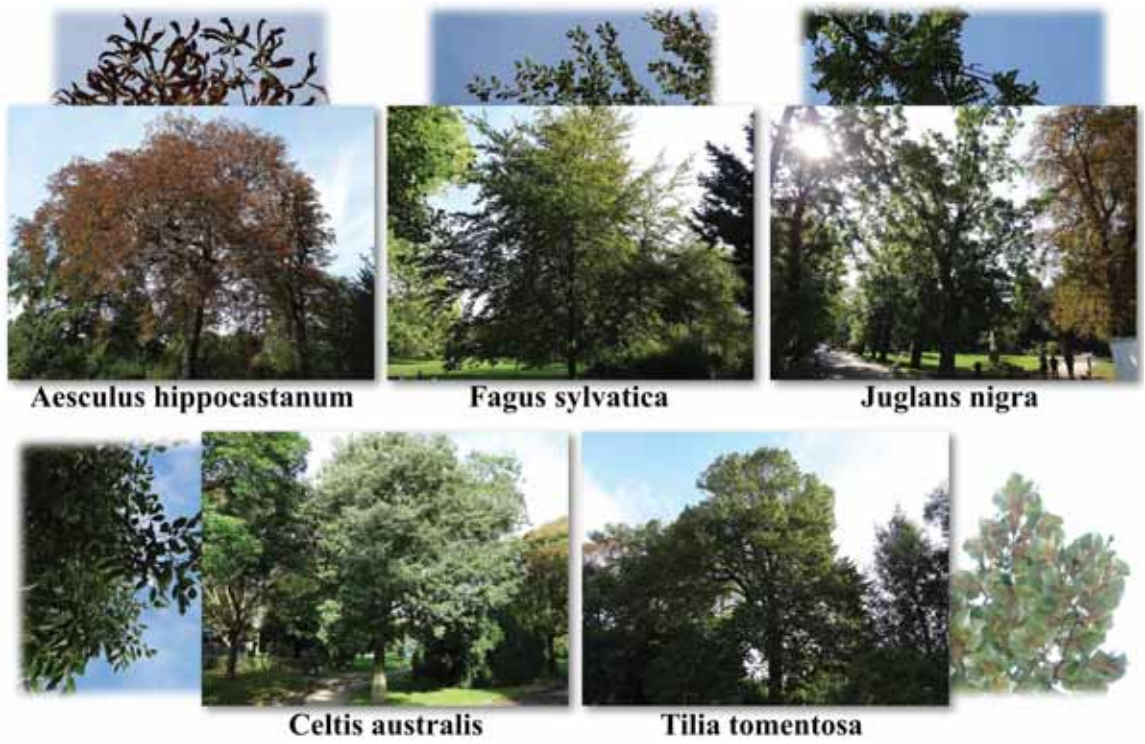

Figure 3.19: Trees subject to field measurements.

particular, the Greeness Index (GI) discriminates this species compared to the others. This is explained by a strong absorption of radiation around 677 nm.

Appendix of chapter III

Features of the street trees

The discussions of the chapter III highlight the interest of including other contextual features such as the distance between the street trees. This distance is expected to be the same within an alignment whereas its value is expected to change from an alignment to another. The Figure 3.23 illustrates that contextual feature. From this histogram, it is reasonable to assume that the distance between the street trees lies within a small range of values. This information can be included in the MPP model through an improvement of the equation

III.4. However, other measurements demonstrate that we can not assume that there is a unique distance value for a city, even if recommendations are given by urban managers. Indeed, each site often requires a specific planning.

Complementary maps

The Figure III.9 highlights the contribution of the GIS data, angle and height in the street tree identification. For a complementary analysis, the Figures 3.24, 3.25 and 3.26 illustrate what happens from a spatial point of view. As visible thanks to the Figure III.9, the height information is more discriminative than the angle one. Maybe there is an interest in weighting differently these two features.

List of scientific productions

Peer-reviewed publications in international journals

Accepted

J. Aval, J. Demuynck, E. Zenou, S. Fabre, D. Sheeren, M. Fauvel, K. Adeline and X. Briottet. Detection of individual trees in urban alignment from airborne data and contextual information: a marked point process approach. ISPRS Journal of Photogrammetry and Remote Sensing, 2018.

J. Aval, S. Fabre, E. Zenou, D. Sheeren, M. Fauvel and X. Briottet. Object-based fusion for urban tree species classification from heterogeneous data: hyperspectral, panchromatic and nDSM. International Journal of Remote Sensing, 2018.

Under review

J. Aval, S. Fabre, E. Zenou, D. Sheeren, M. Fauvel and X. Briottet. An ensemble classifier approach for urban tree species classification from ground-based spectral references. Remote Sensing, 2018.

Communications

Oral communications

J. Aval, S. Fabre, E. Zenou, D. Sheeren, M. Fauvel and X. Briottet. Individual street tree detection from airborne data and contextual information. GEOBIA, Montpellier, France,

2018.

J. Aval, S. Fabre, E. Zenou, D. Sheeren, M. Fauvel and X. Briottet. Urban tree species classification with airborne hyperspectral VNIR and SWIR, PAN and DSM data. EARSel, Zurich, Swiss, 2017.

J. Aval, S. Fabre, E. Zenou, D. Sheeren, M. Fauvel and X. Briottet. Classification des espèces d'arbres en milieu urbain par fusion de données aéroportées hyperspectrales, panchromatiques et d'un MNS. TEMU, Toulouse, France, 2017.

Posters

J. Aval, S. Fabre, E. Zenou, D. Sheeren, M. Fauvel and X. Briottet. An ensemble classifier approach for urban tree species classification from ground-based spectral references. ForestSAT, Washington, USA, 2018.

J. Aval, S. Fabre, E. Zenou, D. Sheeren, M. Fauvel and X. Briottet. Urban tree species classification with airborne hyperspectral VNIR and SWIR, PAN and DSM data by fusion at the object level. IAC, Adelaide, Australia, 2017.

J. Aval, S. Fabre, E. Zenou, D. Sheeren, M. Fauvel and X. Briottet. Étude du potentiel des données hyperspectrales en vue de développer une méthode de cartographie automatique du patrimoine arboré en milieu urbain. SFPT, Grenoble, France, 2016.

Book chapter

J. Aval and T. Erudel. Potentiel de la télédétection optique aéroportée pour la cartographie des piscines en milieu urbain.

In the following book:
N. Baghdadi, C. Mallet, M. Zribi (2018). QGIS et applications en aménagement du territoire (Vol. 3). ISTE Editions.

Figure 3.20: Illustration of the canopy level spectral measurements carried out thanks to a cherry picker of Toulouse city.

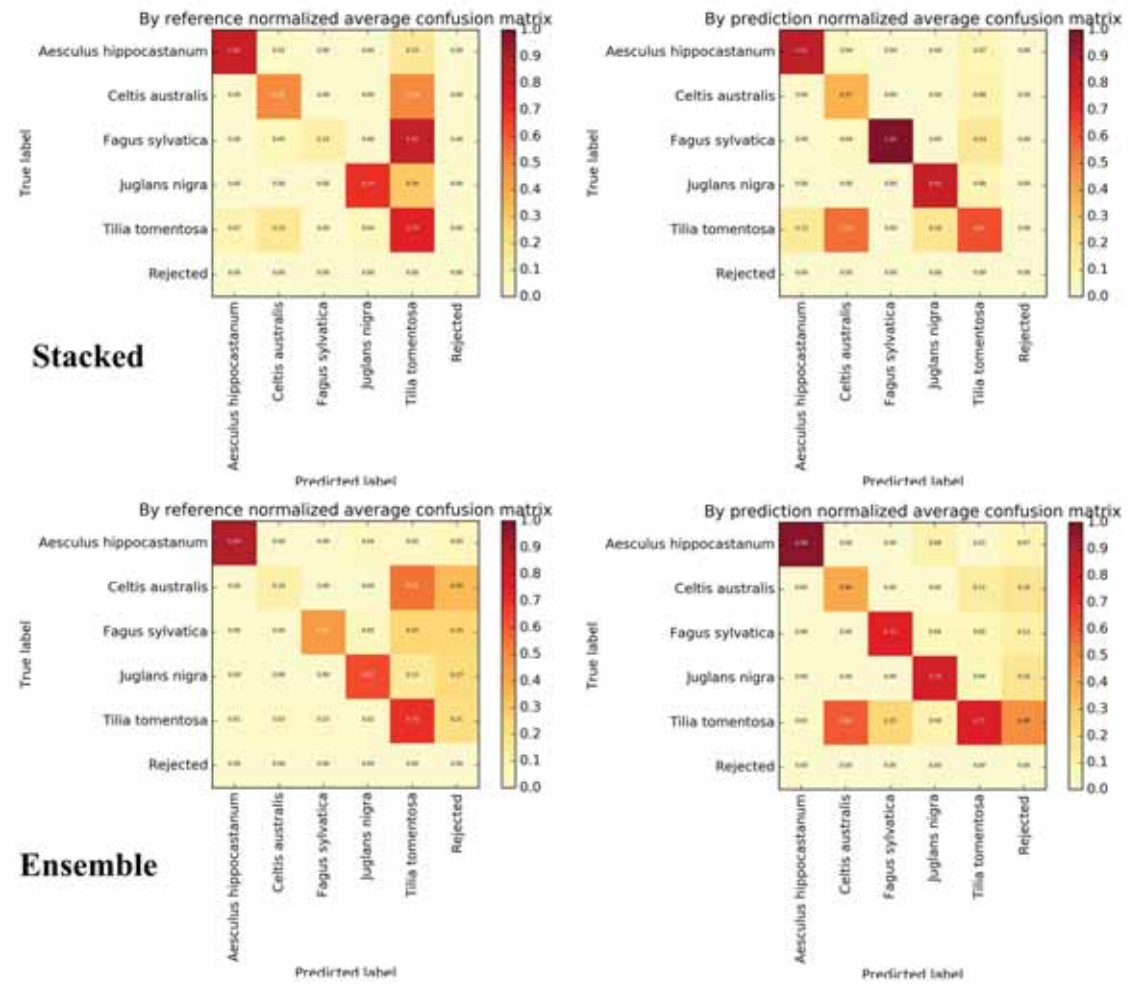

Figure 3.21: Comparison of the confusions matrices of the ensemble and stacked approcahes in the case of canopy level spectral data.

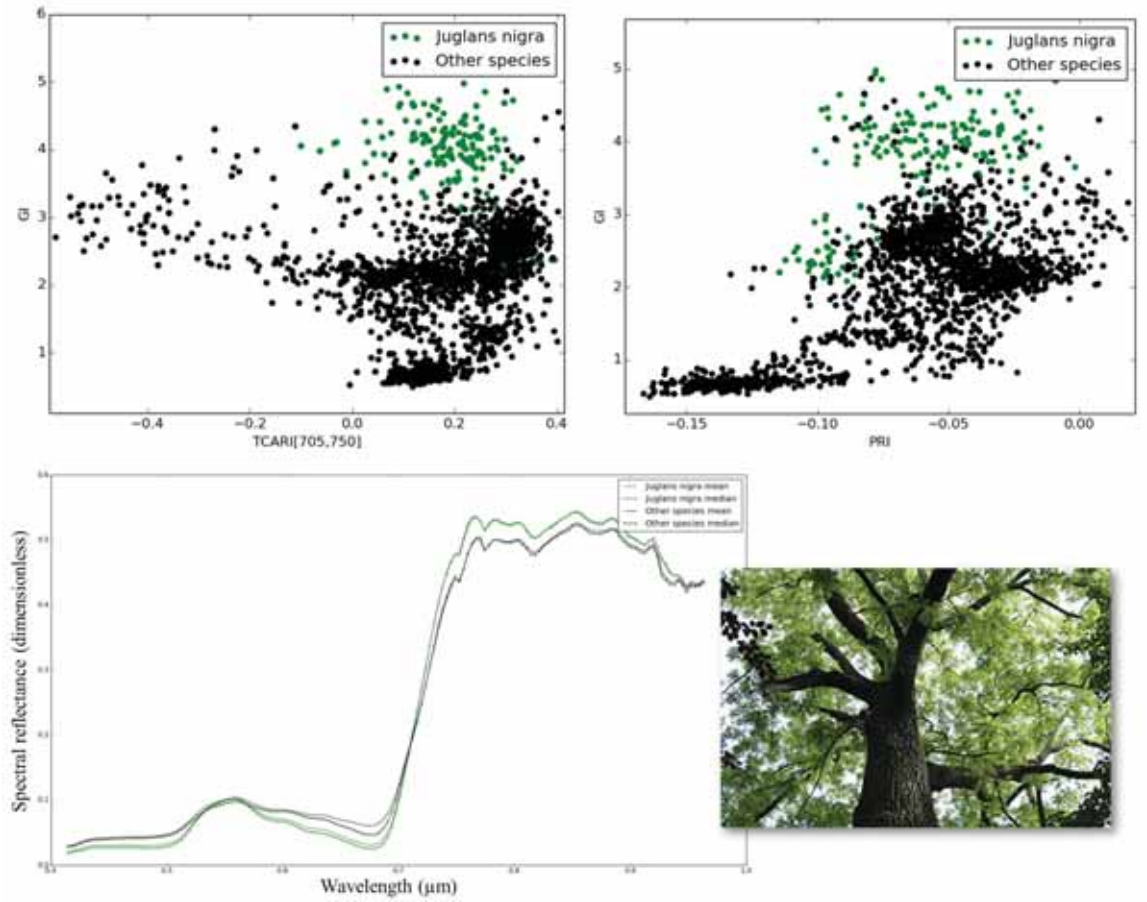

Figure 3.22: Link between the spectral indices, the spectral signatures and the species for airborne level data (*Juglans nigra*). Top: Spectral indices space. Bottom: Spectral signatures space and an illustration of the species.

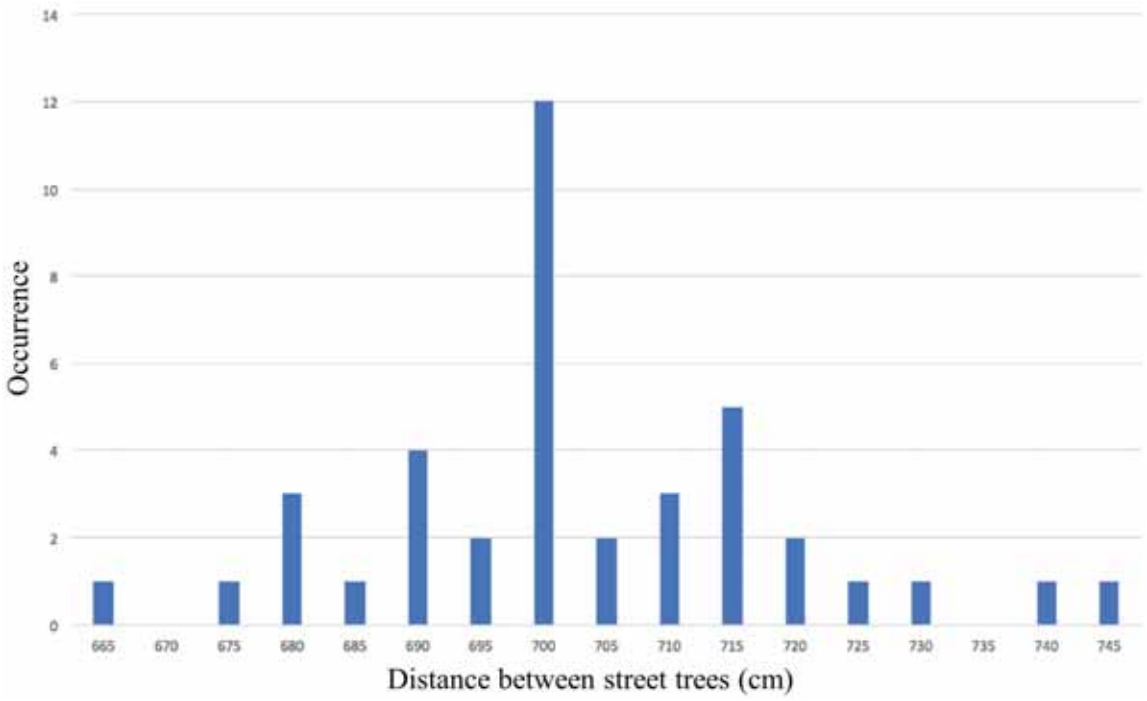

Figure 3.23: Occurrence of the distances between street trees from the urban alignments of *Allées Paul Sabatier*, Toulouse, France. Based on field measurements, the values are concentrated around 7 m. The average is approximately 7.02 m, with an uncertainty assumed to be $\pm$ 5 cm because of the position of the tree trunks that can not be exactly estimated.

Figure 3.24: Top: with all the features. Bottom: without GIS. Left: first study site. Right: second study site.

Figure 3.25: Top: without GIS and without the height. Bottom: without GIS and without the angle. Left: first study site. Right: second study site.

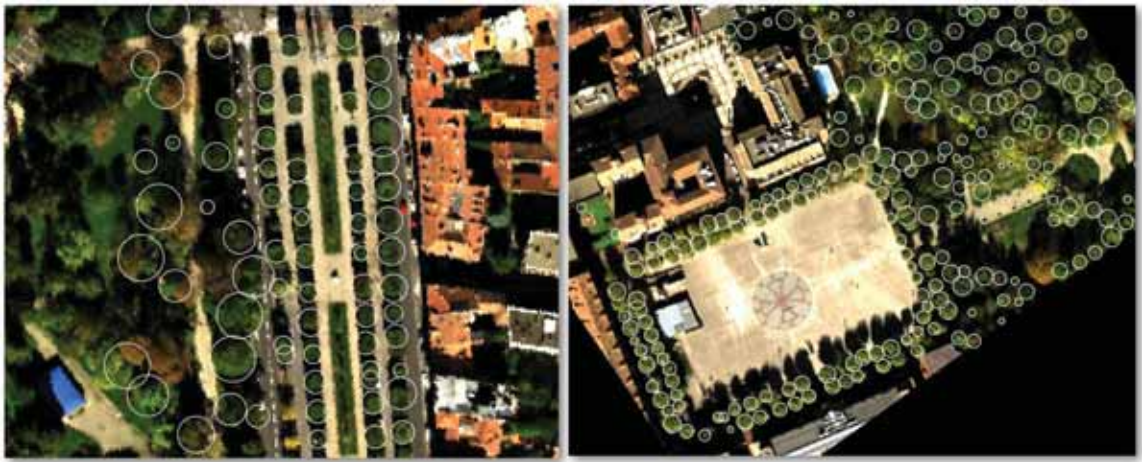

Figure 3.26: Without any feature. Left: first study site. Right: second study site.

Bibliography

[Abbasi et al. 2015] ABBASI, B ; AREFI, H ; BIGDELI, B ; MOTAGH, M ; ROESSNERB, S: FUSION OF HYPERSPECTRAL AND LIDAR DATA BASED ON DIMENSION REDUCTION AND MAXIMUM LIKELIHOOD. In: International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences (2015)

[Adeline 2014] ADELINE, Karine: Classification des matériaux urbains en présence de végétation éparse par télédétection hyperspectrale à haute résolution spatiale, Toulouse, ISAE, Dissertation, 2014

[Adeline et al. 2013] ADELINE, Karine R. ; LE BRIS, Arnaud ; COUBARD, Fabien ; BRIOTTET, Xavier ; PAPARODITIS, Nicolas ; VIALLEFONT, Françoise ; RIVIÈRE, Nicolas ; PAPELARD, Jean-Pierre ; DAVID, Nicolas ; DELIOT, Philippe et al.: Description de la campagne aéroportée UMBRA: étude de l'impact anthropique sur les écosystèmes urbains et naturels avec des images THR multispectrales et hyperspectrales. In: Revue française de photogrammétrie et de télédétection (2013), Nr. 202, S. 79–92

[ADEUS 2014] ADEUS: Les îLots de fraîcheur dans La viLLe - ADEUS (Agence de développement et d'urbanisme de l'agglomération strasbourgeoise). In: Les Notes de l'ADEUS 140 (2014)

[Alonzo et al. 2014] ALONZO, Michael ; BOOKHAGEN, Bodo ; ROBERTS, Dar A.: Urban tree species mapping using hyperspectral and lidar data fusion. In: Remote Sensing of Environment 148 (2014), S. 70–83

[Alonzo et al. 2013] ALONZO, Mike ; ROTH, Keely ; ROBERTS, Dar: Identifying Santa Barbara's urban tree species from AVIRIS imagery using canonical discriminant analysis. In: Remote Sensing Letters 4 (2013), Nr. 5, S. 513–521

[Alvarez-Uria and Körner 2007] ALVAREZ-URIA, P ; KÖRNER, Ch: Low temperature

limits of root growth in deciduous and evergreen temperate tree species. In: Functional ecology 21 (2007), Nr. 2, S. 211–218

[Ardila et al. 2012] ARDILA, Juan P. ; BIJKER, Wietske ; TOLPEKIN, Valentyn A. ; STEIN, Alfred: Context-sensitive extraction of tree crown objects in urban areas using VHR satellite images. In: International Journal of Applied Earth Observation and Geoinformation 15 (2012), S. 57–69

[Ardila et al. 2011] ARDILA, Juan P. ; TOLPEKIN, Valentyn A. ; BIJKER, Wietske ; STEIN, Alfred: Markov-random-field-based super-resolution mapping for identification of urban trees in VHR images. In: ISPRS journal of photogrammetry and remote sensing 66 (2011), Nr. 6, S. 762–775

[Asrar et al. 1984] ASRAR, GQ ; FUCHS, M ; KANEMASU, ET ; HATFIELD, JL: Estimating Absorbed Photosynthetic Radiation and Leaf Area Index from Spectral Reflectance in Wheat 1. In: Agronomy journal 76 (1984), Nr. 2, S. 300–306

[Aval 2018] AVAL: Object-based fusion for urban tree species classification from heterogeneous data: hyperspectral, panchromatic and DSM. In: IJRS (2018)

[Baatz and Schape 2000] BAATZ, M ; SCHAPE, A: Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. (2000), 01, S. 12–23

[Bailey et al. 2002] BAILEY, S-A ; HAINES-YOUNG, RH ; WATKINS, C: Species presence in fragmented landscapes: modelling of species requirements at the national level. In: Biological Conservation 108 (2002), Nr. 3, S. 307–316

[Baraton 2014] BARATON, Alain: Mes trucs et astuces de jardinier. Flammarion, 2014

[Blaschke 2010] BLASCHKE, T.: Object based image analysis for remote sensing. In: ISPRS Journal of Photogrammetry and Remote Sensing 65 (2010), Nr. 1, S. 2 – 16. – URL http://www.sciencedirect.com/science/article/pii/S0924271609000884. – ISSN 0924-2716

[Bolund and Hunhammar 1999] BOLUND, Per ; HUNHAMMAR, Sven: Ecosystem services in urban areas. In: Ecological economics 29 (1999), Nr. 2, S. 293–301

[Briottet et al. 2011] BRIOTTET, Xavier ; MARION, Rodolphe ; CARRERE, Véronique ; JACQUEMOUD, Stéphane ; CHEVREL, Stéphane ; PRASTAULT, Philippe ; D'ORIA, Marc ; GILOUPPE, Philippe ; HOSFORD, Steven ; LUBAC, Bertrand et al.: HYPXIM: A

new hyperspectral sensor combining Science/Defence applications. In: Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2011 3rd Workshop on IEEE (Veranst.), 2011, S. 1–4

[Ceamanos et al. 2010] CEAMANOS, Xavier ; WASKE, Björn ; BENEDIKTSSON, Jon A. ; CHANUSSOT, Jocelyn ; FAUVEL, Mathieu ; SVEINSSON, Johannes R.: A classifier ensemble based on fusion of support vector machines for classifying hyperspectral data. In: International Journal of Image and Data Fusion 1 (2010), Nr. 4, S. 293–307

[Chan and Yao 2008] CHAN, Chak K. ; YAO, Xiaohong: Air pollution in mega cities in China. In: Atmospheric environment 42 (2008), Nr. 1, S. 1–42

[Chanussot et al. 1999] CHANUSSOT, Jocelyn ; MAURIS, Gilles ; LAMBERT, Patrick: Fuzzy fusion techniques for linear features detection in multitemporal SAR images. In: IEEE Transactions on Geoscience and Remote Sensing 37 (1999), Nr. 3, S. 1292–1305

[Chatelain et al. 2009] CHATELAIN, Florent ; DESCOMBES, Xavier ; ZERUBIA, Josiane: Parameter estimation for marked point processes. application to object extraction from remote sensing images. In: International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition Springer (Veranst.), 2009, S. 221–234

[Chen et al. 2006] CHEN, Qi ; BALDOCCHI, Dennis ; GONG, Peng ; KELLY, Maggi: Isolating individual trees in a savanna woodland using small footprint lidar data. In: Photogrammetric Engineering & Remote Sensing 72 (2006), Nr. 8, S. 923–932

[Chiesura 2004] CHIESURA, Anna: The role of urban parks for the sustainable city. In: Landscape and urban planning 68 (2004), Nr. 1, S. 129–138

[Cho et al. 2008] CHO, MA ; SOBHAN, I ; SKIDMORE, AK ; LEEUW, J de: Discriminating species using hyperspectral indices at leaf and canopy scales. In: The International Archives of the Spatial Information Sciences (2008), S. 369–376

[Clark and Roberts 2012] CLARK, Matthew L. ; ROBERTS, Dar A.: Species-level differences in hyperspectral metrics among tropical rainforest trees as determined by a tree-based classifier. In: Remote Sensing 4 (2012), Nr. 6, S. 1820–1855

[Clark et al. 2005] CLARK, Matthew L. ; ROBERTS, Dar A. ; CLARK, David B.: Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. In: Remote sensing of environment 96 (2005), Nr. 3, S. 375–398

[Coburn and Roberts 2004] COBURN, CA ; ROBERTS, ACB: A multiscale texture analysis procedure for improved forest stand classification. In: International journal of remote sensing 25 (2004), Nr. 20, S. 4287–4308

[Corporation 2018] CORPORATION, Satellite I.: GeoEye-1 Satellite Sensor. https://www.satimagingcorp.com/satellite-sensors/geoeye-1/. 2018. – Accessed: 22/05/18

[Cunningham 2018] CUNNINGHAM, Margaret: Population Distribution: Rural vs. Urban Areas. https://study.com/academy/lesson/population-distribution-rural-vs-urban-areas.html. 2018. – Accessed: 18/05/18

[Dalponte et al. 2008] DALPONTE, Michele ; BRUZZONE, Lorenzo ; GIANELLE, Damiano: Fusion of hyperspectral and LIDAR remote sensing data for classification of complex forest areas. In: IEEE Transactions on Geoscience and Remote Sensing 46 (2008), Nr. 5, S. 1416–1427

[Dalponte et al. 2012] DALPONTE, Michele ; BRUZZONE, Lorenzo ; GIANELLE, Damiano: Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. In: Remote sensing of environment 123 (2012), S. 258–270

[Dalponte et al. 2009] DALPONTE, Michele ; BRUZZONE, Lorenzo ; VESCOVO, Loris ; GIANELLE, Damiano: The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas. In: Remote Sensing of Environment 113 (2009), Nr. 11, S. 2345–2355

[Dalponte et al. 2014] DALPONTE, Michele ; ØRKA, Hans O. ; ENE, Liviu T. ; GOBAKKEN, Terje ; NÆSSET, Erik: Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. In: Remote sensing of environment 140 (2014), S. 306–317

[Dalponte et al. 2015] DALPONTE, Michele ; REYES, Francesco ; KANDARE, Kaja ; GIANELLE, Damiano: Delineation of Individual Tree Crowns from ALS and Hyperspectral data: a comparison among four methods. In: European Journal of Remote Sensing 48 (2015), S. 365–382

[Datt 2000] DATT, Bisun: Recognition of eucalyptus forest species using hyperspectral reflectance data. In: Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International Bd. 4 IEEE (Veranst.), 2000, S. 1405–1407

[Descamps et al. 2009] DESCAMPS, Stig ; DESCOMBES, Xavier ; BECHET, Arnaud ; ZERUBIA, Josiane: Flamingo detection using marked point processes for estimating the size of populations. In: Traitement du Signal 26 (2009), Nr. 2, S. 95–108

[Descombes et al. 2009] DESCOMBES, Xavier ; MINLOS, Robert ; ZHIZHINA, Elena: Object extraction using a stochastic birth-and-death dynamics in continuum. In: Journal of Mathematical Imaging and Vision 33 (2009), Nr. 3, S. 347–359

[van Deventer et al. 2013] DEVENTER, Heidi van ; AZONG CHO, Moses ; MUTANGA, Onisimo ; NAIDOO, Laven ; DUDENI-TLHONE, Nontembeko: Reducing Leaf-Level Hyperspectral Data to 22 Components of Biochemical and Biophysical Bands Optimizes Tree Species Discrimination. (2013)

[Doick et al. 2014] DOICK, Kieron J. ; PEACE, Andrew ; HUTCHINGS, Tony R.: The role of one large greenspace in mitigating London's nocturnal urban heat island. In: Science of the total environment 493 (2014), S. 662–671

[Dépêche 2013] DÉPÊCHE, La: Le Palais de Justice. https://www.ladepeche.fr/diaporama/toulouse-vue-du-ciel-1/tse1.html. 2013. – Accessed: 27/07/18

[Dubitzky et al. 2007] DUBITZKY, Werner ; GRANZOW, Martin ; BERRAR, Daniel P.: Fundamentals of data mining in genomics and proteomics. Springer Science & Business Media, 2007

[Duchon 1977] DUCHON, Jean: Splines minimizing rotation-invariant semi-norms in Sobolev spaces. In: Constructive theory of functions of several variables. Springer, 1977, S. 85–100

[Ene et al. 2012] ENE, Liviu ; NÆSSET, Erik ; GOBAKKEN, Terje: Single tree detection in heterogeneous boreal forests using airborne laser scanning and area-based stem number estimates. In: International Journal of Remote Sensing 33 (2012), Nr. 16, S. 5171–5193

[Engler et al. 2013] ENGLER, Robin ; WASER, Lars T. ; ZIMMERMANN, Niklaus E. ; SCHAUB, Marcus ; BERDOS, Savvas ; GINZLER, Christian ; PSOMAS, Achilleas: Combining ensemble modeling and remote sensing for mapping individual tree species at high spatial resolution. In: Forest Ecology and Management 310 (2013), S. 64 – 73. – URL http://www.sciencedirect.com/science/article/pii/S0378112713005173. – ISSN 0378-1127

[EPA 2017] EPA, USA: Climate Change and Heat Islands. https://www.epa.gov/heat-islands/climate-change-and-heat-islands. 2017. – Accessed: 18/05/18

[Erudel et al. 2017] ERUDEL, Thierry ; FABRE, Sophie ; HOUET, Thomas ; MAZIER, Florence ; BRIOTTET, Xavier: Criteria Comparison for Classifying Peatland Vegetation Types Using In Situ Hyperspectral Measurements. In: Remote Sensing 9 (2017), Nr. 7, S. 748

[Fassnacht et al. 2014] FASSNACHT, Fabian E. ; NEUMANN, Carsten ; FORSTER, Michael ; BUDDENBAUM, Henning ; GHOSH, Aniruddha ; CLASEN, Anne ; JOSHI, Pawan K. ; KOCH, Barbara: Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central European test sites. In: Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 7 (2014), Nr. 6, S. 2547–2561

[Fassnacht et al. 2016] FASSNACHT, Fabian E. ; LATIFI, Hooman ; STEREŃCZAK, Krzysztof ; MODZELEWSKA, Aneta ; LEFSKY, Michael ; WASER, Lars T. ; STRAUB, Christoph ; GHOSH, Aniruddha: Review of studies on tree species classification from remotely sensed data. In: Remote Sensing of Environment 186 (2016), S. 64–87

[Fauvel et al. 2006] FAUVEL, Mathieu ; CHANUSSOT, Jocelyn ; BENEDIKTSSON, Jón Atli: Decision fusion for the classification of urban remote sensing images. In: IEEE Transactions on Geoscience and Remote Sensing 44 (2006), Nr. 10, S. 2828–2838

[Feng et al. 2015] FENG, Quanlong ; LIU, Jiantao ; GONG, Jianhua: Urban flood mapping based on unmanned aerial vehicle remote sensing and random forest classifier—A case of Yuyao, China. In: Water 7 (2015), Nr. 4, S. 1437–1455

[Féret et al. 2017] FÉRET, J-B ; GITELSON, AA ; NOBLE, SD ; JACQUEMOUD, S: PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. In: Remote Sensing of Environment 193 (2017), S. 204–215

[Féret and Asner 2013] FÉRET, Jean-Baptiste ; ASNER, Gregory P.: Tree species discrimination in tropical forests using airborne imaging spectroscopy. In: Geoscience and Remote Sensing, IEEE Transactions on 51 (2013), Nr. 1, S. 73–84

[Fini et al. 2015] FINI, A ; FRANGI, P ; FAORO, M ; PIATTI, R ; AMOROSO, G ; FERRINI, F: Effects of different pruning methods on an urban tree species: A four-year-experiment scaling down from the whole tree to the chloroplasts. In: Urban Forestry & Urban Greening 14 (2015), Nr. 3, S. 664–674

[Forestier et al. 2012] FORESTIER, G. ; PUISSANT, A. ; WEMMERT, C. ; GANÇARSKI, P.: Knowledge-based region labeling for remote sensing image interpretation. In: Computers, Environment and Urban Systems 36 (2012), Nr. 5,

S. 470 – 480. – URL http://www.sciencedirect.com/science/article/pii/S019897151200004X. – ISSN 0198-9715

[Franklin et al. 2000] FRANKLIN, S. E. ; HALL, R. J. ; MOSKAL, L. M. ; MAUDIE, A. J. ; LAVIGNE, M. B.: Incorporating texture into classification of forest species composition from airborne multispectral images. In: International Journal of Remote Sensing 21 (2000), Nr. 1, S. 61–79

[Gamal-Eldin et al. 2010] GAMAL-ELDIN, Ahmed ; DESCOMBES, Xavier ; ZERUBIA, Josiane: Multiple birth and cut algorithm for point process optimization. In: Signal-Image Technology and Internet-Based Systems (SITIS), 2010 Sixth International Conference on IEEE (Veranst.), 2010, S. 35–42

[Gamba et al. 2005] GAMBA, Paolo ; DELL'ACQUA, Fabio ; DASARATHY, Belur V.: Urban remote sensing using multiple data sets: Past, present, and future. In: Information Fusion 6 (2005), Nr. 4, S. 319–326

[Gastellu-Etchegorry et al. 2004] GASTELLU-ETCHEGORRY, JP ; MARTIN, E ; GASCON, F: DART: a 3D model for simulating satellite images and studying surface radiation budget. In: International journal of remote sensing 25 (2004), Nr. 1, S. 73–96

[Ghiyamat et al. 2013] GHIYAMAT, Azadeh ; SHAFRI, Helmi Zulhaidi M. ; MAHDIRAJI, Ghafour A. ; SHARIFF, Abdul Rashid M. ; MANSOR, Shattri: Hyperspectral discrimination of tree species with different classifications using single-and multiple-endmember. In: International Journal of Applied Earth Observation and Geoinformation 23 (2013), S. 177–191

[Ghosh et al. 2014a] GHOSH, Aniruddha ; FASSNACHT, Fabian E. ; JOSHI, PK ; KOCH, Barbara: A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. In: International Journal of Applied Earth Observation and Geoinformation 26 (2014), S. 49–63

[Ghosh et al. 2014b] GHOSH, Aniruddha ; FASSNACHT, Fabian E. ; JOSHI, PK ; KOCH, Barbara: A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. In: International Journal of Applied Earth Observation and Geoinformation 26 (2014), S. 49–63

[Ghosh and Joshi 2014] GHOSH, Aniruddha ; JOSHI, PK: A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery. In: International Journal of Applied Earth Observation and Geoinformation 26 (2014), S. 298–311

[Gillner et al. 2015] GILLNER, Sten ; VOGT, Juliane ; THARANG, Andreas ; DETTMANN, Sebastian ; ROLOFF, Andreas: Role of street trees in mitigating effects of heat and drought at highly sealed urban sites. In: Landscape and Urban Planning 143 (2015), S. 33–42

[Gitelson et al. 1996] GITELSON, Anatoly A. ; KAUFMAN, Yoram J. ; MERZLYAK, Mark N.: Use of a green channel in remote sensing of global vegetation from EOS-MODIS. In: Remote sensing of Environment 58 (1996), Nr. 3, S. 289–298

[Gitelson and Merzlyak 1996] GITELSON, Anatoly A. ; MERZLYAK, Mark N.: Signature analysis of leaf reflectance spectra: algorithm development for remote sensing of chlorophyll. In: Journal of plant physiology 148 (1996), Nr. 3-4, S. 494–500

[Gomesa and Maillarda 2014] GOMESA, Marılia F. ; MAILLARDA, Philippe: Integration of Marked Point Processes and Template Matching for the identification of individual tree crowns in an urban and a wooded savanna environment in Brazil. In: Proc. of SPIE Vol Bd. 9245, 2014, S. 92450X–1

[Gong et al. 2003] GONG, Peng ; PU, Ruiliang ; BIGING, Greg S. ; LARRIEU, Mirta R.: Estimation of forest leaf area index using vegetation indices derived from Hyperion hyperspectral data. In: IEEE transactions on geoscience and remote sensing 41 (2003), Nr. 6, S. 1355–1362

[Guo et al. 2016] GUO, Yuming ; ZENG, Hongmei ; ZHENG, Rongshou ; LI, Shanshan ; BARNETT, Adrian G. ; ZHANG, Siwei ; ZOU, Xiaonong ; HUXLEY, Rachel ; CHEN, Wanqing ; WILLIAMS, Gail: The association between lung cancer incidence and ambient air pollution in China: a spatiotemporal analysis. In: Environmental research 144 (2016), S. 60–65

[Gupta et al. 2010] GUPTA, Sandeep ; WEINACKER, Holger ; KOCH, Barbara: Comparative analysis of clustering-based approaches for 3-D single tree detection using airborne fullwave lidar data. In: Remote Sensing 2 (2010), Nr. 4, S. 968–989

[Hadj et al. 2010] HADJ, Saima B. ; CHATELAIN, Florent ; DESCOMBES, Xavier ; ZERUBIA, Josiane: Parameter estimation for a marked point process within a framework of multidimensional shape extraction from remote sensing images. In: ISPRS Technical Commission III Symposium on Photogrammetry Computer Vision and Image Analysis (PCV 2010), 2010, S. 1

[Haralick et al. 1973] HARALICK, Robert M. ; SHANMUGAM, Karthikeyan ; DINSTEIN, Its' H.: Textural features for image classification. In: Systems, Man and Cybernetics, IEEE Transactions on (1973), Nr. 6, S. 610–621

[Hauglin et al. 2014] HAUGLIN, Marius ; GOBAKKEN, Terje ; ASTRUP, Rasmus ; ENE, Liviu ; NÆSSET, Erik: Estimating single-tree crown biomass of Norway spruce by airborne laser scanning: a comparison of methods with and without the use of terrestrial laser scanning to obtain the ground reference data. In: Forests 5 (2014), Nr. 3, S. 384–403

[Heaviside et al. 2016] HEAVISIDE, Clare ; VARDOULAKIS, Sotiris ; CAI, Xiao-Ming: Attribution of mortality to the urban heat island during heatwaves in the West Midlands, UK. In: Environmental Health 15 (2016), Nr. 1, S. S27

[Heinzel et al. 2011] HEINZEL, Johannes N. ; WEINACKER, Holger ; KOCH, Barbara: Prior-knowledge-based single-tree extraction. In: International journal of remote sensing 32 (2011), Nr. 17, S. 4999–5020

[Hester et al. 2008] HESTER, David B. ; CAKIR, Halil I. ; NELSON, Stacy A. ; KHORRAM, Siamak: Per-pixel classification of high spatial resolution satellite imagery for urban land-cover mapping. In: Photogrammetric Engineering & Remote Sensing 74 (2008), Nr. 4, S. 463–471

[Holmgren et al. 2008] HOLMGREN, Johan ; PERSSON, Åsa ; SÖDERMAN, Ulf: Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. In: International Journal of Remote Sensing 29 (2008), Nr. 5, S. 1537–1552

[Horvath et al. 2009] HORVATH, Peter ; JERMYN, Ian H. ; KATO, Zoltan ; ZERUBIA, Josiane: A higher-order active contour model of a 'gas of circles' and its application to tree crown extraction. In: Pattern Recognition 42 (2009), Nr. 5, S. 699–709

[Hughes 1968] HUGHES, Gordon: On the mean accuracy of statistical pattern recognizers. In: IEEE transactions on information theory 14 (1968), Nr. 1, S. 55–63

[INFO 2018] INFO, LAND: LAND INFO High-Resolution Global Satellite Imagery. `http://www.landinfo.com/products_satellite.htm`. 2018. – Accessed: 22/05/18

[Iovan et al. 2008] IOVAN, Corina ; BOLDO, Didier ; CORD, Matthieu: MODELISATION DE LA VEGETATION EN MILIEU URBAIN: DETECTION ET CARACTERISATION A PARTIR D'IMAGES AERIENNES HAUTE RESOLUTION COULEUR ET INFRAROUGE. In: Revue française de photogrammétrie et de télédétection (2008), Nr. 189, S. 17–27

[Jacquemoud and Baret 1990] JACQUEMOUD, S ; BARET, F: PROSPECT: A model of leaf optical properties spectra. In: Remote sensing of environment 34 (1990), Nr. 2, S. 75–91

[Jacquemoud et al. 2009] JACQUEMOUD, Stéphane ; VERHOEF, Wout ; BARET, Frédéric ; BACOUR, Cédric ; ZARCO-TEJADA, Pablo J. ; ASNER, Gregory P. ; FRANÇOIS, Christophe ; USTIN, Susan L.: PROSPECT+ SAIL models: A review of use for vegetation characterization. In: Remote sensing of environment 113 (2009), S. S56–S66

[Jardinier 2016] JARDINIER, Blog: Mineuse du marronnier : petite chenille, gros dégâts ! http://www.jardiniers-professionnels.fr/mineuse-du-marronnier-petite-chenille-gros-degats/. 2016. – Accessed: 27/07/18

[Jensen and Cowen 1999] JENSEN, John R. ; COWEN, Dave C.: Remote sensing of urban/suburban infrastructure and socio-economic attributes. In: Photogrammetric engineering and remote sensing 65 (1999), S. 611–622

[Johansen and Phinn 2006] JOHANSEN, Kasper ; PHINN, Stuart: Mapping structural parameters and species composition of riparian vegetation using IKONOS and Landsat ETM+ data in Australian tropical savannahs. In: Photogrammetric Engineering & Remote Sensing 72 (2006), Nr. 1, S. 71–80

[Jones 2014] JONES, Owain: 8 (Urban) Places of Trees. In: Urban Forests, Trees, and Greenspace: A Political Ecology Perspective (2014), S. 111

[Jones et al. 2010] JONES, Trevor G. ; COOPS, Nicholas C. ; SHARMA, Tara: Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. In: Remote Sensing of Environment 114 (2010), Nr. 12, S. 2841–2852

[Kass et al. 1988] KASS, Michael ; WITKIN, Andrew ; TERZOPOULOS, Demetri: Snakes: Active contour models. In: International journal of computer vision 1 (1988), Nr. 4, S. 321–331

[Ke et al. 2010] KE, Yinghai ; ZHANG, Wenhua ; QUACKENBUSH, Lindi J.: Active contour and hill climbing for tree crown detection and delineation. In: Photogrammetric Engineering & Remote Sensing 76 (2010), Nr. 10, S. 1169–1181

[Koch et al. 2014] KOCH, Barbara ; KATTENBORN, Teja ; STRAUB, Christoph ; VAUHKONEN, Jari: Segmentation of forest to tree objects. In: Forestry Applications of Airborne Laser Scanning. Springer, 2014, S. 89–112

[Köhler 2016] KÖHLER, Claas H.: Airborne Imaging Spectrometer HySpex. In: Journal of large-scale research facilities JLSRF 2 (2016), Nr. A93, S. 1–6

[Kuncheva 2004] KUNCHEVA, Ludmila I.: Combining pattern classifiers: methods and algorithms. John Wiley & Sons, 2004

[Larsen et al. 2011] LARSEN, Morten ; ERIKSSON, Mats ; DESCOMBES, Xavier ; PERRIN, Guillaume ; BRANDTBERG, Tomas ; GOUGEON, François A: Comparison of six individual tree crown detection algorithms evaluated under varying forest conditions. In: International Journal of Remote Sensing 32 (2011), Nr. 20, S. 5827–5852

[Leckie et al. 2003] LECKIE, Don ; GOUGEON, François ; HILL, David ; QUINN, Rick ; ARMSTRONG, Lynne ; SHREENAN, Roger: Combined high-density lidar and multispectral imagery for individual tree crown analysis. In: Canadian Journal of Remote Sensing 29 (2003), Nr. 5, S. 633–649

[Li 2013] LI, Chunyang: Probability Estimation in Random Forests. (2013)

[Li et al. 2015] LI, Dan ; KE, Yinghai ; GONG, Huili ; LI, Xiaojuan: Object-based urban tree species classification using bi-temporal worldview-2 and worldview-3 images. In: Remote Sensing 7 (2015), Nr. 12, S. 16917–16937

[Lin et al. 2011] LIN, Chinsu ; THOMSON, Gavin ; LO, Chien-Shun ; YANG, Ming-Shein: A multi-level morphological active contour algorithm for delineating tree crowns in mountainous forest. In: Photogrammetric Engineering & Remote Sensing 77 (2011), Nr. 3, S. 241–249

[Liu et al. 2017] LIU, Luxia ; COOPS, Nicholas C. ; AVEN, Neal W. ; PANG, Yong: Mapping urban tree species using integrated airborne hyperspectral and LiDAR remote sensing data. In: Remote Sensing of Environment 200 (2017), S. 170–182

[Lu and Weng 2007] LU, Dengsheng ; WENG, Qihao: A survey of image classification methods and techniques for improving classification performance. In: International journal of Remote sensing 28 (2007), Nr. 5, S. 823–870

[da Luz and Crowley 2010] LUZ, Beatriz R. da ; CROWLEY, James K.: Identification of plant species by using high spatial and spectral resolution thermal infrared (8.0–13.5 μm) imagery. In: Remote Sensing of Environment 114 (2010), Nr. 2, S. 404–413

[MacFaden et al. 2012] MACFADEN, Sean W. ; O'NEIL-DUNNE, Jarlath P. ; ROYAR, Anna R. ; LU, Jacqueline W. ; RUNDLE, Andrew G.: High-resolution tree canopy mapping for New York City using LIDAR and object-based image analysis. In: Journal of Applied Remote Sensing 6 (2012), Nr. 1, S. 063567–1

[Mallinis et al. 2008] MALLINIS, Georgios ; KOUTSIAS, Nikos ; TSAKIRI-STRATI, Maria ; KARTERIS, Michael: Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. In: ISPRS Journal of Photogrammetry and Remote Sensing 63 (2008), Nr. 2, S. 237–250

[Martin et al. 1998] MARTIN, ME ; NEWMAN, SD ; ABER, JD ; CONGALTON, RG: Determining forest species composition using high spectral resolution remote sensing data. In: Remote Sensing of Environment 65 (1998), Nr. 3, S. 249–254

[McPherson et al. 2016] MCPHERSON, E G. ; DOORN, Natalie van ; GOEDE, John de: Structure, function and value of street trees in California, USA. In: Urban Forestry & Urban Greening 17 (2016), S. 104–115

[Merzlyak et al. 1999] MERZLYAK, Mark N. ; GITELSON, Anatoly A. ; CHIVKUNOVA, Olga B. ; RAKITIN, Victor Y.: Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. In: Physiologia plantarum 106 (1999), Nr. 1, S. 135–141

[MetLink 2017] METLINK: Urban Heat Island Introduction. `http://www.metlink.org/other-weather/urban-heat-islands/urban-heat-island-background/`. 2017. – Accessed: 27/06/18

[Miyajima and Takahashi 2007] MIYAJIMA, Yutaka ; TAKAHASHI, Koichi: Changes with altitude of the stand structure of temperate forests on Mount Norikura, central Japan. In: Journal of forest research 12 (2007), Nr. 3, S. 187–192

[Nagao et al. 1979] NAGAO, Makoto ; MATSUYAMA, Takashi ; IKEDA, Yoshio: Region extraction and shape analysis in aerial photographs. In: Computer Graphics and Image Processing 10 (1979), Nr. 3, S. 195–223

[Neu et al. 2014] NEU, Laurent ; TROULET, Claire ; VIGOUROUX, André: La sélection variétale au secours du platane décimé par l'épidémie de chancre coloré. In: Journées Jean Chevaugeon 2014. 10es Rencontres de Phytopathologie-Mycologie de la Société Française de Phytopathologie (SFP). 27 au 31 janvier 2014. 2014; 10. rencontres de Phytopathologie-Mycologie, Aussois, FRA, 2014-01-27-2014-01-31, (2014)

[Nidamanuri and Zbell 2011] NIDAMANURI, Rama R. ; ZBELL, Bernd: Use of field reflectance data for crop mapping using airborne hyperspectral image. In: ISPRS journal of photogrammetry and remote sensing 66 (2011), Nr. 5, S. 683–691

[Oke 2011] OKE, TR: Urban heat island. In: The Routledge Handbook of Urban Ecology. Routledge Abingdon, Oxon, 2011, S. 120–131

[Ørka et al. 2009] ØRKA, Hans O. ; NÆSSET, Erik ; BOLLANDSÅS, Ole M.: Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. In: Remote Sensing of Environment 113 (2009), Nr. 6, S. 1163–1174

[Otsu 1975] OTSU, Nobuyuki: A threshold selection method from gray-level histograms. In: Automatica 11 (1975), Nr. 285-296, S. 23–27

[Pal 2005] PAL, Mahesh: Random forest classifier for remote sensing classification. In: International Journal of Remote Sensing 26 (2005), Nr. 1, S. 217–222

[Paterska et al. 2017] PATERSKA, Maja ; BANDURSKA, Hanna ; WYSŁOUCH, Joanna ; MOLIŃSKA-GLURA, Marta ; MOLIŃSKI, Krzysztof: Chemical composition of horsechestnut (Aesculus) leaves and their susceptibility to chestnut leaf miner Cameraria ohridella Deschka & Dimić. In: Acta Physiologiae Plantarum 39 (2017), Nr. 4, S. 105

[Percival et al. 2011] PERCIVAL, Glynn C. ; BARROW, I ; NOVISS, K ; KEARY, I ; PENNINGTON, P: The impact of horse chestnut leaf miner (Cameraria ohridella Deschka and Dimic; HCLM) on vitality, growth and reproduction of Aesculus hippocastanum L. In: Urban Forestry & Urban Greening 10 (2011), Nr. 1, S. 11–17

[Perrin et al. 2004] PERRIN, Guillaume ; DESCOMBES, Xavier ; ZERUBIA, Josiane: Tree crown extraction using marked point processes. In: Signal Processing Conference, 2004 12th European IEEE (Veranst.), 2004, S. 2127–2130

[Perrin et al. 2005] PERRIN, Guillaume ; DESCOMBES, Xavier ; ZERUBIA, Josiane: A marked point process model for tree crown extraction in plantations. In: Image Processing, 2005. ICIP 2005. IEEE International Conference on Bd. 1 IEEE (Veranst.), 2005, S. I–661

[Perrin et al. 2006] PERRIN, Guillaume ; DESCOMBES, Xavier ; ZERUBIA, Josiane ; BOUREAU, Jean-Guy: Forest resource assessment using stochastic geometry. In: International Precision Forestry Symposium, 2006

[Platt et al. 1999] PLATT, John et al.: Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In: Advances in large margin classifiers 10 (1999), Nr. 3, S. 61–74

[Platt and Rapoza 2008] PLATT, Rutherford V. ; RAPOZA, Lauren: An evaluation of an object-oriented paradigm for land use/land cover classification. In: The Professional Geographer 60 (2008), Nr. 1, S. 87–100

[Poutier et al. 2002] POUTIER, Laurent ; MIESCH, Christophe ; LENOT, Xavier ; ACHARD, Véronique ; BOUCHER, Yannick: COMANCHE and COCHISE: two reciprocal atmospheric codes for hyperspectral remote sensing. In: 2002 AVIRIS Earth Science and Applications Workshop Proceedings, 2002

[Puissant et al. 2014] PUISSANT, Anne ; ROUGIER, Simon ; STUMPF, André: Object-oriented mapping of urban trees using Random Forest classifiers. In: International Journal of Applied Earth Observation and Geoinformation 26 (2014), S. 235–245

[Pulighe and Lupia 2016] PULIGHE, Giuseppe ; LUPIA, Flavio: Mapping spatial patterns of urban agriculture in Rome (Italy) using Google Earth and web-mapping services. In: Land Use Policy 59 (2016), S. 49–58

[Rajpoot and Rajpoot 2004] RAJPOOT, Kashif M. ; RAJPOOT, Nasir M.: Wavelets and support vector machines for texture classification. In: Multitopic Conference, 2004. Proceedings of INMIC 2004. 8th International IEEE (Veranst.), 2004, S. 328–333

[Ramdani 2013] RAMDANI, Fatwa: Urban vegetation mapping from fused hyperspectral image and LiDAR data with application to monitor urban tree heights. In: Journal of Geographic Information System 5 (2013), Nr. 04, S. 404

[Roberts et al. 2004] ROBERTS, Dar A. ; USTIN, Susan L. ; OGUNJEMIYO, Segun ; GREENBERG, Jonathan ; DOBROWSKI, Solomon Z. ; CHEN, Jiquan ; HINCKLEY, Thomas M.: Spectral and structural measures of northwest forest vegetation at leaf to landscape scales. In: Ecosystems 7 (2004), Nr. 5, S. 545–562

[Rochery et al. 2006] ROCHERY, Marie ; JERMYN, Ian H. ; ZERUBIA, Josiane: Higher Order Active Contours. In: International Journal of Computer Vision 69 (2006), Aug, Nr. 1, S. 27–42. – URL https://doi.org/10.1007/s11263-006-6851-y. – ISSN 1573-1405

[Rol-Tanguy et al. 2010] ROL-TANGUY, Francis ; ALBA, Dominique ; DRAGONI, Maria ; TER MINASSIAN, Hovig ; BLANCOT, Christiane: Essai de bilan sur le

développement des arbres d'alignement dans Paris. https://www.apur.org/fr/nos-travaux/bilan-developpement-arbres-alignement-paris. 2010. – Accessed: 24/10/17

[Rouse Jr et al. 1974] ROUSE JR, J_W ; HAAS, RH ; SCHELL, JA ; DEERING, DW: Monitoring vegetation systems in the Great Plains with ERTS. (1974)

[Savitzky and Golay 1964] SAVITZKY, Abraham ; GOLAY, Marcel J.: Smoothing and differentiation of data by simplified least squares procedures. In: Analytical chemistry 36 (1964), Nr. 8, S. 1627–1639

[Sebestyen et al. 2008] SEBESTYEN, D ; NEMETH, M ; HANGYAL, R ; KRIZBAI, L ; EMBER, I ; NYERGES, K ; KOLBER, M ; KISS, E ; BESE, G: Ornamental prunus species as new natural hosts of plum pox virus and their importance in the spread of the virus in hungary. In: Journal of Plant Pathology (2008), S. S57–S61

[Seinfeld and Pandis 2016] SEINFELD, John H. ; PANDIS, Spyros N.: Atmospheric chemistry and physics: from air pollution to climate change. John Wiley & Sons, 2016

[Seto et al. 2012] SETO, Karen C. ; GÜNERALP, Burak ; HUTYRA, Lucy R.: Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. In: Proceedings of the National Academy of Sciences 109 (2012), Nr. 40, S. 16083–16088

[Shackelford and Davis 2003] SHACKELFORD, Aaron K. ; DAVIS, Curt H.: A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. In: IEEE Transactions on GeoScience and Remote sensing 41 (2003), Nr. 10, S. 2354–2363

[Shafer 1976] SHAFER, Glenn: A mathematical theory of evidence. Bd. 42. Princeton university press, 1976

[Shang and Chisholm 2014] SHANG, Xiao ; CHISHOLM, Laurie A.: Classification of Australian native forest species using hyperspectral remote sensing and machine-learning classification algorithms. In: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7 (2014), Nr. 6, S. 2481–2489

[Sheeren et al. 2016] SHEEREN, David ; FAUVEL, Mathieu ; JOSIPOVIĆ, Veliborka ; LOPES, Maïlys ; PLANQUE, Carole ; WILLM, Jérôme ; DEJOUX, Jean-François: Tree Species Classification in Temperate Forests Using Formosat-2 Satellite Image Time Series. In: Remote Sensing 8 (2016), Nr. 9, S. 734

[Shojanoori and Shafri 2016] SHOJANOORI, Razieh ; SHAFRI, Helmi Z.: Review on the use of remote sensing for urban forest monitoring. In: Arboric Urban For 42 (2016), Nr. 6, S. 400–17

[Shojanoori et al. 2018] SHOJANOORI, Razieh ; SHAFRI, Helmi Z. ; MANSOR, Shattri ; ISMAIL, Mohd H.: Generic rule-sets for automated detection of urban tree species from very high-resolution satellite data. In: Geocarto International 33 (2018), Nr. 4, S. 357–374

[Smith et al. 2003] SMITH, M-L ; MARTIN, Mary E. ; PLOURDE, Lucie ; OLLINGER, Scott V.: Analysis of hyperspectral data for estimation of temperate forest canopy nitrogen concentration: comparison between an airborne (AVIRIS) and a spaceborne (Hyperion) sensor. In: IEEE Transactions on Geoscience and Remote Sensing 41 (2003), Nr. 6, S. 1332–1337

[Song et al. 2013] SONG, Chorong ; JOUNG, Dawou ; IKEI, Harumi ; IGARASHI, Miho ; AGA, Mariko ; PARK, Bum-Jin ; MIWA, Masayuki ; TAKAGAKI, Michiko ; MIYAZAKI, Yoshifumi: Physiological and psychological effects of walking on young males in urban parks in winter. In: Journal of physiological anthropology 32 (2013), Nr. 1, S. 18

[Souchon et al. 2010] SOUCHON, Jean-Philippe ; THOM, Christian ; MEYNARD, Christophe ; MARTIN, Olivier ; PIERROT-DESEILLIGNY, Marc: The IGN CAMv2 System. In: The Photogrammetric Record 25 (2010), Nr. 132, S. 402–421

[Stavrakoudis et al. 2014] STAVRAKOUDIS, Dimitris G. ; DRAGOZI, Eleni ; GITAS, Ioannis Z. ; KARYDAS, Christos G.: Decision fusion based on hyperspectral and multispectral satellite imagery for accurate forest species mapping. In: Remote Sensing 6 (2014), Nr. 8, S. 6897–6928

[Suárez et al. 2005] SUÁREZ, Juan C. ; ONTIVEROS, Carlos ; SMITH, Steve ; SNAPE, Stewart: Use of airborne LiDAR and aerial photography in the estimation of individual tree heights in forestry. In: Computers & Geosciences 31 (2005), Nr. 2, S. 253–262

[Tosics 2017] TOSICS, Ivan: Densification beyond the city centre: urban transformation against sprawl. `http://urbact.eu/densification-beyond-city-centre-urban-transformation-against-sprawl`. 2017. – Accessed: 18/05/18

[Tsunetsugu et al. 2013] TSUNETSUGU, Yuko ; LEE, Juyoung ; PARK, Bum-Jin ; TYRVÄINEN, Liisa ; KAGAWA, Takahide ; MIYAZAKI, Yoshifumi: Physiological and

psychological effects of viewing urban forest landscapes assessed by multiple measurements. In: Landscape and Urban Planning 113 (2013), S. 90–93

[Tuia et al. 2016] TUIA, Devis ; PERSELLO, Claudio ; BRUZZONE, Lorenzo: Domain adaptation for the classification of remote sensing data: An overview of recent advances. In: IEEE geoscience and remote sensing magazine 4 (2016), Nr. 2, S. 41–57

[Tupin et al. 1999] TUPIN, Florence ; BLOCH, Isabelle ; MAÎTRE, Henri: A first step toward automatic interpretation of SAR images using evidential fusion of several structure detectors. In: IEEE Transactions on Geoscience and Remote Sensing 37 (1999), Nr. 3, S. 1327–1343

[USGS 2011] USGS: Earth Observing 1 (EO-1). https://eo1.usgs.gov/sensors/hyperion. 2011. – Accessed: 22/05/18

[Ustin and Gamon 2010] USTIN, Susan L. ; GAMON, John A.: Remote sensing of plant functional types. In: New Phytologist 186 (2010), Nr. 4, S. 795–816

[Vailshery et al. 2013] VAILSHERY, Lionel S. ; JAGANMOHAN, Madhumitha ; NAGENDRA, Harini: Effect of street trees on microclimate and air pollution in a tropical city. In: Urban forestry & urban greening 12 (2013), Nr. 3, S. 408–415

[Van Leeuwen et al. 2010] VAN LEEUWEN, Martin ; COOPS, Nicholas C. ; WULDER, Michael A.: Canopy surface reconstruction from a LiDAR point cloud using Hough transform. In: Remote Sensing Letters 1 (2010), Nr. 3, S. 125–132

[Van Lieshout 2000] VAN LIESHOUT, MNM: Markov point processes and their applications. World Scientific, 2000

[Verhoef 1984] VERHOEF, Wouter: Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. In: Remote sensing of environment 16 (1984), Nr. 2, S. 125–141

[Vigouroux 2014] VIGOUROUX, André: Le chancre coloré du platane. http://agriculture.gouv.fr/sites/minagri/files/point_dactu_-_le_chancre_du_platane_-_decembre_2014.pdf. 2014. – Accessed: 07/11/16

[Voelcker 2014] VOELCKER, John: 1.2 Billion Vehicles On World's Roads Now, 2 Billion By 2035: Report. https://www.greencarreports.com/news/1093560_1-2-billion-vehicles-on-worlds-roads-now-2-billion-by-2035-report. 2014. – Accessed: 18/05/18

[Wang et al. 2009] WANG, Xin-Lu ; WASKE, Björn ; BENEDIKTSSON, Jon A.: Ensemble methods for spectral-spatial classification of urban hyperspectral data. In: Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009 Bd. 4 IEEE (Veranst.), 2009, S. IV–944

[Weber and Puissant 2003] WEBER, Christiane ; PUISSANT, Anne: Urbanization pressure and modeling of urban growth: example of the Tunis Metropolitan Area. In: Remote sensing of environment 86 (2003), Nr. 3, S. 341–352

[Welch 1982] WELCH, Roy: Spatial resolution requirements for urban studies. In: International Journal of Remote Sensing 3 (1982), Nr. 2, S. 139–146

[Wen et al. 2017] WEN, Dawei ; HUANG, Xin ; LIU, Hui ; LIAO, Wenzhi ; ZHANG, Liangpei: Semantic Classification of Urban Trees Using Very High Resolution Satellite Imagery. In: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 10 (2017), Nr. 4, S. 1413–1424

[Wikipédia 2018] WIKIPÉDIA: Mineuse du marronnier. `https://fr.wikipedia.org/wiki/Mineuse_du_marronnier`. 2018. – Accessed: 27/07/18

[Wu et al. 2004] WU, Ting-Fan ; LIN, Chih-Jen ; WENG, Ruby C.: Probability estimates for multi-class classification by pairwise coupling. In: Journal of Machine Learning Research 5 (2004), Nr. Aug, S. 975–1005

[Xia et al. 2015] XIA, Junshi ; CHANUSSOT, Jocelyn ; DU, Peijun ; HE, Xiyan: Spectral–spatial classification for hyperspectral data using rotation forests with local feature extraction and Markov random fields. In: IEEE Transactions on Geoscience and Remote Sensing 53 (2015), Nr. 5, S. 2532–2546

[Xia et al. 2017] XIA, Junshi ; YOKOYA, Naoto ; IWASAKI, Akira: A novel ensemble classifier of hyperspectral and LiDAR data using morphological features. In: Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on IEEE (Veranst.), 2017, S. 6185–6189

[Xiao et al. 2004] XIAO, Q ; USTIN, SL ; MCPHERSON, EG: Using AVIRIS data and multiple-masking techniques to map urban forest tree species. In: International Journal of Remote Sensing 25 (2004), Nr. 24, S. 5637–5654

[Yang et al. 2005] YANG, Jun ; MCBRIDE, Joe ; ZHOU, Jinxing ; SUN, Zhenyuan: The urban forest in Beijing and its role in air pollution reduction. In: Urban Forestry & Urban Greening 3 (2005), Nr. 2, S. 65–78

[Yin et al. 2011] YIN, Shan ; SHEN, Zhemin ; ZHOU, Pisheng ; ZOU, Xiaodong ; CHE, Shengquan ; WANG, Wenhua: Quantifying air pollution attenuation within urban parks: An experimental approach in Shanghai, China. In: Environmental pollution 159 (2011), Nr. 8, S. 2155–2163

[Zhang and Xie 2012] ZHANG, Caiyun ; XIE, Zhixiao: Combining object-based texture measures with a neural network for vegetation mapping in the Everglades from hyperspectral imagery. In: Remote Sensing of Environment 124 (2012), S. 310–320

[Zhang 2010] ZHANG, Jixian: Multi-source remote sensing data fusion: status and trends. In: International Journal of Image and Data Fusion 1 (2010), Nr. 1, S. 5–24

[Zhang and Hu 2012] ZHANG, Kongwen ; HU, Baoxin: Individual urban tree species classification using very high spatial resolution airborne multi-spectral imagery using longitudinal profiles. In: Remote Sensing 4 (2012), Nr. 6, S. 1741–1757

[Zhen et al. 2015] ZHEN, Zhen ; QUACKENBUSH, Lindi J. ; STEHMAN, Stephen V. ; ZHANG, Lianjun: Agent-based region growing for individual tree crown delineation from airborne laser scanning (ALS) data. In: International Journal of Remote Sensing 36 (2015), Nr. 7, S. 1965–1993

[Zhen et al. 2014] ZHEN, Zhen ; QUACKENBUSH, Lindi J. ; ZHANG, Lianjun: Impact of tree-oriented growth order in marker-controlled region growing for individual tree crown delineation using airborne laser scanner (ALS) data. In: Remote Sensing 6 (2014), Nr. 1, S. 555–579

[Zhen et al. 2016] ZHEN, Zhen ; QUACKENBUSH, Lindi J. ; ZHANG, Lianjun: Trends in Automatic Individual Tree Crown Detection and Delineation—Evolution of LiDAR Data. In: Remote Sensing 8 (2016), Nr. 4, S. 333

[Zhou et al. 2010] ZHOU, Jia ; PROISY, Christophe ; DESCOMBES, Xavier ; HEDHLI, Ihssen ; BARBIER, Nicolas ; ZERUBIA, Josiane ; GASTELLU-ETCHEGORRY, Jean-Philippe ; COUTERON, Pierre: Tree crown detection in high resolution optical and LiDAR images of tropical forest. In: Remote Sensing International Society for Optics and Photonics (Veranst.), 2010, S. 78240Q–78240Q

[Zhu et al. 2001] ZHU, A-Xing ; HUDSON, B ; BURT, J ; LUBICH, K ; SIMONSON, D: Soil mapping using GIS, expert knowledge, and fuzzy logic. In: Soil Science Society of America Journal 65 (2001), Nr. 5, S. 1463–1472

Résumé en français

Avec l'expansion des zones urbaines, la pollution de l'air et les effets des îlots de chaleur augmentent, entraînant des problèmes d'état de santé pour les habitants et des changements climatiques globaux. Dans ce contexte, les arbres urbains sont une ressource précieuse à la fois pour améliorer la qualité de l'air et promouvoir les îlots de fraîcheur. D'autre part, les canopées sont soumises à des conditions spécifiques dans l'environnement urbain, provoquant la propagation de maladies et une diminution de l'espérance de vie parmi les arbres. Cette thèse explore le potentiel de la télédétection pour la cartographie automatique des arbres urbains, depuis la détection des couronnes d'arbres jusqu'à l'estimation de leur espèce. C'est une tâche préliminaire essentielle pour la conception des futures villes vertes et pour un suivi efficace de la végétation. Fondé sur des données hyperspectrales aéroportées, panchromatiques et Digital Surface Model, le premier objectif de cette thèse consiste à exploiter plusieurs sources de données pour améliorer les cartes urbaines existantes, en testant différentes stratégies de fusion (fusion de caractéristiques et de décision). La nature des résultats nous a conduit à optimiser la complémentarité des sources. En particulier, le second objectif est d'étudier en profondeur la richesse des données hyperspectrales, en développant un ensemble de classifieurs fondé sur des classifieurs spécifiques aux espèces. Les caractéristiques sont construites grâce à une sélection d'indices de végétation, spécifique à chaque espèce. Enfin, la première partie a mis en évidence l'intérêt de discriminer les arbres d'alignement des autres structures d'arbres urbains. Dans un cadre de processus de points marqués, le troisième objectif est de détecter les arbres appartenant à un alignement. Grâce au premier objectif, cette thèse démontre que les données hyperspectrales (en particulier le VNIR) sont le principal moteur de la précision de la prédiction de l'espèce. La stratégie de fusion au niveau de la décision est la plus appropriée pour améliorer les performances par rapport aux données hyperspectrales seules, mais de légères améliorations sont obtenues (quelques pourcents) en raison de la faible complémentarité des caractéristiques texturales et structurales en plus des caractéristiques spectrales. L'approche d'ensemble développée dans la seconde partie permet de classer les espèces d'arbres à partir de mesures spectrales terrain (à l'échelle de la feuille ou de la canopée), avec des améliorations significatives par rapport à une approche standard de classification au niveau des caractéristiques. Enfin, les arbres d'alignement peuvent être cartographiés grâce au modèle proposé intégrant des caractéristiques contextuelles (alignement et hauteurs similaires). Ce travail pourrait être étendu au suivi phénologique de la végétation urbaine et à l'analyse de l'état de santé.

Mots clés : Urbain, Arbre, Télédétection, Hyperspectral, Panchromatique, Modèle Numérique de Surface, Orienté objet, Fusion, Ensemble de classifieurs, MPP, Indices de végétation.

English abstract

With the expansion of urban areas, air pollution and heat island effects are increasing, leading to state of health issues for the inhabitants and global climate changes. In this context, urban trees are a valuable resource for both improving air quality and promoting freshness islands. On the other hand, canopies are subject to specific conditions in the urban environment, causing the spread of diseases and life expectancy decreases among the trees. This thesis explores the potential of remote sensing for the automatic urban tree mapping, from the detection of the individual tree crowns to their species estimation. This is an essential preliminary task for designing the future green cities, and for an effective vegetation monitoring. Based on airborne hyperspectral, panchromatic and Digital Surface Model data, the first objective of this thesis consists in taking advantage of several data sources for improving the existing urban tree maps, by testing different fusion strategies (feature and decision level fusion). The nature of the results led us to optimize the complementarity of the sources. In particular, the second objective is to investigate deeply the richness of the hyperspectral data, by developing an ensemble classifier approach based on species specific classifiers. The features are built owing to vegetation indices selection, specific for each species. Finally, the first part highlighted to interest of discriminating the street trees from the other structures of urban trees. In a Marked Point Process framework, the third objective is to detect trees in urban alignment. Through the first objective, this thesis demonstrates that the hyperspectral data (especially the VNIR) are the main driver of the species prediction accuracy. The decision level fusion strategy is the most appropriate one for improving the performance in comparison the hyperspectral data alone, but slight improvements are obtained (a few percent) due to the low complementarity of textural and structural features in addition to the spectral ones. The ensemble classifier approach developed in the second part allows the tree species to be classified from ground-based spectral references (leaf and canopy levels), with significant improvements in comparison to a standard feature level classification approach. Finally, the street trees can be mapped thanks to the proposed model integrating contextual features (alignment and similar heights). This work could be extended to the phenological monitoring of urban vegetation and the analysis of the state of health.

Keywords: Urban, Tree, Remote sensing, Hyperspectral, Panchromatic, Digital Surface Model, Object-based, Fusion, Ensemble classifier, MPP, Vegetation indices.