



**HAL**  
open science

## Detection of individual trees in urban alignment from airborne data and contextual information : A marked point process approach

Josselin Aval, Jean Demuynck, Emmanuel Zenou, Sophie Fabre, David Sheeren, Mathieu Fauvel, Karine Adeline, Xavier Briottet

### ► To cite this version:

Josselin Aval, Jean Demuynck, Emmanuel Zenou, Sophie Fabre, David Sheeren, et al.. 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, 146, pp.197-210. 10.1016/j.isprsjprs.2018.09.016 . hal-01991252

**HAL Id: hal-01991252**

**<https://hal.science/hal-01991252>**

Submitted on 23 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of some Toulouse researchers and makes it freely available over the web where possible.

This is an author's version published in: <https://oatao.univ-toulouse.fr/21691>

**Official URL :** <https://doi.org/10.1016/j.isprs.2018.09.016>

### To cite this version :

Aval, Josselin and Jean, Demuynck and Zenou, Emmanuel and Sophie, Fabre and David, Sheeren and Mathieu, Fauvel and Karine, Adeline and Xavier, Briottet Detection of individual trees in urban alignment from airborne data and T contextual information: A marked point process approach. (2018) ISPRS Journal of Photogrammetry and Remote Sensing, 146. 197-210. ISSN 0924-2716

Any correspondence concerning this service should be sent to the repository administrator:

[tech-oatao@listes-diff.inp-toulouse.fr](mailto:tech-oatao@listes-diff.inp-toulouse.fr)

# Detection of individual trees in urban alignment from airborne data and contextual information: A marked point process approach

Josselin Aval<sup>1a,\*</sup>, Jean Demuynck<sup>b</sup>, Emmanuel Zenou<sup>b</sup>, Sophie Fabre<sup>a</sup>, David Sheeren<sup>c</sup>,  
Mathieu Fauvel<sup>c</sup>, Karine Adeline<sup>a</sup>, Xavier Briottet<sup>a</sup>

<sup>a</sup> ONERA-DOTA, University of Toulouse, Toulouse 31055, France

<sup>b</sup> ISAE-SUPAERO, University of Toulouse, Toulouse 31400, France

<sup>c</sup> DYNAFOR, INP-ENSAT, INP-EI Purpan, INRA, University of Toulouse, Auzeville 31320, France

---

## A B S T R A C T

### Keywords:

Street tree  
Urban remote sensing  
Airborne data  
Geographic information system  
Individual tree crown delineation  
Marked point process

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<sub>2</sub> 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 modeling 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. 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<sup>2</sup> (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

---

\* Corresponding author.

E-mail address: [josselin.aval@onera.fr](mailto:josselin.aval@onera.fr) (J. Aval).

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 data bases 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 multispectral 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 modeled. 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 2 presents the study area and the data used for individual street tree mapping, followed by Section 3 with the description of the proposed method and a baseline ITCD method used for comparison. Afterwards, the results are showed in Section 4 and discussed in Section 5. Finally, main conclusions of the study are detailed and the perspectives of the work highlighted in Section 6.

## 2. Materials

### 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 (Fig. 1).

The three study cases are presented in Fig. 2. In all cases, the street trees form lines along roads and are pruned as it will be highlighted in Section 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 × 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 × hispanica*).

### 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 2000 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  $\mu\text{m}$ ). About the spatial

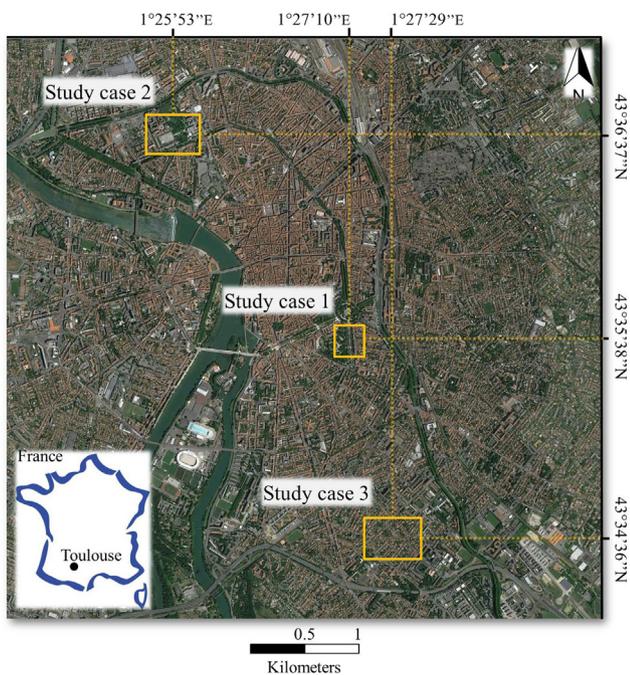


Fig. 1. Study area with the three study cases represented on a Google Earth image. The yellow rectangles correspond to each study case. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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 (Fig. 3).

### 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 × 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.

### 3. Methods

The description of the proposed method is carried out in Section 3.1, followed by Section 3.2 with the description of a baseline ITCO method used for comparison.

#### 3.1. Proposed street tree mapping

Fig. 4 presents the proposed street tree mapping scheme. First, the high vegetation close to the streets is detected (Section 3.1.1). Secondly, the street tree crowns are delineated based on MPP which allow a prior contextual information to be modeled via an interaction term (Section 3.1.2). 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 has the same height as its neighbors.

##### 3.1.1. 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  $U_d(x)$  of the MPP, defined in Section 3.1.2 (Fig. 4). The mask is the MPP input data that allows to reduce the

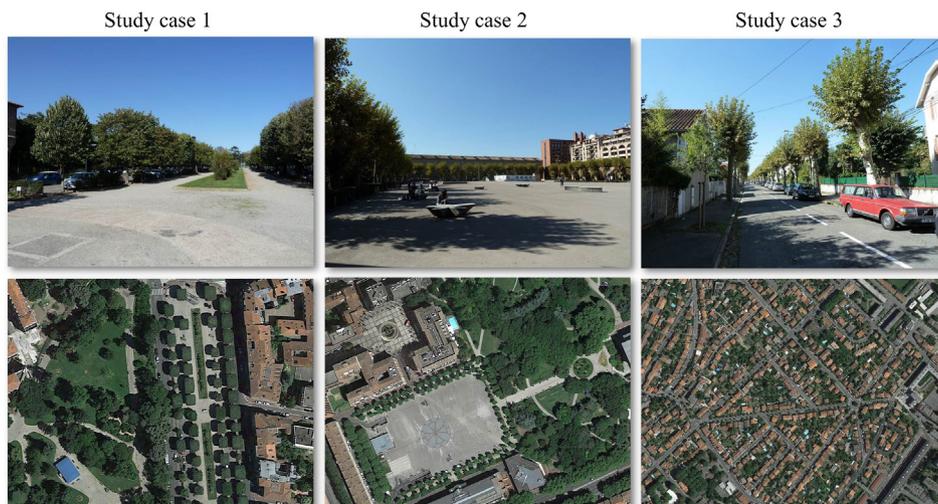


Fig. 2. Description of the study cases. At the top, field view is showed for each case. At the bottom, Google Earth images illustrate each case.

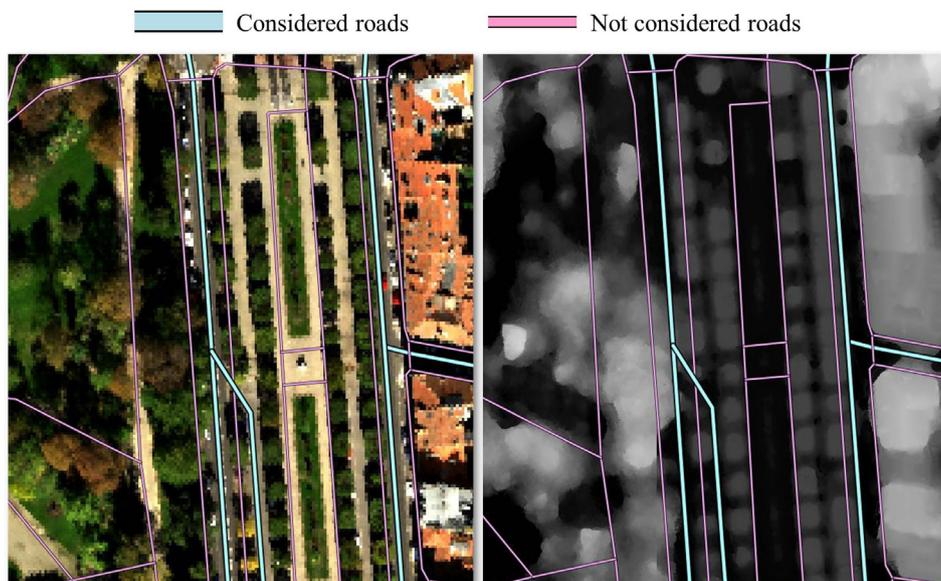


Fig. 3. Illustration of the data used in this study for the first site. Left: VNIR reflectance image. Right: DSM. The GIS data is represented by blue and violet lines indicated over the airborne data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

computational time, by restricting the search space  $P$  (introduced in Section 3.1.2). For the vegetation mask, the NDVI (Normalized Difference Vegetation Index) index (Rouse 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 (Fig. 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

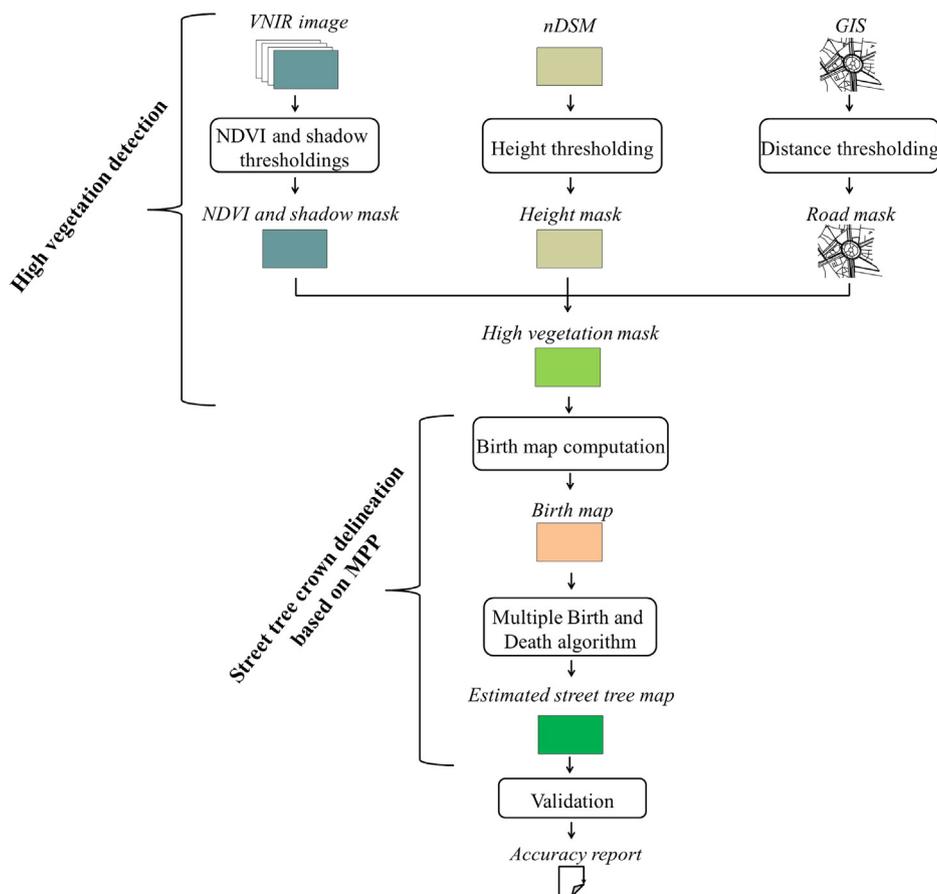


Fig. 4. Principle of the street tree mapping method.

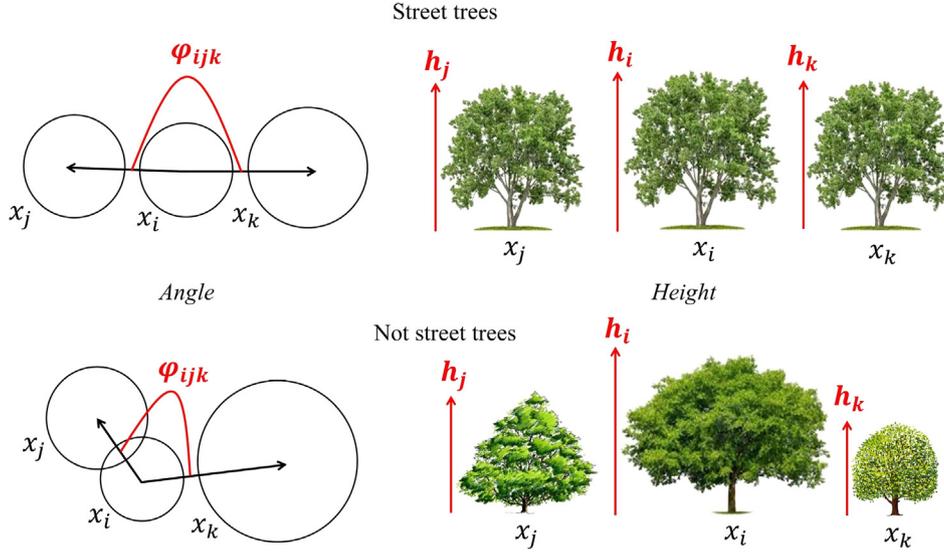


Fig. 5. Illustration of the contextual features used to compute the interaction energy. At the top, the case of street trees is described (similar height, tree species and alignment). At the bottom, the case of trees which are not street trees is presented (various height, different tree species and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

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).

### 3.1.2. 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  $\chi$  in which  $\mathbf{x}$  is a realization can be defined such as:

$$\chi = P \times M = [1, X_M] \times [1, Y_M] \times [r_m, r_M] \quad (1)$$

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) (Eq. (2)).

$$d(x_i) = \frac{N_{x_i} - N_{a_{x_i}}}{N_{t_{x_i}} - N_{t_{a_{x_i}}}}$$

$$U_d(x_i) = \begin{cases} 1 - \frac{d(x_i)}{d_0} & \text{if } d(x_i) < d_0 \\ \exp[-(d(x_i) - d_0)] - 1 & \text{if } d_0 \leq d(x_i) \end{cases} \quad (2)$$

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) (Eq. (3)).

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

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 Fig. 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 neighboring street trees, i.e. an alignment is modeled from three street trees (Fig. 5 and Eq. (4)). We choose three trees because such a model is more flexible in case of curved roads for example.

$$U_{ij} \left[ x_i, \left( x_j, x_k \right) \right] = \frac{1}{a+b} \left( a \cdot \frac{\theta_{ijk}}{\pi/2} + b \cdot \frac{|h_i - h_j| + |h_i - h_k|}{2 \cdot \max(h_i, h_j, h_k)} \right) \in [0, 1]$$

$$U_{ij(x_i)} = \begin{cases} 1 & \text{if } V_{x_i} = \emptyset \\ \min_{(x_j, x_k) \in V_{x_i}} U_{ij} [x_i, (x_j, x_k)] & \text{otherwise} \end{cases}$$

where  $\theta_{ijk} = \pi - \varphi_{ijk}$  if  $\pi/2 < \varphi_{ijk} \leq \pi$   
and  $\theta_{ijk} = \varphi_{ijk}$  if  $0 \leq \varphi_{ijk} \leq \pi/2$  (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  $\theta$ ” whereas is referred as “without  $\theta$ ” (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).  $\varphi_{ijk}$  is the angle between the two segments joining the center of  $x_i$  with the centers of  $x_j$  and  $x_k$ . In particular,  $U_{ij} [x_i, (x_j, x_k)]$  is computed for all pairs of neighbors  $x_j$  and  $x_k$  in the neighborhood  $V_{x_i}$  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  $U_{ij} [x_i, (x_j, x_k)]$  is then retained as  $U_{ij}(x_i)$ . If there is no pair of neighbors,  $U_{ij}(x_i)$  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 neighboring street trees of similar heights,  $U_{ij}(x_i)$  is close to 0. Other configurations are penalized as they result in higher values of  $U_{ij}(x_i)$ .

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  $\delta$  (equal to 200 which corresponds to the order of magnitude of the number of trees 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. There fore, this choice is empirical and may thus be specific to our dataset and study areas. 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  $U_{ij}(x_i)$  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  $U_{ij}(x_i)$  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:

$$d(x_i) = \frac{\delta \cdot e(x \setminus x_i)}{1 + \delta \cdot e(x \setminus x_i)}$$

$$\text{where } e(x \setminus x_i) = \exp \left( \frac{\alpha \cdot U_{ij}(x_i) + \beta \cdot U_{is}(x_i) + \gamma \cdot U_{ij}(x_i)}{T} \right) \quad (5)$$

with  $\alpha$ ,  $\beta$  and  $\gamma$  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  $\theta$  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 to reduce computational time,  $\gamma$  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,  $\gamma$  is set to 3.5 to map the street trees. At the end of each iteration,  $T$  and  $\delta$  are multiplied by a factor of 0.997 (similar to that used in Descombes et al. (2009)).

### 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 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  $\sigma_{Gauss}$  is equal to 2 here 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.

### 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 1).

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  $F\text{-score}(\%) = 100 \cdot \frac{2 \cdot PA \cdot UA}{PA + UA}$ .

**Table 1**

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

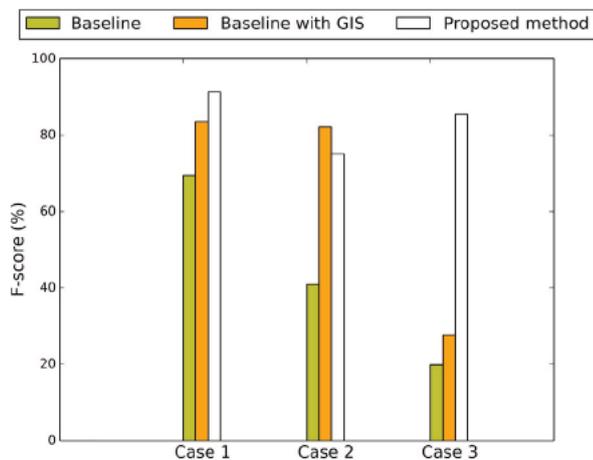


Fig. 6. Comparison of F-score between the proposed method and the baseline method for each study case.

#### 4. Results

In this section, two main results are presented. First, the proposed method with all the features (with GIS,  $\theta$  and  $h$ , Section 3.1.2, Eq. (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 analyzed in order to identify the set of the most discriminative contextual features for street tree mapping.

##### 4.1. Comparison between the proposed method and the baseline method

Fig. 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, Fig. 7 shows the confusion matrices of the two approaches. In particular, there are a lot of “False Positive” trees (Table 1) 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 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 Fig. 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

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

Fig. 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 Fig. 6. The prediction is per column. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

reference method. Focusing on the baseline, this figure demonstrates also that a lot of not street trees can be filtered just in 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 1) 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 over segmented by the reference method. The important overlaps of the 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.

##### 4.2. Contributions of the angle, the height and the GIS data in the street tree mapping

Fig. 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,  $\theta$  and  $h$ , Section 3.1.2, Eq. (4)), the F score values are 91%, 75% and 85% instead of 76%, 58% and 26% without any feature (without GIS,  $\theta$  and  $h$ ) for the three cases respectively. Without GIS but with  $\theta$  and  $h$ , the F score values become 88%, 79% and 62%. With GIS but without  $\theta$  and  $h$ , the F score 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

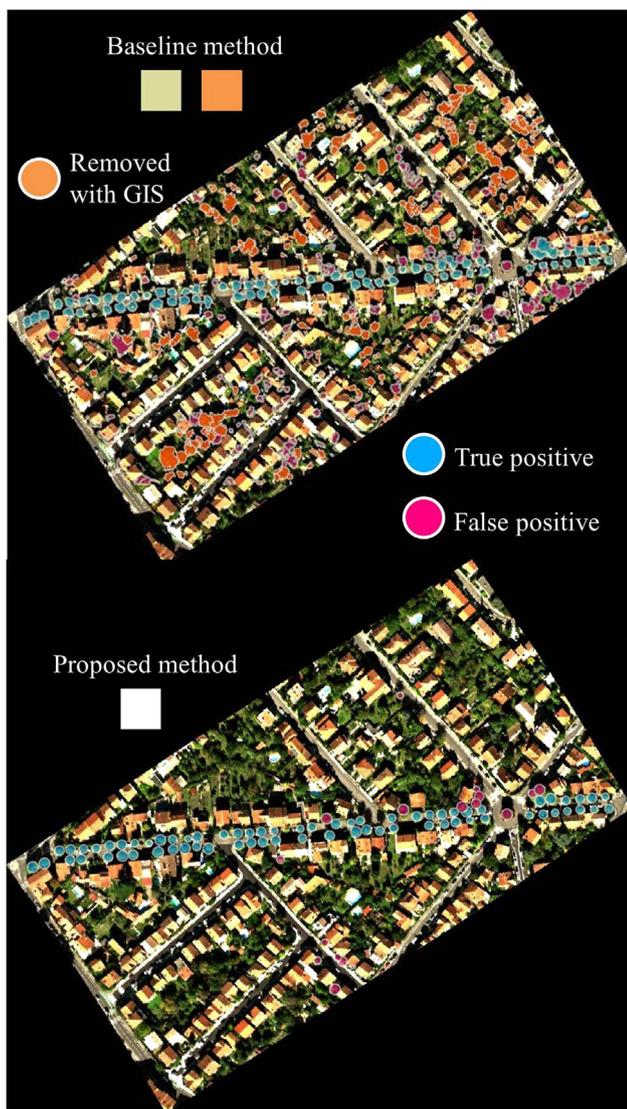


Fig. 8. Comparison of the produced maps for the baseline method and the proposed method for the third case. The colors under “Baseline method” and “Proposed method” refer to the colors used in Fig. 6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

improvements of 15p.p., 17p.p. and 59p.p. compared to the case where no feature is used). However, using only  $\theta$  and  $h$  gives already good results. On the other hand, using only the GIS data is not appropriate and needs the integration of  $\theta$  and  $h$ . Focusing on  $\theta$  and  $h$ , they have to be used together. Fig. 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.

In order to go further in the analysis, Fig. 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  $\theta$  and  $h$ . Especially, there are 3 further “False Negative” trees (PA decreases of 6p.p. and 4p.p.) 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

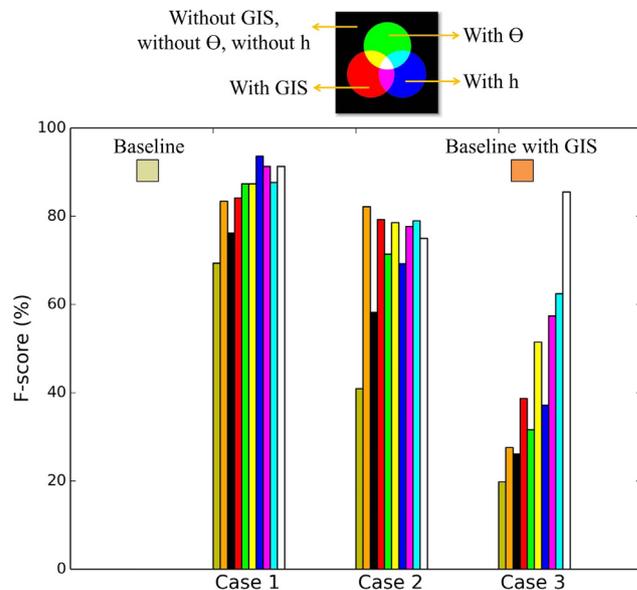


Fig. 9. Contribution of the angle, the height and the GIS data in terms of F-score for each study case. Each color corresponds to a combination of red (if GIS used), green (if  $\theta$  used) and blue (if  $h$  used) colors. The color under “Baseline method” and the white color refer to the colors used in Fig. 6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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  $\gamma$  in the proposed method, Section 3.1.2, Eq. (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 (Fig. 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).

## 5. Discussions

### 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 specific categories of urban trees

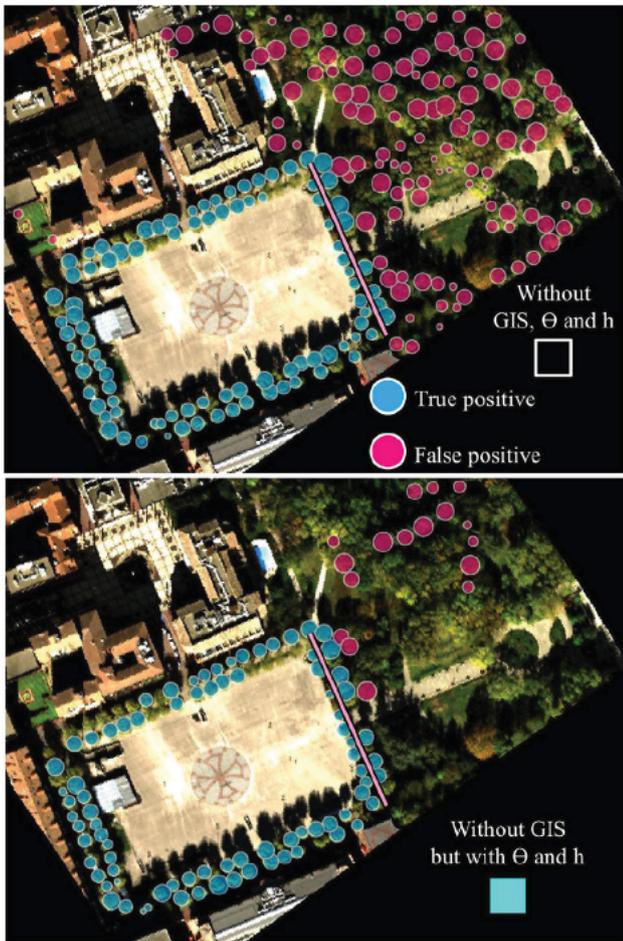


Fig. 10. Comparison of the produced maps with and without the use of  $\theta$  and  $h$  for the second case. The colors under “Without GIS but with  $\theta$  and  $h$ ” and “Without GIS,  $\theta$  and  $h$ ” refer to the colors used in Fig. 9. The pink road is a not considered pathway. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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 neighborhood properties (Alvarez Uriá and Körner, 2007). As an example, it is well known that deciduous trees cannot 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, modeling 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 modeling 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 datasources, which can cause errors in cases of significant overlap.

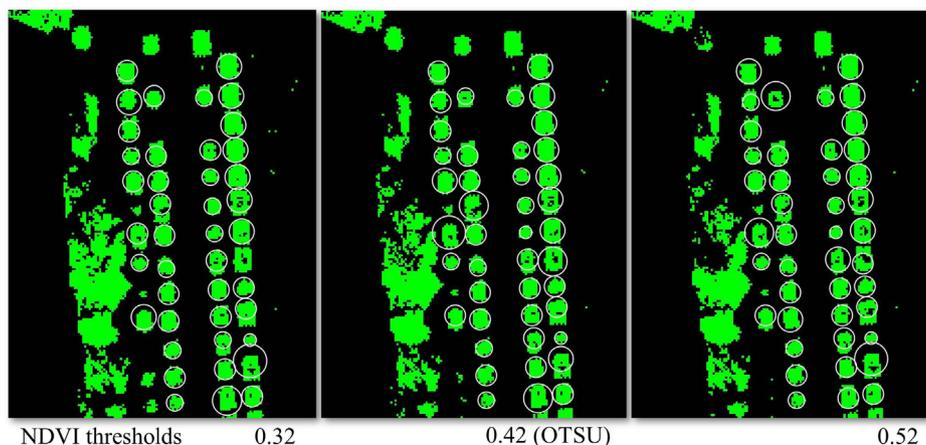
Case 1	Case 2	Case 3
With GIS but without $\theta$ and $h$ (■)		
$\hat{st}$ $\widehat{st}$	$\hat{st}$ $\widehat{st}$	$\hat{st}$ $\widehat{st}$
$st$ 45 5	$st$ 74 33	$st$ 72 0
$\overline{st}$ 12 -	$\overline{st}$ 6 -	$\overline{st}$ 229 -
Without GIS but with $\theta$ and $h$ (■)		
$\hat{st}$ $\widehat{st}$	$\hat{st}$ $\widehat{st}$	$\hat{st}$ $\widehat{st}$
$st$ 39 11	$st$ 82 25	$st$ 70 2
$\overline{st}$ 0 -	$\overline{st}$ 19 -	$\overline{st}$ 83 -
With GIS, with $\theta$ and $h$ (□)		
$\hat{st}$ $\widehat{st}$	$\hat{st}$ $\widehat{st}$	$\hat{st}$ $\widehat{st}$
$st$ 42 8	$st$ 64 43	$st$ 69 3
$\overline{st}$ 0 -	$\overline{st}$ 0 -	$\overline{st}$ 21 -

Fig. 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  $\theta$  and  $h$ . Middle: without GIS but with  $\theta$  and  $h$ . Bottom: with GIS, with  $\theta$  and  $h$ . The colors refer to the colors used in Fig. 9. The prediction is per column. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 12 shows the behavior of the proposed method with all the features (with GIS,  $\theta$  and  $h$ , Section 3.1.2, Eq. (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.

### 5.2. Discriminative contextual features of the street trees

This paper 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 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



**Fig. 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,  $\theta$  and  $h$ , Section 3.1.2, Eq. (4)) are used. Values of shadow, height and distance thresholds are fixed.

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 (RJCMCMC) or Multiple Birth and Cut (MBC) could have been used. While the MBD outperforms the RJCMCMC 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.

### 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 its applicability, the proposed method (with GIS,  $\theta$  and  $h$ ) is tested on the ISPRS dataset (Cramer, 2010), keeping the same set of parameters than for Toulouse city for assessing the robustness of the method. The Vaihingen data corresponding to the 2D semantic labelling benchmark test are considered. Initially equal to 8 cm and 9 cm respectively, the spatial resolutions of the digital aerial images and DSM are resampled (nearest neighbor) to 50 cm, in order to reduce computational time. On the other hand, as the ground is not flat, the DSM is normalized using the histogram based thresholding detailed in Section 2.3, except that the histogram is computed in a 60 m window (instead of the whole DSM initially) to ensure that we include ground areas in cases of overlapping canopies. The results of this test are illustrated in Fig. 13. The proposed approach reaches a F score value of 61%, against 55% (−6p.p.) for the baseline. UA values of 89% and 40% are obtained, while PA values of 47% and 88% are achieved, respectively. Fig. 13 illustrates these accuracies. The proposed approach leads to only few false positive cases, while the baseline one detects all the trees that are close to the streets, leading to many false positive trees. For that reason, the alignment trees are not omitted by the reference method. However, our approach omits many street trees, for example

an alignment of very small tree crowns in the lower right corner of Fig. 13. This is the same behavior as in Toulouse: the proposed approach does not confuse the street trees and the other ones, but tends to underestimate the number of street trees, especially because the street trees are not perfectly circular, aligned and of the same height. Overall, from the user's point of view, our method only detects street trees, which is the objective of this study. In addition, our method allows curved alignments to be identified, thanks to the flexibility of using triplets of trees.

Moreover, we have computed the street tree feature energy (Section 3.1.2, Eq. (4)) of all the inventoried street trees in the tree database of Paris (Fig. 14). This database contains, among other information, the location of the street tree trunks and the heights. 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 neighbors and have exactly the same heights than their neighbors. 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 neighbors with heights of 11 m, or a neighbor with a height of 11 m and the other with a height of 13 m, or a neighbor 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. Table 2 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 10 s, 10 min and 30 min (10 s, 1 min and 20 min with the GIS data) for the three study cases, instead of 20 min, 50 min and 3 h (20 min, 30 min and 1.5 h 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 behavior 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



Fig. 13. Comparison of the produced maps for the baseline method and the proposed method for the ISPRS dataset. The colors under “Baseline method” and “Proposed method” refer to the colors used in Fig. 6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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.

#### 5.4. Limits of the proposed approach

In summary, the limits of the proposed approach are illustrated in Fig. 15. On the one hand, the first limit of the proposed method occurs for cases of significant overlaps between the tree crowns (Fig. 15, left). In such a situation, the crowns are difficult to discern in the high vegetation mask, and it is accentuated in our autumn based experiment because of certain trees having a lot of senescent leaves (Fig. 15, left). Thus this implies having data outside winter/autumn. For these cases, nDSM based features modeling the 3D shape of the trees could make the approach more accurate. An alternative would be also to consider

the radiometric levels of the optical images directly, for example, by considering that the boundaries of the crowns appear darker in the images. On the other hand, a significant spatial connection between the street trees and the other ones can lead to failure cases, as illustrated in Fig. 15, middle. Focusing on this figure, the magenta individual is aligned with one of its neighboring street tree, and has the same height. Including that the street trees are aligned along the roads, in addition to be close to them, could be an effective solution. However, this is an unlikely thus less critical issue. Moreover, the street trees can have different heights. This is the case in Fig. 15, right, and the smallest tree is omitted by our approach. This happens when new trees are planted. A solution could be to consider that a tree aligned with its neighbors along a road is probably a street tree, even if it has a different height. Finally, the results of the proposed method on the Vaihingen data demonstrate that very small crowns can be omitted. While degrading the spatial resolution allows to reduce computational time, this leads very small crowns to have few pixels, thus a non circular shape which is not well handled by our approach. A solution would be to improve the modeling of the crown shape.

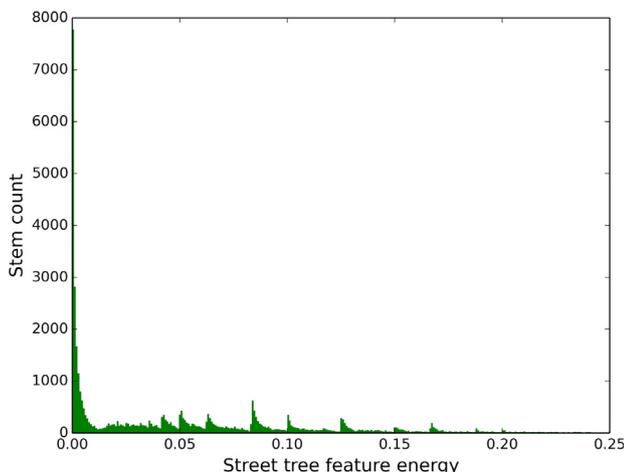


Fig. 14. Histogram of the street tree feature energy (Section 3.1.2, Eq. (4)) computed from 43,168 street trees of the tree database of Paris. Only the main species are taken into account: plane tree, horse chestnut, pagoda tree and lime tree (*Platanus × hispanica*, *Aesculus hippocastanum*, *Sophora japonica* and *Tilia tomentosa*).

Table 2 Comparison of the baseline and proposed methods in terms of computational burden.

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

### 5.5. 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 × hispanica* trees. In particular, the tree crowns are first estimated thanks to the proposed method with all the features (with GIS,  $\theta$  and  $h$ , Section 3.1.2, Eq. (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 Fig. 16 (top: before regularization), there are mainly errors in the right part of the alignment (18 trees on the

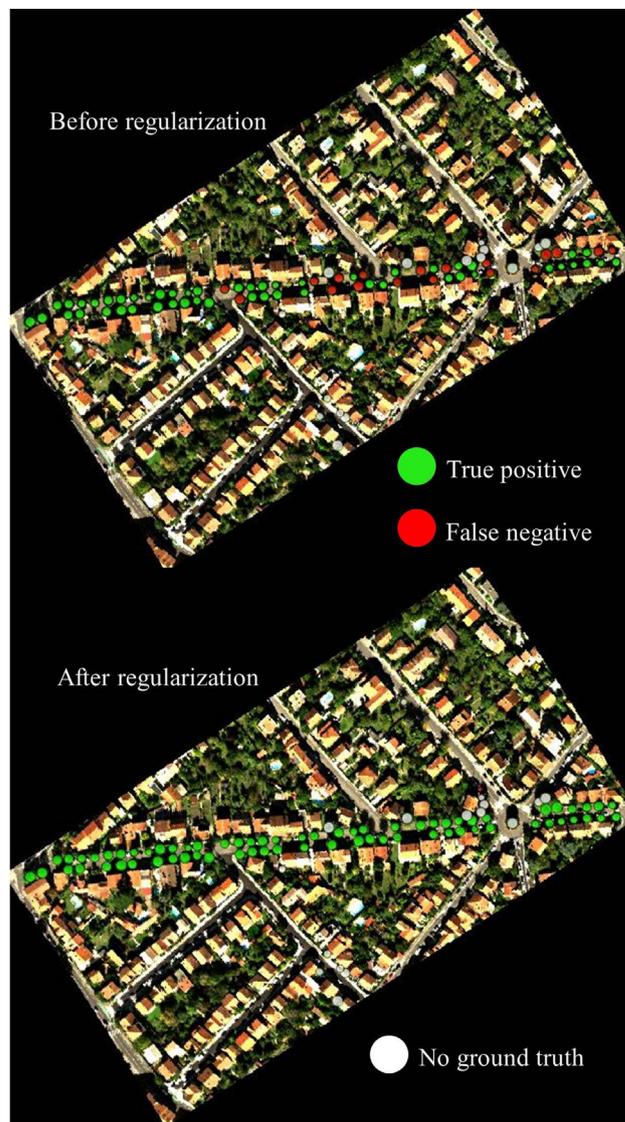


Fig. 16. Improvement of tree species classification in urban alignment thanks to the proposed method for the third case (whose alignment is mainly composed of *Platanus × hispanica* trees).

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  $U_{ij}(x)$  (Section 3.1.2, Eq. (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, Fig. 16 (bottom: after regularization) demonstrates the potential of the proposed method

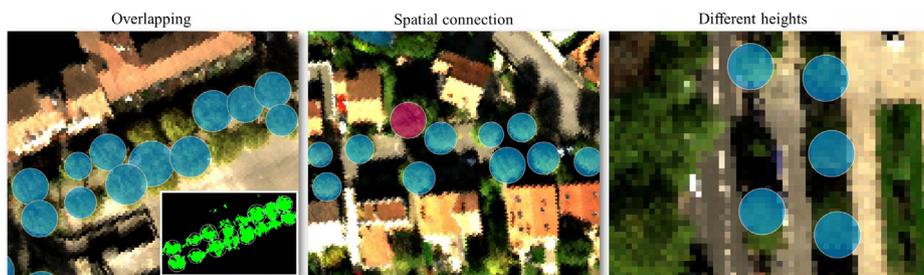


Fig. 15. Illustration of the limits of the proposed approach. Left: overlapping with an illustration of the corresponding high vegetation mask in the bottom right-hand corner, showing that the street trees are difficult to discern. Middle: spatial connection. Right: different heights.

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).

## 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 15p.p., 38p.p. and 65p.p. 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,  $\theta$  and  $h$ ). Without GIS but with  $\theta$  and  $h$ , the F score values become 88%, 79% and 62% ( $-3$ p.p.,  $+4$ p.p. and  $-23$ p.p.). It is thus more appropriate to exploit together the GIS data, the angle between the trees and the heights. Nevertheless, using only  $\theta$  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 these 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 Philippe Deliot and the French Mapping Agency for their participation in the UMBRA campaign. The Vaihing data set was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) (Cramer, 2010): [http://www.ifp.uni-stuttgart.de/dgpf/DKEP Allg.html](http://www.ifp.uni-stuttgart.de/dgpf/DKEP>Allg.html). Thanks to the members of the HY perspective 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.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.isprsjprs.2018.09.016>.

## References

- Adeline, K., 2014. Classification des matériaux urbains en présence de végétation éparse par télédétection hyperspectrale à haute résolution spatiale (Ph.D. thesis). Toulouse, ISAE.
- Adeline, K.R., Le Bris, A., Coubar, F., Briottet, X., Paparoditis, N., Viallefont, F., Rivière, N., Papelard, J.-P., David, N., Deliot, P., et al., 2013. Description de la campagne aéroportée UMBRA: étude de l'impact anthropique sur les écosystèmes urbains et naturels avec des images THR multispectrales et hyperspectrales. *Revue française de photogrammétrie et de télédétection* (202), 79–92.
- Alonzo, M., Bookhagen, B., Roberts, D.A., 2014. Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sens. Environ.* 148, 70–83.
- Alvarez-Uria, P., Körner, C., 2007. Low temperature limits of root growth in deciduous and evergreen temperate tree species. *Funct. Ecol.* 21 (2), 211–218.
- Ardila, J.P., Tolpekin, V.A., Bijker, W., Stein, A., 2011. Markov-random-field-based super-resolution mapping for identification of urban trees in vhr images. *ISPRS J. Photogramm. Remote Sens.* 66 (6), 762–775.
- Baraton, A., 2014. Mes trucs et astuces de jardinier. Flammarion.
- Bolund, P., Hunhammar, S., 1999. Ecosystem services in urban areas. *Ecol. Econ.* 29 (2), 293–301.
- Chatelain, F., Descombes, X., Zerubia, J., 2009. 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, pp. 221–234.
- Chen, Q., Baldocchi, D., Gong, P., Kelly, M., 2006. Isolating individual trees in a savanna woodland using small footprint lidar data. *Photogramm. Eng. Remote Sens.* 72 (8), 923–932.
- Chiesura, A., 2004. The role of urban parks for the sustainable city. *Landscape Urban Plann.* 68 (1), 129–138.
- Cramer, M., 2010. The dgpf-test on digital airborne camera evaluation—overview and test design. *Photogramm.-Fernerkund.-Geoinform.* 2010 (2), 73–82.
- Descamps, S., Descombes, X., Bechet, A., Zerubia, J., 2009. Flamingo detection using marked point processes for estimating the size of populations. *Traitement. Signal* 26 (2), 95–108.
- Descombes, X., Minlos, R., Zhizhina, E., 2009. Object extraction using a stochastic birth-and-death dynamics in continuum. *J. Math. Imaging Vision* 33 (3), 347–359.
- Doick, K.J., Peace, A., Hutchings, T.R., 2014. The role of one large greenspace in mitigating London's nocturnal urban heat island. *Sci. Total Environ.* 493, 662–671.
- Duchon, J., 1977. Splines minimizing rotation-invariant semi-norms in Sobolev spaces. In: *Constructive Theory of Functions of Several Variables*. Springer, pp. 85–100.
- Fassnacht, F.E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L.T., Straub, C., Ghosh, A., 2016. Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.* 186, 64–87.
- Fini, A., Frangi, P., Faoro, M., Piatti, R., Amoroso, G., Ferrini, F., 2015. Effects of different pruning methods on an urban tree species: a four-year-experiment scaling down from the whole tree to the chloroplasts. *Urban For. Urban Green.* 14 (3), 664–674.
- Gamal-Eldin, A., Descombes, X., Zerubia, J., 2010. Multiple birth and cut algorithm for point process optimization. In: *2010 Sixth International Conference on Signal-Image Technology and Internet-Based Systems (SITIS)*. IEEE, pp. 35–42.
- Gillner, S., Vogt, J., Tharang, A., Dettmann, S., Roloff, A., 2015. Role of street trees in mitigating effects of heat and drought at highly sealed urban sites. *Landscape Urban Plann.* 143, 33–42.
- Gomesa, M.F., Maillarda, P., 2014. 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. 9245, p. 92450X-1.
- Gupta, S., Weinacker, H., Koch, B., 2010. Comparative analysis of clustering-based approaches for 3-d single tree detection using airborne fullwave lidar data. *Remote Sens.* 2 (4), 968–989.
- Hadj, S.B., Chatelain, F., Descombes, X., Zerubia, J., 2010. 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)*, p. 1.
- Iovan, C., Boldo, D., Cord, M., 2008. Modélisation de la végétation en milieu urbain: Détection et caractérisation à partir d'images aériennes haute résolution couleur et infrarouge. *Revue française de photogrammétrie et de télédétection* (189), 17–27.
- Köhler, C.H., 2016. Airborne imaging spectrometer hypspec. *J. Large-scale Res. Facil. JLSRF* 2 (A93), 1–6.
- Leckie, D., Gougeon, F., Hill, D., Quinn, R., Armstrong, L., Shreenan, R., 2003. Combined high-density lidar and multispectral imagery for individual tree crown analysis. *Can. J. Remote Sens.* 29 (5), 633–649.
- MacFaden, S.W., O'Neil-Dunne, J.P., Royar, A.R., Lu, J.W., Rundle, A.G., 2012. High-resolution tree canopy mapping for New York city using lidar and object-based image analysis. *J. Appl. Remote Sens.* 6 (1) 063567-1.
- McPherson, E.G., van Doorn, N., de Goede, J., 2016. Structure, function and value of street trees in California, USA. *Urban For. Urban Green.* 17, 104–115.
- Miyajima, Y., Takahashi, K., 2007. Changes with altitude of the stand structure of temperate forests on Mount Norikura, Central Japan. *J. For. Res.* 12 (3), 187–192.
- Nagao, M., Matsuyama, T., Ikeda, Y., 1979. Region extraction and shape analysis in aerial photographs. *Comput. Graph. Image Process.* 10 (3), 195–223.

- Otsu, N., 1975. A threshold selection method from gray-level histograms. *Automatica* 11 (285–296), 23–27.
- Perrin, G., Descombes, X., Zerubia, J., 2005. A marked point process model for tree crown extraction in plantations. In: *IEEE International Conference on Image Processing, 2005. ICIP 2005*, vol. 1. IEEE, pp. 1–661.
- Perrin, G., Descombes, X., Zerubia, J., Boureau, J.-G., 2006. Forest resource assessment using stochastic geometry. In: *International Precision Forestry Symposium*.
- Poutier, L., Miesch, C., Lenot, X., Achard, V., Boucher, Y., 2002. Comanche and cochise: two reciprocal atmospheric codes for hyperspectral remote sensing. In: *2002 AVIRIS Earth Science and Applications Workshop Proceedings*.
- Pulighe, G., Lupia, F., 2016. Mapping spatial patterns of urban agriculture in Rome (Italy) using Google earth and web-mapping services. *Land Use Policy* 59, 49–58.
- Ramdani, F., 2013. Urban vegetation mapping from fused hyperspectral image and lidar data with application to monitor urban tree heights. *J. Geograph. Inform. Syst.* 5 (04), 404.
- Rol-Tanguy, F., Alba, D., Dragoni, M., Ter Minassian, H., Blancot, C., 2010. Essai de bilan sur le développement des arbres d'alignement dans paris, 2010. < <https://www.apur.org/fr/nos-travaux/bilan-developpement-arbres-alignement-paris> > (accessed: 24/10/17).
- Rouse Jr. J., Haas, R., Schell, J., Deering, D., 1974. Monitoring vegetation systems in the great plains with ERTS. In: *Third ERTS Symposium, NASA SP-351 I*. pp. 309–317.
- Sebestyen, D., Nemeth, M., Hangyal, R., Krizbai, L., Ember, I., Nyerges, K., Kolber, M., Kiss, E., Bese, G., 2008. Ornamental prunus species as new natural hosts of plum pox virus and their importance in the spread of the virus in Hungary. *J. Plant Pathol.* S57–S61.
- Seto, K.C., Güneralp, B., Hutyra, L.R., 2012. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proc. Nat. Acad. Sci.* 109 (40), 16083–16088.
- Souchon, J.-P., Thom, C., Meynard, C., Martin, O., Pierrot-Deseilligny, M., 2010. The IGN CAMv2 system. *Photogram. Rec.* 25 (132), 402–421.
- Vailshery, L.S., Jaganmohan, M., Nagendra, H., 2013. Effect of street trees on micro-climate and air pollution in a tropical city. *Urban For. Urban Green.* 12 (3), 408–415.
- Van Leeuwen, M., Coops, N.C., Wulder, M.A., 2010. Canopy surface reconstruction from a LiDAR point cloud using hough transform. *Remote Sens. Lett.* 1 (3), 125–132.
- Van Lieshout, M., 2000. *Markov Point Processes and their Applications*. World Scientific.
- Welch, R., 1982. Spatial resolution requirements for urban studies. *Int. J. Remote Sens.* 3 (2), 139–146.
- Wen, D., Huang, X., Liu, H., Liao, W., Zhang, L., 2017. Semantic classification of urban trees using very high resolution satellite imagery. *IEEE J. Sel. Top. Appl. Earth Obser. Remote Sens.* 10 (4), 1413–1424.
- Xiao, Q., Ustin, S., McPherson, E., 2004. Using aviris data and multiple-masking techniques to map urban forest tree species. *Int. J. Remote Sens.* 25 (24), 5637–5654.
- Yang, J., McBride, J., Zhou, J., Sun, Z., 2005. The urban forest in Beijing and its role in air pollution reduction. *Urban For. Urban Green.* 3 (2), 65–78.
- Zhen, Z., Quackenbush, L.J., Stehman, S.V., Zhang, L., 2015. Agent-based region growing for individual tree crown delineation from airborne laser scanning (als) data. *Int. J. Remote Sens.* 36 (7), 1965–1993.
- Zhen, Z., Quackenbush, L.J., Zhang, L., 2016. Trends in automatic individual tree crown detection and delineation-evolution of lidar data. *Remote Sens.* 8 (4), 333.
- Zhou, J., Proisy, C., Descombes, X., Hedhli, I., Barbier, N., Zerubia, J., Gastellu-Etchegorry, J.-P., Couteron, P., 2010. Tree crown detection in high resolution optical and lidar images of tropical forest. In: *Remote Sensing. International Society for Optics and Photonics* pp. 78240Q–78240Q.