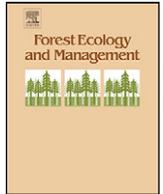




Contents lists available at [ScienceDirect](#)

## Forest Ecology and Management

journal homepage: [www.elsevier.com/locate/foreco](http://www.elsevier.com/locate/foreco)



# Multiple attribute decision making for individual tree detection using high-resolution laser scanning

Giovanni Forzieri<sup>a,b,\*</sup>, Leonardo Guarnieri<sup>c</sup>, Enrique R. Vivoni<sup>b</sup>, Fabio Castelli<sup>a</sup>, Federico Preti<sup>c</sup>

<sup>a</sup> Dipartimento di Ingegneria Civile e Ambientale, University of Florence, Italy

<sup>b</sup> Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology, Socorro, NM, USA

<sup>c</sup> Dipartimento di Ingegneria Agraria e Forestale, University of Florence, Italy

### ARTICLE INFO

#### Article history:

Received 11 June 2009

Received in revised form 2 September 2009

Accepted 2 September 2009

#### Keywords:

Single-tree identification

Forest monitoring

Image segmentation

LiDAR

Remote sensing

Decision making

Simple additive weighting

### ABSTRACT

A canopy height model (CHM) is a standard LiDAR-derived product for deriving relevant forest inventory information, including individual tree positions, crown boundaries and plant density. Several image-processing techniques for individual tree detection from LiDAR data have been extensively described in literature. Such methods show significant performance variability depending on the vegetation characteristics of the monitored forest. Moreover, over regions of high vegetation density, existing algorithms for individual tree detection do not perform well for overlapping crowns and multi-layered forests. This study presents a new time and cost-efficient procedure to automatically detect the best combination of the morphological analysis for reproducing the monitored forest by estimating tree positions, crown boundaries and plant density from LiDAR data. The method needs an initial calibration phase based on multi attribute decision making-simple additive weighting (MADM-SAW). The model is tested over three different vegetation patterns: two riparian ecosystems and a small watershed with sparse vegetation. The proposed approach allows exploring the dependences between CHM filtering and segmentation procedures and vegetation patterns. The MADM architecture is able to self calibrate, automatically finding the most accurate de-noising and segmentation processes over any forest type. The results show that the model performances are strongly related to the vegetation characteristics. Good results are achieved over areas with a ratio between the average plant spacing and the average crown diameter (TCI) greater than 0.59, and plant spacing larger than the remote sensing data spatial resolution. The proposed algorithm is thus shown a cost effective tool for forest monitoring using LiDAR data that is able to detect canopy parameters in complex broadleaves forests with high vegetation density and overlapping crowns and with consequent significant reduction of the field surveys, limiting them over only the calibration site.

© 2009 Elsevier B.V. All rights reserved.

## 1. Introduction

Single tree-level forest information plays a crucial role in hydrological, meteorological, and ecological applications sensitive to vegetation evolution at local and regional scales. Some of these applications consist of monitoring forest regeneration and damage evaluation (Chen et al., 2006), biomass and carbon stock estimation (Popescu and Wynne, 2004), wildfire simulation models (Finney, 1998), quantifying woodland structure and habitat quality for birds (Hinsley et al., 2002) flow resistance models for hydraulic roughness estimation (e.g., Petryk and Bosmajian, 1975; Thompson and Roberson, 1976; Kouwen and Fathi-Moghadam, 2000;

Järvelä, 2004; Baptist et al., 2007) and for atmospheric turbulent flux modeling (Eagleson, 2002). Tree crown boundaries, individual tree positions and the spatial variability of plant density represent important forest parameters. Traditional methods of investigating such parameters include labor-intensive forest inventories and complex sampling designs (Shivers and Borders, 1996). Moreover, the existing methods are time-consuming, subjective and more applicable primarily to small areas (Avery and Burkhart, 1994).

New technologies, such as remote sensing and new computer vision algorithms, have enabled the introduction of semi-automated forest assessments based on delineation of single tree crowns and individual tree detection. Several large area inventories have been achieved with very high-resolution remote sensing using automated pattern recognition (e.g., Gougeon and Leckie, 2003). However, as the spectral and textural characteristics derived from remotely sensed images are not directly related to tree morphology, these methods can produce inaccurate estimates

\* Corresponding author at: Dipartimento di Ingegneria Civile e Ambientale, University of Florence, Italy. Tel.: +39 055 4796458; fax: +39 055 495333.

E-mail address: [gioforz@dicea.unifi.it](mailto:gioforz@dicea.unifi.it) (G. Forzieri).

(Gong et al., 2002). In recent years, airborne LiDAR technology has been used to detect individual tree crowns and biophysical characteristics (Andersen et al., 2005; Hyyppä et al., 2001; Popescu and Wynne, 2004; Næsset and Økland, 2002). Compared with passive images, LiDAR has the advantage of directly measuring the three-dimensional coordinates of canopies, providing information for crown geometric shapes. Despite LiDAR-derived products can never reveal certain tree patterns measured in the field, such as suppressed trees, grouped trees in dense forests and understorey (Popescu and Zhao, 2008; Zhao et al., 2009), they represent powerful tools for deriving relevant forest inventory information.

Several methods for tree detection developed for optical imagery have been extended to LiDAR data. An efficient method for automated segmentation is the morphological watershed algorithm (Vincent and Soille, 1991; Soille, 1999), and its recent variants (Osma-Ruiz et al., 2007; Rambabu and Chakrabarti, 2007). Watershed segmentation has also been used to detect tree crown boundaries and positions through several formulations: classical approach (Andersen et al., 2005), marker-controlled watershed segmentation (Chen et al., 2006) and watershed using the transformation distance (Kwak et al., 2007). Such methods show significant performance variability depending on the vegetation characteristics of the monitored forest. To avoid “over-segmentation” in the application of a watershed algorithm to LiDAR data, the canopy height model (CHM, a digital crown height model) is usually pre-processed using different filters such as Gaussian (Dralle and Rudemo, 1996; Persson et al., 2002) or convolution (Hyyppä et al., 2001), with a static or variable window (Popescu and Wynne, 2004). Despite the fact that filter and window size influence significantly the CHM smoothing process (Chen et al., 2006), the linkage between filtering parameters and segmentation performances is still poorly explored.

Although the use of LiDAR data for producing CHM estimates has produced encouraging results over coniferous forests, similar performances have not been assessed over broadleaved woodlands or multi-layered forest canopies, characterized by a complex plant morphology with overlapping crowns (Maltamo et al., 2004). As a result, new methodologies for detecting tree crown characteristics and positions need to be validated over a range of different forest conditions.

The overall goal of this study was to develop a new time and cost-efficient procedure to automatically detect the best combination of the image morphological analysis for reproducing the monitored forest by estimating tree positions, crown boundaries and plant density variability using airborne LiDAR data. The model needs an initial calibration phase based on multiple attribute decision making (MADM) simple additive weighting method (SAW) (Hwang and Yoon, 1981). The main novelty is that the MADM architecture can be easily applied on several forest patterns detecting automatically, in an ensemble of the most used segmentation algorithms and de-noising filters, the optimal image processing depending on the canopy characteristics of the investigated area. The model is tested over three different vegetation patterns: two riparian ecosystems along the Serchio and the Sieve River floodplains (Tuscany Region, Italy) and a small watershed with sparse vegetation located in the Sevilleta National Wildlife Refuge (New Mexico, USA).

## 2. Methods

### 2.1. Study areas

The study area location is important for the description of the different forest environments that are considered necessary to calibrate the model. The first study area is located along the Serchio River in the Township of Galliciano, 28 km north of the city of Lucca

(Tuscany Region, Italy) (Fig. 1A). The site is made up of a small area ( $\approx 0.0025 \text{ km}^2$ ) containing a fluvial island with a very high vegetation density downstream of the Campia bridge. This mature woodland results from the evolution of riparian ecosystems that have not been altered by flood events for some time. The most common species are Lombardy poplar (*Populus nigra*), willow (*Salix alba*), black alder (*Alnus glutinosa*), and field maple (*Acer campestre*). The vertical distribution of tree crowns indicates a non-coetaneous biplanar riparian formation, originating from natural dissemination, formerly managed via cutting, and currently in the pole forest evolutionary stage.

The second study area is situated along the Sieve River, 20 km north-east of the city of Florence (Tuscany Region, Italy) (Fig. 1B). The site is composed of a small area ( $\approx 0.001 \text{ km}^2$ ) set on the river bank and is characterized by high vegetation density with overlapping tree crowns. Common species are Lombardy poplar (*P. nigra*), black locust (*Robinia pseudoacacia* L.) while the undergrowth is comprised of field maple (*A. campestre* L.), elder (*Sambucus nigra* L.), cornel (*Cornus mas*) and bramble (*Rubus ulmifolius* S.). The vertical distribution of the crowns indicates a mixed riparian formation of the biplanar type, originating from natural dissemination and currently in a high forest evolutionary stage. In this phase, social differentiation is reduced, conditioning first the longitudinal and then the diametrical growth. Stalks are larger in diameter and are less flexible to winds, with the result that the crown is limited to the upper part of the tree.

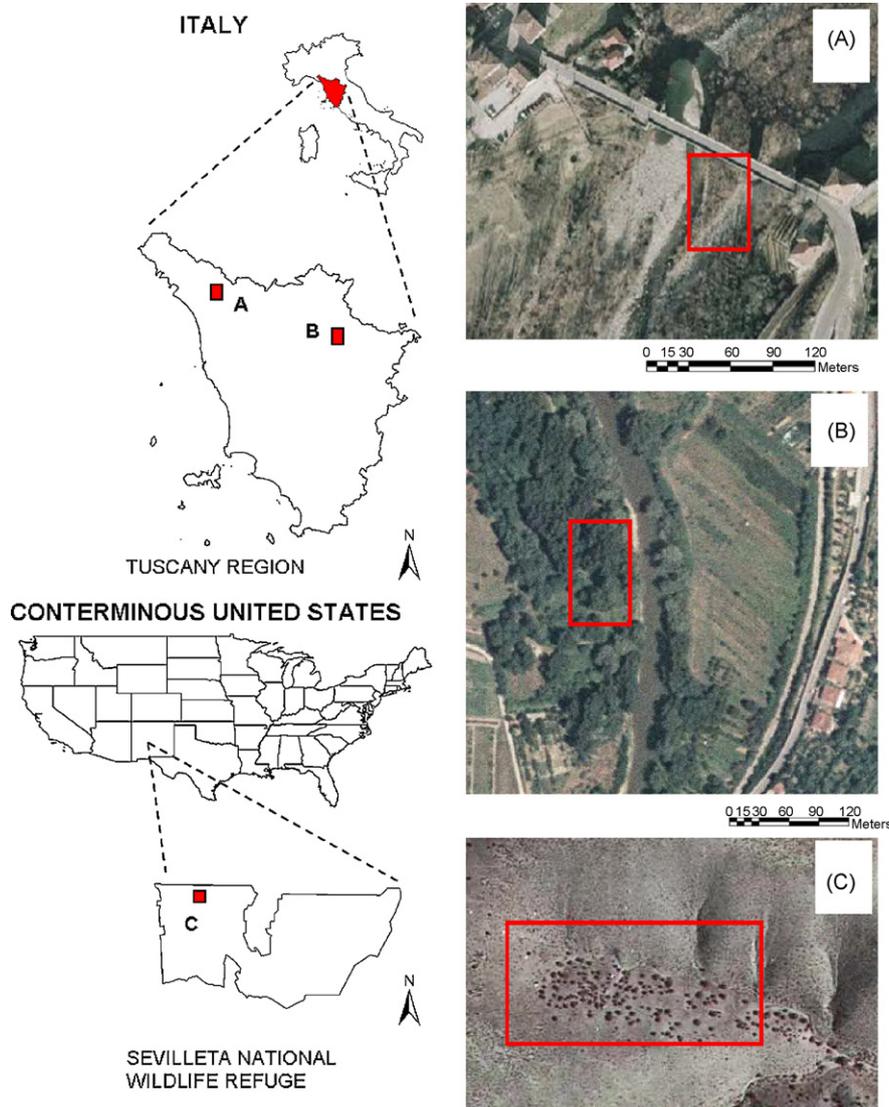
The third test area is located in the northwest part of the Sevilleta National Wildlife Refuge (SNWR) site in central New Mexico, 90 km south of Albuquerque (Fig. 1C). The site comprises a small ( $\approx 0.1 \text{ km}^2$ ) first-order catchment dissected by an east flowing ephemeral channel giving rise to opposing north and south-facing slopes and an east facing headslope (Gutiérrez-Jurado et al., 2007). In the study basin, opposing hillslopes are characterized by marked differences in ecosystem composition and soil profile properties, with the south-facing slope consisting of creosote bush (*Larrea tridentata*) and the north-facing hillslope dominated by one seed juniper (*Juniperus monosperma*) (Gutiérrez-Jurado et al., 2006) with sparse or open canopies vegetation.

Fig. 2 shows some pictures of the canopy for each study area and highlights the significant variability of vegetation density between the three investigated areas, as afore-mentioned.

### 2.2. Datasets

Light detection and ranging (LiDAR) is a remote sensing technology that can provide highly accurate measurements of both the forest canopy and ground surface. Airborne laser scanning systems can provide terrain elevation data for open areas with a vertical accuracy of  $\sim 15 \text{ cm}$  and  $23 \text{ cm}$  under conifer forest canopies (Reutebuch et al., 2003). Laser scanning is based on distance measurements and precise orientation of these measurements between a sensor, whose position is known, and a reflecting object, whose position is unknown. The orientation and the position of the sensor at the time of each emitted pulse is known through the use of an integrated inertial navigation system and a differential global positioning system. By classifying the laser pulses iteratively into terrain and non-terrain returns, it was possible to produce a digital terrain model (DTM) and a digital surface model (DSM) (Brandtberg et al., 2003). The difference between the DSM and DTM models is called, in this study, canopy height model, and represents a 3D representation of the tree heights within the target forest area. We derive canopy height models with 1-m spatial resolution for each study area.

The area of the calibration sites mainly depends on the vegetation characteristics. To capture the variability of vegetation



**Fig. 1.** Study areas. The red boxes show the monitored test sites. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

patterns, forest scenes with high spatial heterogeneity obviously need larger sample areas than more homogeneous stands. Geostatistical explorative analysis represents useful tools to verify the adequacy of the sample area. In our study we chose the minimum sample area above which the coefficient of variation of forest attributes saturates (Piussi, 1994). Such analysis was performed independently for each study area. As a thumb rule, based on the standard silvicultural practise, the ratio between the forest area to be analyzed and the sample area is  $\approx 15$ .

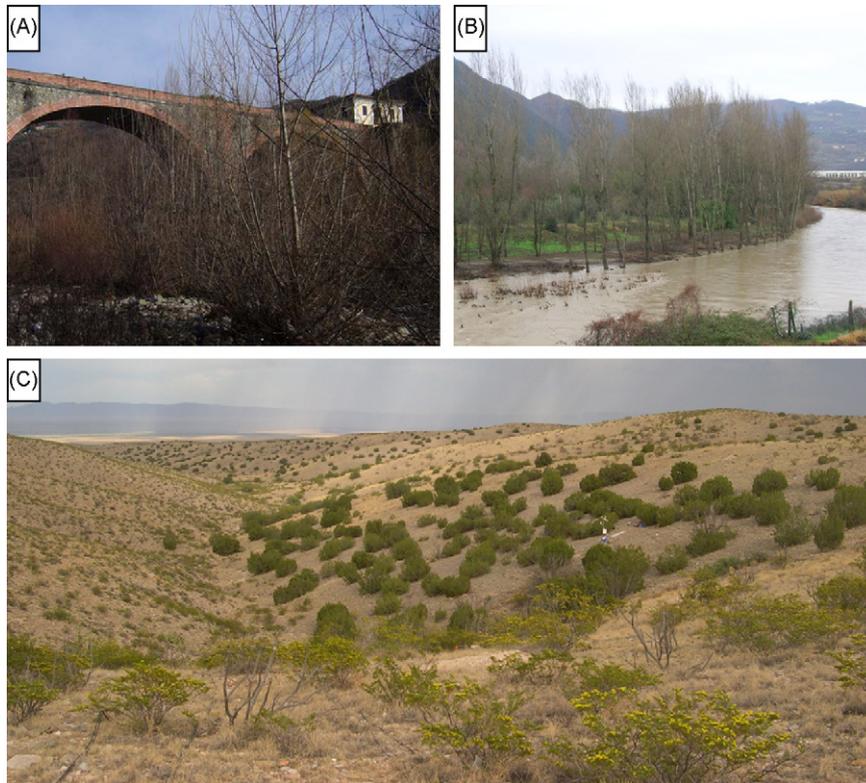
The vegetation within the calibration sites was characterized using a total Nikon station DTM-A2LG with a reflecting prism and forestry instrumentation (vertex, hypsometer with lens, dendrometric tripod). The estimated tree parameters were: identity code, scientific name (genus, species), position in absolute coordinates, diameter at the base, diameter at 1.3 m, tree height, crown insertion height, planimetry of the crown diameter (average radius via measurement in the four main directions).

### 2.3. Theoretical formulation of MADM-SAW

Multiple attribute decision making (MADM) techniques are used to evaluate a finite number of alternatives ( $i$ ) with multiple

attributes, with the goal of identifying a preferred option or to distinguish acceptable possibilities. An attribute is a measurable quantity whose value reflects the degree to which a particular objective is achieved (Chankong and Haines, 1983); therefore each alternative has a performance rating for each attribute. The best solution is driven by the attributes ( $j$ ), which are calculated based on a predefined decision rule to detect the characteristics of the best solution.

Simple additive weighting (SAW) is a popular multiple attribute decision technique (Malczewski, 1999; Hwang and Yoon, 1981) based on the weighted average. An evaluation score ( $x_{i,j}$ ) is calculated for each alternative and attribute based on the decision rule. If attribute scores are calculated on different scales, these must be normalized to a common dimensionless unit before using the SAW method. Different approaches for the normalization process exist (e.g., Voogd, 1983; Chakraborty and Yeh, 2007). In the following, we will use the notation  $x'_{i,j}$  as the normalized score of the  $i$ th alternative with respect to the  $j$ th attribute. The weights ( $W$ ) of relative importance are assigned for each attribute by using the relation  $\sum_{i=1}^M w_i = 1$ , where  $M$  is the total number of attributes. Subsequently, an evaluation rank is calculated for each alternative by multiplying the normalized scores with the relative importance



**Fig. 2.** Vegetation density on the three investigated study areas. Serchio floodplain (A), with very high vegetation density; Sieve floodplain (B) with high vegetation density and overlapping crowns; Sevilleta (C), a small watershed with sparse vegetation (courtesy of Hugo Gutierrez-Jurado).

weights and by summing the obtained products. The SAW method evaluates each alternative ( $A_i$ ) as:

$$A_i = \sum_j^M x'_{i,j} w_j, \quad (1)$$

where  $w_j$  is the weight respect to the  $j$ th attribute. If the attributes are considered of similar “benefit” ( $B$ ) or “cost” ( $C$ ), the best solution ( $K$ ) is detected by analyzing the maximum or the minimum rank between all the analyzed alternatives as:

$$\begin{aligned} K &= \max(A_i) & j \in B \\ K &= \min(A_i) & j \in C \end{aligned} \quad (2)$$

#### 2.4. Methodology

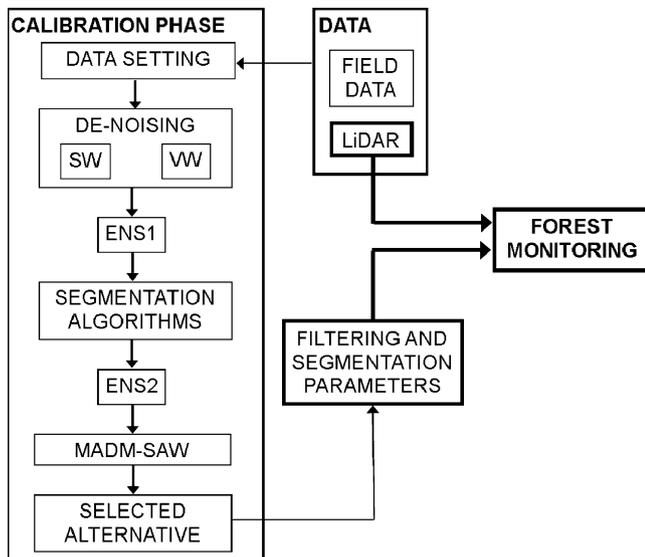
The proposed method includes an initial calibration phase performed over a restricted sample area (calibration site) and a subsequent analysis extended over all the monitored forest. The calibration phase needs LiDAR data and field surveys over the calibration site, with a vegetation pattern representative of all the monitored forest. The use of the observed dataset allows to drive the selection of the best forest modeling, obtained via image-processing techniques, by comparing observed and simulated forest scenarios. The methodology encompasses several existing image-processing techniques well documented and extensively described in the literature. The proposed MADM-SAW method systematically detects the best combination of such morphological analysis for reproducing the monitored forest by adaptively selecting the most accurate de-noising and segmentation processes over each forest. Then, the proposed MADM architecture drives the selection of existing algorithms. Optimal parameters and suitable algorithms to reproduce the calibration site are selected through sequential processes, and then applied over all

the LiDAR extensions. The methodology allows significant reduction of the field surveys, limiting them to the calibration site only.

The procedure, written in MATLAB, interfaces with input and output files in GIS format, facilitating use of different data sources. The canopy height model is transformed to a gray level image (Andersen et al., 2005) and analyzed through image-processing techniques. The object recognition system, in the calibration phase, performs four sequential steps: (1) data setting, (2) de-noising filters, (3) segmentation algorithms and (4) MADM-SAW method (Fig. 3). In the setting phase, LiDAR and field data are pre-processed and converted to a digital format. In the de-noising step, the CHM is filtered using different image-processing operations, generating a first ensemble (ENS1) of CHM-derived digital surfaces. In the segmentation phase, the previous set is analyzed through different watershed segmentation algorithms producing a new ensemble of segmented images (ENS2), with each pattern representing a different alternative for individual tree positions and crown boundary detection. The MADM-SAW method detects the best alternative (the best de-noising/segmentation process) of the ENS2 ensemble. In order to investigate the model capability to explore the spatial variability of the plant density, simulated and observed scenarios are compared. In the following paragraphs, the procedure steps are explained in more detail.

##### 2.4.1. Data setting

In order to avoid the influence of the adjacent plants around the test site, a gridded mask is digitized by visual inspection to include all the sampled trees. In addition, the CHM is masked over a subset window to reduce the computing time and boundary effects that could compromise the segmentation performance. Then, the clipped CHM is inverted to show each crown as a local depression (Andersen et al., 2005) and the vector-based field data (tree position, height, crown diameter) are converted to grid format to apply the image-processing techniques. We used ground mea-



**Fig. 3.** Flow diagram of the proposed MADM-SAW approach. The proposed method includes an initial calibration phase and a subsequent analysis extended over all the monitored forest. The calibration phase needs LiDAR data and field surveys over a restricted sample area (calibration site), with a vegetation pattern representative of all the monitored forest. The object recognition system, in the training phase, performs four sequential steps: data setting, de-noising filters, segmentation algorithms and MADM-SAW method. Then, suitable parameters and algorithms to reproduce the calibration site are selected and applied over all the LiDAR extensions.

surements to investigate the correlation between crown diameter and tree height through a simple regression analysis based on the following (Popescu et al., 2002):

$$C_D = \alpha_0 + \alpha_1 h^2, \quad (3)$$

where  $C_D$  is the crown diameter,  $h$  is the tree height and  $\alpha_0, \alpha_1$  are the regression parameters. We performed the afore-mentioned regressive analysis on 85, 29 and 97 monitored trees, for the Serchio, Sieve and Sevilleta areas respectively. The carried out equations could be improved by providing allometric relationships depending on the plant species and the vertical distribution of branches and foliage elements (Baker, 1950; Eagleson, 2002). Using the LiDAR-derived height and the  $C_D-h$  relationship, we can generate a *pixel-oriented crown diameter* for each cell, that represents the window size in the filtering procedure, explained in more detail in the following paragraph.

#### 2.4.2. De-noising processes

To reduce local noise and suppress irrelevant local maxima in the treetop detection, the inverted CHM is filtered using different image-processing techniques. We systematically applied linear and non-linear spatial filters such as Prewitt (PF), Sobel (SF), Average (AF), Disk (DF), Laplacian (LF), Unsharp (UF), Gaussian (GF) and Log (LoF), by varying the filtering parameters (see Table 1) and the window size. The neighborhood processing analyses the input-image both with static and variable window (SW and VW). In the first case, the inverted CHM is transformed iteratively using the above-mentioned filters and a static window. In the second case, a variable window size for each cell is extracted by the *pixel-oriented crown diameter*, through the  $C_D-h$  relationship previously explained, and applied for the filtering processes. In the de-noising phase, a first ensemble of filtered images (ENS1) is generated as an input for the subsequent step.

#### 2.4.3. The morphological watershed algorithm

In the classical formulation, the watershed algorithm (Vincent and Soille, 1991; Soille, 1999) finds the boundaries of basins, or

**Table 1**

Ranges of the explored filtering parameters ( $par_i$ ) for different image-processing analysis for the static and variable windows. Max and Min indicate the extreme values for each parameter. Note some filters do not need explicated parameter (/ symbol).

Filter	Static window			
	$par_1$		$par_2$	
	Min	Max	Min	Max
PF	/	/	/	/
SF	/	/	/	/
AF	3	9	/	/
DF	1	7	/	/
LF	0	1	/	/
UF	0	1	/	/
GF	3	9	0.1	3
LoF	3	9	0.1	3

Filter	Variable window		
	$par_1$	$par_2$	
		Min	Max
AF	$C_D$	/	/
DF	$C_D$	/	/
GF	$C_D$	0.1	3
LoF	$C_D$	0.1	3

watersheds, within a surface model. If the LiDAR-derived canopy height model is inverted, each tree crown is analogous to a small basin and it is possible to delineate the local depression (Andersen et al., 2005). In this phase, we applied six different watershed segmentation algorithms on the ensemble of filtered images (ENS1), generating a new ensemble of segmented images (ENS2), each representing an alternative for tree position and crown boundary patterns.

All the watershed transforms are processed with 8- and 4-connected neighborhoods. The six watershed segmentation algorithms are: (1) *morphological watershed algorithm* (WS) (Vincent and Soille, 1991; Soille, 1999), (2) *watershed segmentation using the distance transform* (WSDT) (Gonzalez et al., 2004), (3) *watershed segmentation using gradients* (WSG) (Gonzalez et al., 2004), (4) *marker-controlled watershed segmentation* (MCWS) (Gonzalez et al., 2004), (5) *marker-controlled watershed segmentation, modified* (MCMWS) (Chen et al., 2006) and (6) *principal curvature-based region detector* (PCBR) (Deng et al., 2007). To detect robust watershed regions, the principal curvature image is processed by combining a grayscale morphological close with a new “eigenvector flow” hysteresis threshold. In Table 2, we list the segmentation parameters used in this study.

#### 2.4.4. Decision making process on forest scenarios

In the MADM-SAW phase, the best simulated forest scenario is identified. The ensemble of derived segmented images (ENS2) represents the available alternatives ( $i$ ) generated through different de-noising and segmentation analyses. We use a three-parameter decision rule and, in order to calibrate the model limited over the ground-monitored area, the three attributes are estimated only for the simulated trees within the envelope mask. Then, the score ( $x_{i,j}$ ) is calculated for each alternative and for each attribute based on the following relationships.

$$x_{i,1} = \frac{|n_{sim} - n_{obs}|}{n_{obs}} \quad (4)$$

$$x_{i,2} = RMSE(V_{sim} - V_{obs})$$

$$x_{i,3} = RMSE(D_{C,sim} - D_{C,obs})$$

where the *sim* means simulated through segmentation algorithms, *obs* means observed by field surveys,  $n$  is the number of trees,  $V$  is the tree position,  $D_C$  is the crown diameter and  $RMSE$  is the root mean square error. All of the afore-mentioned parameters are

**Table 2**  
Ranges of the six explored segmentation parameters ( $par_i$ ) for different watershed-derived algorithms.  $q_{max} = \text{round}(\log 2(\min(2n, 2m)) - 3)$ , where  $n$  and  $m$  represent row and column number of the CHM. In brackets, the notation used by the original authors is presented for each parameter. Note some segmentation algorithms do not need explicated parameter (/symbol).

Watershed-derived algorithm	$par_1$ (conn)	$par_2$ (h)	$par_3$ [se = strel('disk', $par_3$ )]	$par_4$ [se2 = strel(ones( $par_4$ ))]	$par_5$ (q)	$par_6$ (MP <sub><i>j</i></sub> )
WS	[4, 8]	/	/	/	/	/
WSDT	[4, 8]	/	/	/	/	/
WSG	[4, 8]	/	/	/	/	/
MCWS	[4, 8]	3	/	/	/	/
MCMWS	[4, 8]	/	1	1	/	/
PCBR	[4, 8]	/	[2, 5]	/	[1, $q_{max}$ ]	[2, 5]

directly detected from LiDAR data. The centroid of each segmented object is assumed to be the simulated stem position. The values of  $x_2$  are estimated as the mean of the distances between each simulated position and the closest observed stem position and the values of  $x_3$  are calculated as the mean of the differences between each simulated crown and the closest observed crown. The three selected attributes depend on the number of trees, tree positions and crown diameters. We retain that these parameters mainly influence the spatial variability of vegetation density, which in turn significantly affects LiDAR-based segmentation algorithms for individual tree identification. The use of a more extended set of vegetation attributes, including for example crown base height and stem diameter in addition to the afore-mentioned attributes, could provide further improvements giving an enhanced representation of the accuracies of the segmentation results.

To speed up the ranking process, a screening procedure is applied by taking into account only the simulated scenarios with  $0.2n_{obs} < n_{sim} < 3n_{obs}$ . As the scores for the criteria are calculated on different measurement scales, these are normalized using the formula proposed by Chakraborty and Yeh (2007):

$$x'_{i,j} = \frac{x_{max,j} - x_{i,j}}{x_{max,j} - x_{min,j}}, \quad (5)$$

where  $i$  is the  $i$ th alternative (simulated scenario) and  $j$  is the  $j$ th considered forest attribute. The weights of relative importance ( $W$ ) are assigned depending on the attributes that better describe the monitored forest scenario. For each different forest scenario, the optimal weight is obtained through an iterative trial and error procedure by simple visual inspection of the results in the necessary calibration phase. The final rank is obtained by summing the products between normalized attributes and weights. Since the normalized attributes can be considered as a “benefit”, the best solution is found through the following relationship:

$$(K, ix) = \max(A_i), \quad (6)$$

where  $K$  is the higher score obtained and  $ix$  is the index of the best-segmented image, which identifies the selected de-noising/segmentation procedure. The identified best de-noising/segmentation process can be applied over more extended areas with the same vegetation pattern. The procedure exports, as the final result, two output files in GIS-compatible vector format: the crown boundaries and the tree positions.

#### 2.4.5. Spatial variability of plant density

The spatial variability of plant density represents a factor for assessing the model accuracy. The simulated (and observed) density is estimated for each stem position ( $i$ ) by counting the number of trees localized within the circular area  $C$  with predefined radius:

$$d_i = \frac{\sum_{k \in C_i} p_k}{C} \begin{cases} p_k = 1 \Leftrightarrow k = \text{stem position} \\ p_k = 0 \text{ otherwise} \end{cases}, \quad (7)$$

where  $d_i$  is the plant density calculated at the stem position  $i$  and  $p_k$  is the cell included in the circular area ( $C$ ). The disk radius depends on the forest characteristics and its size should explain the spatial variability of the biophysics properties. We take the radius to be equal to the average plant spacing estimated by field measurements, using:

$$R = \left( \frac{A_{TOT}}{\pi n_{obs}} \right)^{1/2}, \quad (8)$$

where  $A_{TOT}$  is the total area of the test site and  $n_{obs}$  is the number of observed trees. The local density parameter is interpolated using a cubic 2D method, resulting in an observed and simulated plant density maps. For each study area we calculated the RMSE of the spatial variability of the plant density by using the following relationship:

$$RMSE = \sqrt{\frac{\sum_{i \in D} (d_{i,obs} - d_{i,sim})^2}{N}}, \quad (9)$$

where  $d_i$  is plant density of the  $i$ th pixel,  $D$  is the domain delineated by the envelope mask,  $sim$  means derived from the simulated plant density map,  $obs$  means derived from the observed plant density map,  $N$  is the total number of pixels.

### 3. Results

#### 3.1. Linkage between morphological analysis and performance accuracies

Attribute scores ( $x_{i,j}$ ) are indicators on the model capacity to reproduce observed forest scenarios. To explore the role played by the weighting scheme, the attribute estimation of the best alternative is obtained by simulating four cases (Table 3). The first three cases use vectors of canonical base ([1 0 0]; [0 1 0]; [0 0 1]), while the fourth uses weights calibrated for each forest scenario. Values show a large performance variability in the same

**Table 3**  
Attribute values of the selected image-processing techniques for each study area using four different weightings. The notation mentioned in the text is used:  $x_1 = |n_{sim} - n_{obs}|/n_{obs}$ ,  $x_2 = RMSE(V_{sim} - V_{obs})$ ,  $x_3 = RMSE(D_{C,sim} - D_{C,obs})$ .

	$W$	$X_1$	$X_2$ (m)	$X_3$ (m)
Serchio	[1.00 0.00 0.00]	0.4824	1.24	3.45
	[0.00 1.00 0.00]	0.7647	0.84	2.53
	[0.00 0.00 1.00]	0.6824	1.12	1.43
	[1.00 0.00 0.00]	0.4824	1.24	3.45
Sieve	[1.00 0.00 0.00]	0	3.75	6.85
	[0.00 1.00 0.00]	0.6897	1.54	5.23
	[0.00 0.00 1.00]	0.6552	2.40	3.26
	[0.50 0.25 0.25]	0.0690	3.33	4.84
Sevilleta	[1.00 0.00 0.00]	0.0103	5.64	10.67
	[0.00 1.00 0.00]	0.3505	2.45	108.66
	[0.00 0.00 1.00]	1.7216	4.56	5.83
	[0.60 0.40 0.00]	0.1340	2.84	3.00

study area and highlight the improvements obtained through the suggested weight settings. Despite the proposed method by selecting the suitable weight setting inevitably introduces subjectivity in the proposed procedure, such suggested weight settings represent driver values to apply the MADM-SAW method to any forest scenes with vegetation characteristics similar to the monitored study areas. A systematic approach to identify the optimal weight setting could represent a further improvement of the proposed methodology. Differences between the model accuracies at the three areas are also evident, showing a dependence between performance and forest characteristics.

Fig. 4 presents model results at the different test sites: Serchio (A), Sieve (B) and Sevilleta (C). To evaluate the image-processing techniques, the final scores ( $A_i$ ) of the simulated scenarios are related to the filtering window and filtering and segmentation procedures. It is evident that the image-processing methods produce different results on the same vegetation pattern, showing significant variability in the average score, variance and extreme values. This highlights the role played by the morphological parameters and the usefulness in estimating the best alternative through a systematic MADM-SAW procedure. Fig. 4 also indicates that the image-processing techniques provide different performances on each forest scenarios, since the morphological analysis depends on the vegetation characteristics. As a result, applying the same procedures on different forests could over or under-estimate vegetation characteristics. This suggests that the principal usefulness of the proposed approach is the capability to find automatically the best morphological analysis for distinct forest types.

In comparing the monitored sites, it is clear that the variability between different image-processing analyses depends on the vegetation characteristics. Forest scenarios with very high vegetation density (Fig. 4A) show a higher rank variability, while forest scenarios with overlapping crowns and sparse vegetation (Fig. 4B and C) show similar performances. The filtering/segmentation parameters and suggested weights producing the best simulations for each different vegetation patterns are shown in Table 4, which would be useful for sites with similar vegetation patterns to the studied areas.

### 3.2. Distributed model accuracy

To analyze the distributed accuracy, we compare simulated and observed density maps by simple subtraction and calculate the relative root mean square error (RMSE), referred to as the *spatial variability* (or  $x_4$ ). Figs. 5–7 show the spatial variability of the plant density, the observed and simulated tree positions (in red triangles and black circles, respectively), the tree crown boundaries (in gray polygons) and the envelope masks (in black line) for each study site, Serchio, Sieve and Sevilleta, respectively. In the Serchio and Sieve areas, the tree crowns are strongly overlapping and thus the simulated canopy boundaries are not clearly displayed. In the following, we discuss the distributed model accuracy for each site separately.

In the Serchio test site (Fig. 5), several computational problems, related to the tree location, prevent an accurate reproduction of the monitored forest. In particular, boundary effects generated by the bridge in close proximity the study area influence the results in the de-noising and segmentation processes. In addition, the observed (ground) vegetation density is characterized by near-meter plant spacing, which cannot be captured through LiDAR sensor with 1-m spatial resolution. The spatial variability map shows high values and model inaccuracy where the observed plants are in close proximity ( $RMSE = 4.32 \#m^2$ ).

In the Sieve site (Fig. 6), the individual tree and crown boundary detection is difficult due to the overlapping canopies and complex plant architecture. In a broadleaved forest, many treetops merge

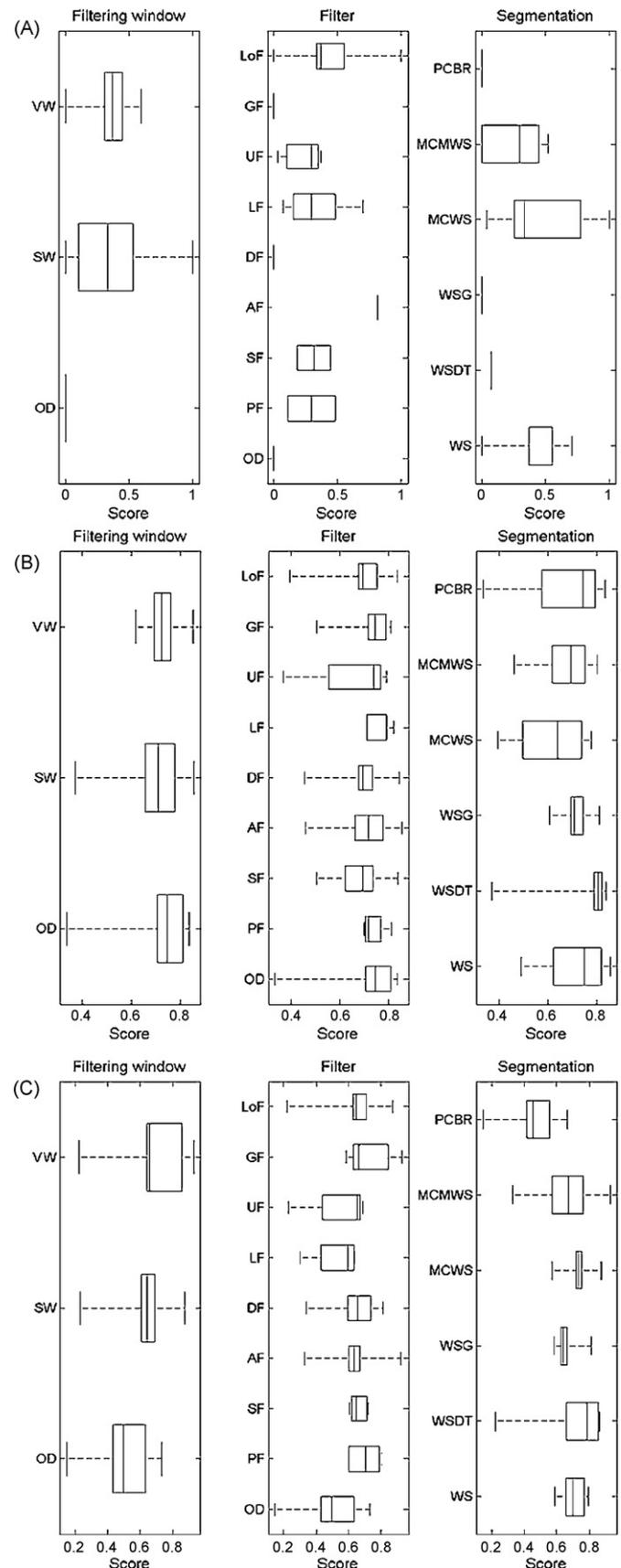
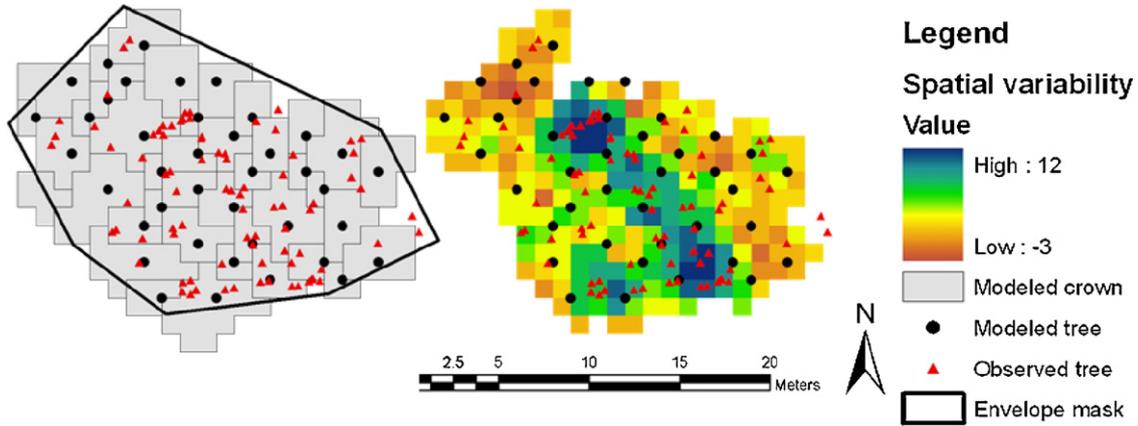


Fig. 4. Model results for different areas (A, B, C are Serchio, Sieve and Sevilleta test sites) by relating the score ( $A_i$ ) of simulated scenarios with the filtering window, filtering and segmentation procedures. The boxes have vertical lines at the lower quartile, median, and upper quartile values. The whiskers extending from each end of the box show the extent of the remaining data. The notation OD means original data without filtering.

**Table 4**  
Filtering and segmentation parameters and suggested weights for monitoring vegetation patterns at the test sites.

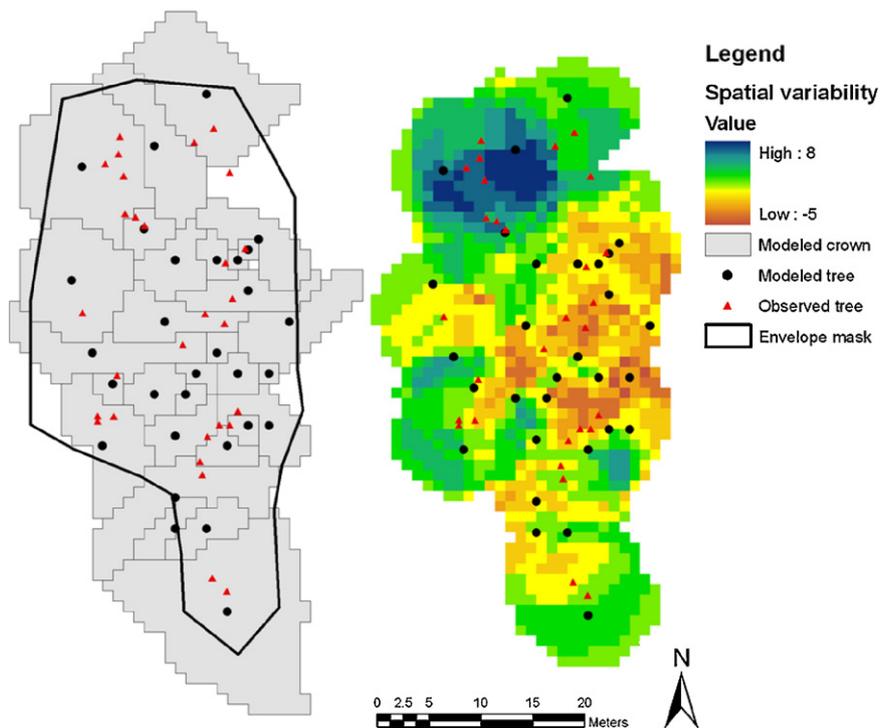
Vegetation patterns	$w_1$	$w_2$	$w_3$	Filtering window	Filter ( $par_1, par_2$ )	Segmentation ( $par_1, par_2, par_3, par_4, par_5, par_6$ )
High vegetation density (Serchio)	1.00	0.00	0.00	SW	LoF(4, 0.1)	WSG(4, 3, /, /, /, /)
High vegetation density with overlapping crowns (Sieve)	0.50	0.25	0.25	SW	AF(6, /)	WS(4, /, /, /, /, /)
Sparse vegetation (Sevilleta)	0.60	0.40	0.00	VW	GF(0, 1.07)	MCMWS(8, 1, 1, /, /, /)



**Fig. 5.** Spatial variability of plant density (right) and modeled crown boundaries (gray polygons), observed and modeled tree positions (red and black triangles) and envelope mask (left). Note the different measurement scale of the spatial variability. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

into a crow and the stem position cannot be located in the center of the canopy, in particular for areas with high density. Despite the complex scenario, the proposed model is able to explain the spatial variability of the plant density ( $RMSE = 2.35 \text{ \#/m}^2$ ). Fig. 6 shows the good agreement between observed and simulated density maps with low values on most of the area. However, the model underestimates the plant density in a zone characterized by closely spaced trees.

The best model performances are achieved over the Sevilleta area (Fig. 7), characterized by a sparse vegetation canopy with intercanopy spaces larger than the LiDAR resolution. The model can detect individual tree and crown boundaries and capture the spatial variability of the plant density ( $RMSE = 1.87 \text{ \#/m}^2$ ). Despite the encouraging overall results, there are two different zones, on the west and in the center of the test site, where the model underestimates slightly the real vegetation pattern.



**Fig. 6.** Same as Fig. 5, but for the Sieve study site.

**Table 5**  
Vegetation characteristics for the three areas.

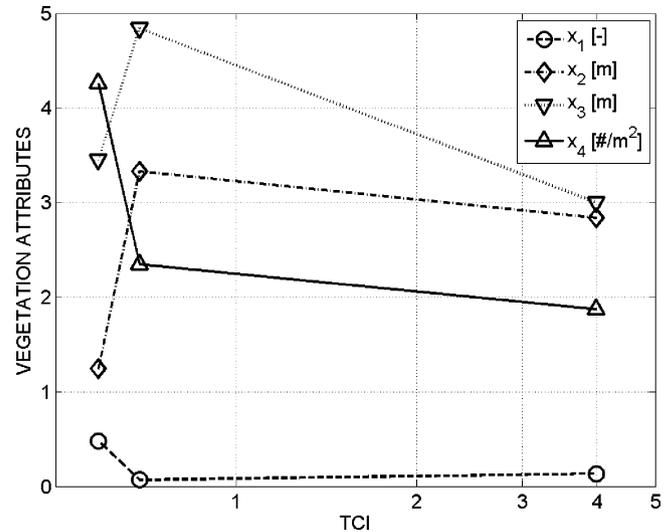
	Number of trees, $n_{obs}$	Test site area, $A_{TOT}$ (m <sup>2</sup> )	Individual area, $A_{TOT}/n_{obs}$ (m <sup>2</sup> )	Average spacing, $\bar{S}_V$ (m)	Average crown diameter, $\bar{D}_C$ (m)	Tree canopy index, $\bar{S}_V/\bar{D}_C$
Serchio	85	225.86	2.64	1.83	3.07	0.59
Sieve	29	1016	35.04	6.68	9.63	0.69
Sevilleta	97	10937	112.75	11.98	2	3.99

### 3.3. Linkage between vegetation characteristics and model performances

The three investigated areas present significant variability in vegetation density. Table 5 lists for each monitored area: number of trees ( $n_{obs}$ ), test area ( $A_{TOT}$ ), individual area ( $A_{TOT}/n_{obs}$ ), average vegetation spacing ( $\bar{S}_V$ ) and average crown diameter ( $\bar{D}_C$ ). It is evident an increasing average vegetation spacing (decreasing vegetation density) from A to C sites:  $\bar{S}_V(A) = 1.83$  m,  $\bar{S}_V(B) = 6.68$  m and  $\bar{S}_V(C) = 11.98$  m.

Since there is evidence of a dependence of model performance on vegetation characteristics, we explore this linkage by introducing an index called *tree canopy index* (TCI). The TCI is the ratio between the average plant spacing and the average crown diameter, both estimated by field measurements (Table 5). TCI is related to the vegetation characteristics and expresses the degree of crown overlap. Low values represent forest scenarios with overlapping crowns (Serchio and Sieve areas), whereas higher values indicate vegetation patterns with sparse plants and crown boundaries that are easily identified (Sevilleta area).

Fig. 8 shows the linkage between the TCI and the model performances in the form of the explored parameters ( $x_1, x_2, x_3, x_4$ ). Lower values of  $x_1, x_2, x_3$  and  $x_4$  indicate better model accuracy. The performances improve with increasing TCI. It is important to note that unidentified trees in each site do not contribute to estimates of  $x_2$  and  $x_3$ . As a result, at the Serchio site where nearly half of the trees are not detected, the values  $x_2$  and  $x_3$  are not representative. A sharp improvement in model performance between the Serchio and Sieve areas is evident despite similar vegetation characteristics as expressed by TCI. In the Serchio area, the near-meter plant



**Fig. 8.** Linkage between vegetation characteristics (tree canopy index) and model performances in the terms of attribute values ( $x_1, x_2, x_3$  and  $x_4$ ).

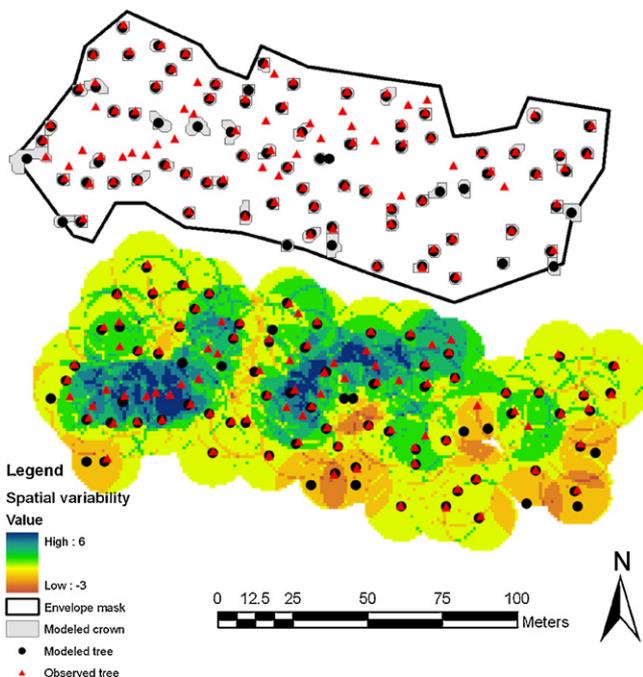
spacing does not allow the reproduction of a correct forest scenario. In the Sieve area, despite the overlapping crowns, the average plant spacing is large enough to allow an accurate vegetation pattern simulation. The Sieve and Sevilleta sites show comparable results with respect to  $x_1$ , highlighting the model capability to detect trees in broadleaved forests and in sparse conifer forests.

A preliminary cost–benefit analysis highlighted that the estimation of vegetation parameters based on the proposed LiDAR-derived methodology appears to be two orders of magnitude cheaper than the correspond field surveys (Forzieri et al., 2008).

### 4. Conclusions

This study describes a new time and cost-efficient approach for detecting individual tree positions, crown boundaries and plant density using airborne LiDAR data, allowing a significant reduction of the field surveys, limiting them over only calibration sites. We demonstrate, also, that varying image-processing techniques applied on the same canopy height model achieve different performances. As a result, a systematic approach, such as the proposed MADM-SAW method, is a powerful tool to detect the best combination of morphological analyses for reproducing the monitored forest. Since the model was tested over three vegetation patterns, the MADM architecture adaptively selected the most accurate de-noising and segmentation processes over each forest. By assigning site-specific weights to the attributes, improvements were obtained in the SAW approach for reproducing monitored forest scenarios. We identify the specific de-noising and segmentation parameters and useful weights to reproduce vegetation patterns similar to the tested areas.

The results show that the model performances are strongly related to vegetation characteristics. Good results are achieved in forest areas characterized by a high ratio between the average plant spacing and the average crown diameter (TCI), and plant



**Fig. 7.** Same as Fig. 5, but for the Sevilleta study site.

spacing larger than the remote sensing data spatial resolution. The model reproduces the investigated vegetation parameters well in complex broadleaves forests with high vegetation density and overlapped crowns, as well as in sparse conifer forest canopies. Nevertheless, limitations exist for vegetation monitoring using LiDAR data over forested areas with plant spacing comparable to the spatial resolution of the remote sensing data. Additional testing of the proposed approach over a broader range of tree canopy indices (TCI) would yield insight to specific thresholds for its applicability.

The main limit of the present study is the lack of another independent dataset to test the method. More extended field surveys would permit further verification of the MADM-SAW approach. In the decision making process on forest scenarios we use a three-parameter decision rule depending on the number of trees, tree positions and crown diameters. The use of a more extended set of vegetation attributes, including for example crown base height and stem diameter in addition to the afore-mentioned attributes, could provide further improvements giving an enhanced representation of the accuracies of the segmentation results. Besides, a systematic approach to identify the optimal weight setting for the MADM-SAW application could represent a further improvement of the proposed methodology, actually only partly objective. Nevertheless, we believe these preliminary results are already a useful contribution to the delineation of operational procedures for remotely characterizing vegetation patterns for forest management purposes. The proposed method is shown to be appropriate for selecting the best image-processing techniques for tree identification from LiDAR dataset by quantifying vegetation parameters such as tree position, crown diameter, and vegetation density.

## Acknowledgments

We thank H2CU for the financial support of the lead author during a visit to the New Mexico Institute of Mining and Technology. The LiDAR data for the Sevilleta site was provided through the National Center for Airborne Laser Mapping (NCALM), whose support is greatly appreciated. We thank Hugo Gutierrez-Jurado (Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology) and Andrea Dani (Dipartimento di Ingegneria Agraria e Forestale, University of Florence) for the collaborative assistance. The manuscript was also improved by the constructive comments of two anonymous referees.

## References

Andersen, H.E., McGaughey, R.J., Reutebuch, S.E., 2005. Estimating forest canopy fuel parameters using LiDAR data. *Remote Sensing of Environment* 94, 441–449.

Avery, T.E., Burkhardt, H.E., 1994. *Forest Measurements*. McGraw-Hill, Boston.

Baker, F.S., 1950. *Principles of Silviculture*. McGraw-Hill, New York.

Baptist, M.J., Babovich, V., Rodríguez Uthurburu, J., Keijzer, M., Uittenbogaard, R.E., Mynett, A., 2007. On inducing equations for vegetation resistance. *Journal of Hydraulic Research* 45 (4), 435–450.

Brandtberg, T., Warner, T., Landenberger, R., McGraw, J., 2003. Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density LiDAR data from the eastern deciduous forest in North America. *Remote Sensing of Environment* 85, 290–303.

Chankong, V., Haimes, Y.Y., 1983. *Multiobjective Decision Making: Theory and Methodology*. Elsevier, New York.

Chakraborty, S., Yeh, C., 2007. A simulation based comparative study of normalization procedures in multiattribute decision making. In: *Proceedings of the 6th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases*, Corfu Island, Greece, February 16–19, 2007.

Chen, Q., Baldocchi, D., Gong, P., Kelly, M., 2006. Isolating individual trees in a Savanna woodland using small footprint LiDAR data. *Photogrammetric Engineering & Remote Sensing* 72 (8), 923–932.

Deng, H., Zhang, W., Mortensen, E., Dietterich, T., 2007. Principal curvature-based region detector for object recognition. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR-2007)*, Minneapolis.

Dralle, K., Rudemo, M., 1996. Stem number estimation by kernel smoothing in aerial photos. *Canadian Journal of Forest Research* 26, 1228–1236.

Eagleson, P.S., 2002. *Ecohydrology, Darwinian Expression of Vegetation Form and Function*. Cambridge Press.

Finney, M.A., 1998. FARSITE: Fire Area Simulator—model development and evaluation. Research Paper RMRSRP-4. Ogden, UT. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, 47 pp.

Forzieri, G., Guarnieri, L., Castelli, F., Preti, F., 2008. Corsi d'acqua e vegetazione: tecniche di parametrizzazione mediante analisi di dati LiDAR. In: *Proceedings of the 31st Convegno Nazionale di Idraulica e Costruzioni Idrauliche*. Perugia, Italy, 9–12 September 2008 in Italian.

Gong, P., Sheng, Y., Biging, G.S., 2002. 3D model-based tree measurement from high-resolution aerial imagery. *Photogrammetric Engineering and Remote Sensing* 68, 1203–1212.

Gonzalez, R.C., Woods, R.E., Eddins, S.L., 2004. *Digital Image Processing Using MATLAB*. Pearson Prentice Hall.

Gougeon, F., Leckie, D., 2003. Forest information extraction from high spatial resolution images using an individual tree crown approach. Information report BC-X-396, Natural Resources Canada, Canadian Forest Service, 26 pp.

Gutiérrez-Jurado, H.A., Vivoni, E.R., Istanbuluoglu, E., Bras, R.L., 2007. Ecohydrological response to a geomorphologically significant flood event in a semiarid catchment with contrasting ecosystems. *Geophysical Research Letters* 34, L24S25, doi:10.1029/2007GL030994.

Gutiérrez-Jurado, H.A., Vivoni, E.R., Harrison, J.B.J., Guan, H., 2006. Ecohydrology of root zone water fluxes and soil development in complex semiarid rangelands. *Hydrological Processes* 20, 3289–3316.

Hinsley, S.A., Hill, R.A., Gaveau, D.L.A., Bellamy, P.E., 2002. Quantifying woodland structure and habitat quality for birds using airborne laser scanning. *Functional Ecology* 16, 851–857.

Hyyppä, J., Kelle, O., Lehtikoinen, M., Inkinen, M., 2001. A segmentation based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing* 39, 969–975.

Hwang, C.-L., Yoon, K., 1981. *Multiple Attribute Decision Making: Methods and Applications*. Springer, Berlin.

Järvelä, J., 2004. Determination of flow resistance caused by nonsubmerged woody vegetation. *International Journal of River Basin Management* 2 (1), 61–70.

Kouwen, N., Fathi-Moghadam, M., 2000. Friction factors for coniferous trees along rivers. *Journal of Hydraulic Engineering* 126 (10), 732–740.

Kvak, D., Lee, W., Lee, J., Biging, G.S., Gong, P., 2007. Detection of individual trees and estimation of tree height using LiDAR data. *Journal of Forest Research* 12, 425–434.

Maltamo, M., Mustonen, K., Hyyppä, J., Pitkänen, J., Yu, X., 2004. The accuracy of estimating individual tree variables with airborne laser scanning in a boreal nature reserve. *Canadian Journal of Forest Research* 34, 1791–1801.

Malczewski, J., 1999. *GIS and Multicriteria Decision Analysis*. John Wiley & Sons Inc., New York.

Næsset, E., Økland, T., 2002. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment* 79, 105–115.

Osma-Ruiz, V., Godino-Llorente, J.I., Sáenz-Lechón, N., Gómez-Vilda, P., 2007. An improved watershed algorithm based on efficient computation of shortest paths. *Pattern Recognition* 40, 1078–1090.

Persson, Å., Holmgren, J., Söderman, U., 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing* 68, 925–932.

Petryk, S., Bosmajian, G.B., 1975. Analysis of flow through vegetation. *Journal of Hydraulic Division, ASCE* 101 (7), 871–884.

Piussi, P., 1994. *Selvicoltura generale*. UTET.

Popescu, S.C., Wynne, R.H., 2004. Seeing the trees in the forest: using LiDAR and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogrammetric Engineering & Remote Sensing* 70 (5), 589–604.

Popescu, S., Wynne, R., Nelson, R., 2002. Estimating plot-level tree heights with LiDAR: local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture* 37, 71–95.

Popescu, S., Zhao, K., 2008. A voxel-based LiDAR method for estimating crown base height for deciduous and pine trees. *Remote Sensing of Environment* 112 (3), 767–781.

Rambabu, C., Chakraborty, I., 2007. An efficient immersion-based watershed transform method and its prototype architecture. *Journal of Systems Architecture* 53, 210–226.

Reutebuch, S.E., McGaughey, R.J., Andersen, H.E., Carson, W.W., 2003. Accuracy of a high-resolution LiDAR terrain model under a conifer forest canopy. *Canadian Journal of Remote Sensing* 29 (5), 527–535.

Shivers, B.D., Borders, B.E., 1996. *Sampling Techniques for Forest Resource Inventory*. Wiley, New York.

Soille, P., 1999. *Morphological Image Analysis*. Springer-Verlag, New York, NY.

Thompson, G.T., Roberson, J.A., 1976. A theory of flow resistance for vegetated channels. *Transactions, ASAE* 19 (2), 288–293.

Vincent, L., Soille, P., 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (6).

Voogd, H., 1983. *Multi-criteria Evaluations for Urban and Regional Planning*. London Princeton University.

Zhao, K., Popescu, S., Nelson, R., 2009. LiDAR remote sensing of forest biomass: a scale-invariant approach using airborne lasers. *Remote Sensing of Environment* 112, 182–196.