Fuzzy Analysis of Airborne LiDAR Data for Rainforest Boundary Determination

Zhenyu Zhang
School of Civil Engineering and Surveying
University of Southern Queensland
Toowoomba, Queensland 4350
Australia
zhenyu.zhang@usq.edu.au

Xiaoye Liu
School of Civil Engineering and Surveying
University of Southern Queensland
Toowoomba, Queensland 4350
Australia
xiaoye.liu@usq.edu.au

Kevin McDougall
School of Civil Engineering and Surveying
University of Southern Queensland
Toowoomba, Queensland 4350
Australia
kevin.mcdougall@usq.edu.au

Wendy Wright
Faculty of Science and Technology
Federation University
Churchill, Victoria 3842
Australia
wendy.wright@federation.edu.au

ABSTRACT

Airborne LiDAR data have advantages over passive remote sensing data in detailed description of vertical forest structure. LiDAR-derived information can potentially be used to solve such problems as forest type classification and forest boundary determination. Forest boundaries were usually represented as sharp lines that attempt to distinguish areas with different forest types. In reality, however, forest boundaries are rarely sharp or crisp, especially in the forest area with multiple canopy layers where species compositions change gradually. Fuzzy analysis offers great potential for characterising the transition zones and determining realistic forest boundaries. This study developed ways of using fuzzy analysis of airborne LiDAR data for determining rainforest boundaries. LiDAR variables were derived and used to define and calculate membership function values for both rainforest and non-rainforest. The confusion index values were then derived to illustrate the transition zones. Finally, the rainforest boundaries were successfully determined in the study area. The results demonstrated the success of proposed method for rainforest boundary determination.

KEYWORDS

LiDAR, Laser scanning, Fuzzy logic, Rainforest, Forest classification

1 INTRODUCTION

Fuzzy set theory developed by Zadeh [1] has been widely adapted to deal with natural phenomena with gradual changes. Fuzziness is often a concomitant of complexity, deal with ambiguity, vagueness, and ambivalence. Fuzz analysis can be used to characterise classes that cannot have or do not have sharply defined boundaries [2]. For example, the forest boundaries were usually represented as sharp lines that attempt to distinguish areas with different forest types. In reality, however, forest boundaries are rarely sharp or crisp, especially in the forest area with multiple canopy layers where species compositions change gradually. Therefore, the forest boundaries should be more realistically described as transition zones rather than boundaries with abrupt changes [3]. Fuzzy analysis offers a great potential for characterising the transition zones and determining realistic forest boundaries within the transition zones.

Application of fuzzy analysis into the forest can include information about the nature of the forest boundaries [2]. The nature information of the boundary can be used to describe the spatial changes of some important natural phenomena. For example, ecotone which describes a transitional area between different vegetation types or ecological communities [4, 5] is of significant importance in biogeographical investigations, especially in monitoring environmental change [6].

Forest type classification is a prerequisite for sustainable forest management and native forest conservation. Remotely sensed data have been widely explored for forest applications. However, passive remote sensing techniques are limited in their ability to capture forest structure complexity. It has been shown that active remote sensing techniques via airborne LiDAR (Light Detection and Ranging) with capability of canopy penetration yields such high-density sampling that detailed description of the forest structure in three dimensions can be obtained. Accordingly, much interest is attached to exploring the application of this approach for identifying the distribution of designated vegetation communities [7].

Rainforests in Victoria, Australia are protected from the impacts of timber harvesting through the imposition of buffers appropriate to the maintenance of their key environmental parameters. Boundaries between rainforest and adjacent forests must be well positioned and monitored for effective rainforest conservation and protection. In our study area, the ecotone is the zone where the rainforest and sclerophyll forest overlap [8], forming a transition zone with mixed forest types. In this transition zone, Mountain Ash dominates the higher storey of the canopy while the rainforest dominates the lower storeys. The boundary between the rainforest and surrounding non-rainforest must be somewhere within this transition zone.

Traditional methods for rainforest boundary determination in an area similar to our study area were based on the interpretation of aerial photographs or field work. These methods are labour intensive and time consuming. In addition, aerial photographs (and even the satellite images) are only able to provide information on horizontal forest extent. The LiDAR data have advantages over these traditional remote sensing data in that vertical forest structure (or the interior of the canopy and understorey vegetation) can be delineated in much detail [7]. However, there is still considerable scope for developing advanced methodology to take maximum advantage of the information extracted from the LiDAR data for effective determination of rainforest boundaries. Therefore, the overall objective of this study is to develop a fuzzy logic model using airborne LiDAR data for determining the boundary between cool temperate rainforest and adjacent forests in the study area. Fuzzy analysis of LiDAR-derived variables will be performed to determine fuzzy membership values describing degrees to which each area belongs to a certain forest type in the study area, and the confusion index values will also be calculated to illustrate the transition zones (or ecotone areas) and determine the boundaries between the rainforest and non-rainforest.

To our best knowledge, this is the first time applying fuzzy analysis to the airborne LiDAR data in Australian cool temperate rainforest environment. This study contributes to the knowledge and understanding of the development of fuzzy logic model using the airborne LiDAR data for effective rainforest boundary determination. The results from this study will strongly warrant the operational adoption of the fuzzy analysis of the airborne LiDAR data in the management of forestry resources. The remainder of this paper is organised as follows. Next section describes the study area and data used. Section 3 presents the fuzzy logical model development in detail. It is followed by two sections presenting the results and discussion. The final section draws the conclusions.

2 MATERIALS

2.1 Study Area

The study area is in the eastern Strzelecki Ranges, southeast Victoria, Australia. The Strzelecki Ranges are an isolated series of mountains in the southern section of the Gippsland region that are surrounded by the Gippsland Plain. Before European settlement, the Strzelecki Ranges were densely vegetated by wet forest (or wet sclerophyll forest) and cool temperate rainforest. Wet forest is most commonly dominated by Mountain Ash (*Eucalyptus regnans*) [9], characterised by a tall eucalypt overstorey, a broad-leaved shrubby understorey and a moist, shaded, fern-rich ground layer that is usually dominated by tree-ferns [10]. In eucalypt-free areas, Silver Wattle (*Acacia dealbata*) may be locally dominant [9]. Cool temperate rainforest is defined as a closed, non-eucalypt forest, which occurs in high rainfall areas and within wet forest areas which have not been exposed to fire [11]. Myrtle Beech (*Nothofagus cunninghamii*) is the dominant species of cool temperate rainforest in the study area. The understorey is characterised by tree ferns and a rich epiphytic flora. The ground layer is dominated by a diversity of ground ferns such as Mother Shield-fern, Hard Water-fern, and Leathery Shield-fern [10].

These forests have experienced widespread land clearing since European settlement. Subsequent agricultural abandonment and a significant fire history have resulted in severe landscape disturbance in the Strzelecki Ranges. There was extensive regeneration of eucalypt forest following catastrophic wildfires in 1939 and 1944. Currently, areas bordering patches of cool temperate rainforest in the eastern Strzeleckis are a mosaic of different land use histories involving both natural and human disturbances, and so a very complex forest structure in the remnant patches of cool temperate rainforest and adjacent forests including wet sclerophyll and plantation forests prevails [7]. This study focuses on an area with cool temperate rainforest distribution in the Eastern Strzeleckis, which covers an area of 25 hectares.

2.2 Data

Airborne LiDAR data were collected using an Optech ALTM Gemini LiDAR system at a flying height of 1,100 m above ground between 11 and 23 October 2009 (for the whole Strzelecki Ranges). The laser pulse repetition frequency is 70 kHz. The wavelength of the LiDAR laser is 1.064 μm. The laser scanner was configured to record up to 4 returns for one laser pulse. The average point density was 4 points per square metre, and the laser footprint diameter was 0.3 m. The LiDAR data used for this project was documented as 0.20 m for vertical accuracy and 0.25 m for horizontal accuracy. The LiDAR data were classified into ground and non-ground points by the vendor and were delivered in binary LAS 1.2 file format. The EVCs (Ecological Vegetation Classes) data in the region provided by the HVP Plantations Pty Ltd were used as reference data in this study.

3 MODEL DEVELOPMENT

3.1 Normalising LiDAR Data

A digital elevation model (DEM) with two-metre horizontal resolution (grid size) was generated using the LiDAR ground data. The height of laser returns above the ground is calculated from the difference between the laser returns (including those from tree canopy and understorey vegetation) and the corresponding DEM value. It is these normalised laser returns that provide an effective way to depict the vertical structure of the forests throughout the whole canopy structure. The structural differences will affect the distribution of the laser returns from the forests [12, 13]. Therefore, the variables or metrics derived from the LiDAR data can be used for tree species identification and forest type classification.

3.2 Deriving LiDAR Variables

A grid of square columns with 4×4 m horizontal resolution (grid size) covering the study area were generated to quantitatively describe the height distribution of LiDAR data. The none-ground LiDAR points within each of these square columns were extracted to quantify the vertical distribution of LiDAR points by calculating height variables as listed in Table 1.

Table 1: LiDAR-derived variables in each of the columns.

Variables	Description
Max	Maximum height in the column
Mean	Mean height in the column
StdDev	Standard deviation of height
Variance	Variance of height
Skewness	Skewness of height
AAD	Average absolute deviation from mean
P30	30th percentile of height
P60	60th percentile of height
P90	90th percentile of height
L2	The second L-moment

3.3 Defining Fuzzy Membership Functions

Let X be a set, with a generic element of X denoted by x: $X=\{x\}$. A fuzzy set A in X is characterised by a membership function $MF(x)$ which associates with each element $x \in X$ a real number in [0, 1]. The value of $MF(x)$ at x represents the grade of membership of x in A. The closer the value of $MF(x)$ to 1, the higher the degree to which the x belongs to A [1]. In other words, the $MF(x)$ of x in A specifies the extent to which x can be regarded as belonging to A [2]. Using the fuzzy set, each forest area can be assigned a membership grade for every forest property value to indicate the degree to which it is associated with the different forest types [3].

After the LiDAR variables have been derived from normalised LiDAR data in each square column, 4 typical rainforest areas and 4 typical non-rainforest areas in the study area were selected, called rainforest training area and non-rainforest training area. The mean μ and standard deviation σ of each variable were then calculated from rainforest training area and non-rainforest training area respectively. Both rainforest and non-rainforest training areas have 10 pairs of mean and standard deviation values which will be used to form the membership functions for fuzzy analysis.

In order to define the fuzzy membership functions and calculate the membership function values for each of the variables, the mean value of a variable was selected as the centre of the fuzzy set. The membership value is 1 at the centre of the set and it falls off in a way through the fuzzy boundaries to the region outside the set, where the membership value takes the value 0. Here, the width of the fuzzy set was defined by two times of the standard deviation value of the variable, so-called two-sigma rule. The values of the variable less than two times of the standard deviation value away from the mean accounting for 95% of the values were used for fuzzy analysis here. Therefore, the fuzzy membership function for the rainforest was defined by:

$$MF(x) = 1 \quad \text{if } x < \mu \tag{1.a}$$

$$MF(x) = \frac{(\mu+2\sigma)-x}{2\sigma} \quad \text{if } \mu \leq x \leq (\mu+2\sigma) \tag{1.b}$$

$$MF(x) = 0 \quad \text{if } x > (\mu+2\sigma) \tag{1.c}$$

The fuzzy membership function for the non-rainforest was defined by:

$$MF(x) = 0 \quad \text{if } x < \mu - 2\sigma \tag{2.a}$$

$$MF(x) = \frac{x-(\mu-2\sigma)}{2\sigma} \quad \text{if } \mu - 2\sigma \leq x \leq \mu \tag{2.b}$$

$$MF(x) = 1 \quad \text{if } x > \mu \tag{2.c}$$

For example, if the mean and standard deviation for the variable of maximum height in the rainforest training area were calculated as being 18.94 and 7.01, the membership value is 1 if the variable values are less than μ (= 18.94 here). The membership values are between 0 and 1 in the region defined by μ and (μ+2σ), here between 18.94 and 32.96. The membership values take 0 if the values of the variable are greater than (μ+2σ). The first set of rainforest membership function values for individual square columns over the study area were calculated using the membership function formulae 1.a to 1.c and the mean and standard deviation values of this variable. Similarly, other 9 set of rainforest membership function values can be calculated using other 9 variables. In same way, 10 set of non-rainforest membership function values were also calculated using the formulae 2.a to 2.c.

3.4 Computing Integrated Membership Function Values

Once the membership values were calculated for the whole study area, there were 20 membership function values in each square column, with 10 rainforest membership function values representing the degree to which this column belongs to rainforest and the other 10 non-rainforest membership function values representing the degree to which this column belongs to non-rainforest. 20 raster images (10 for rainforest and 10 for non-rainforest) with 2×2 m horizontal resolution (grid or cell size) were then created using these membership values, called rainforest membership images and non-rainforest membership images respectively. A raster image of combined or integrated rainforest membership function value (MFV_{RF}) was generated by averaging the 10 rainforest membership images. Similarly, a raster image of combined non-rainforest membership function value (MFV_{NRF}) was also created by averaging the 10 non-rainforest membership images.

3.5 Calculating Confusion Indices

A confusion index (*CI*) which is a measure of the *confusion* of two membership function values in a specific location was computed using the MFV_{RF} and the MFV_{NRF} for each cell as below [14]:

$$CI = 1 - |MFV_{RF} - MFV_{NRF}| \tag{3}$$

Calculations using Formula 3 which is a modified formula from Burrough [14] resulted in a raster image showing the confusion index values in individual cells. If the two membership function values in a cell are significantly different, the confusion index will be very close to zero, indicating less confusion, i.e., either the rainforest or the non-rainforest is the dominant in the cell. On the other hand, however, the closer the two membership function values in a cell, the bigger the confusion about to which forest type the cell belongs. These confusion cells indicate the mixed areas or transition zones between the rainforest and non-rainforest in the study area.

3.6 Determining Rainforest Boundary

The confusion index image created from the above step illustrated the similarity and/or difference between the rainforest membership function values and the non-rainforest membership function values. The boundary between the rainforest and the non-rainforest must occur in the transition zone where the maximum confusion index values were presented. Next we used low pass filtering and reclassification to the confusion index image to extract the boundaries. Finally, the raster image was converted to vector to determine the boundary polygon.

4 RESULTS

The raster image of the integrated rainforest membership function values was shown in Fig. 1. The cells with darker colour got bigger membership function values for the rainforest, indicating to a greater extent the cells belong to the rainforest. On the other hand, the cells with brighter colour have smaller membership function values, implying a lesser extent to which the cells belong to the rainforest.

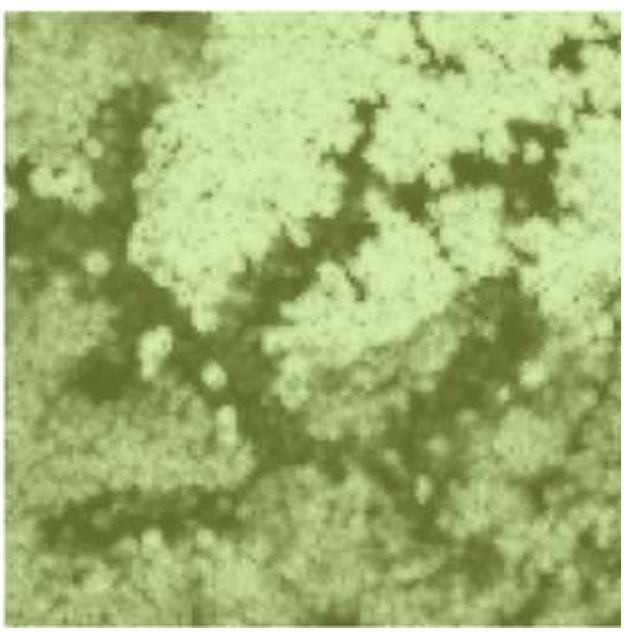

Figure 1: Integrated rainforest membership function values.

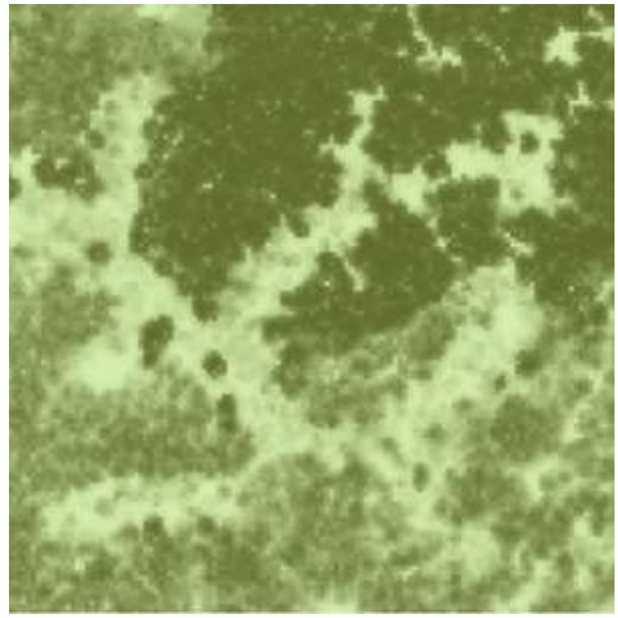

Figure 2: Integrated non-rainforest membership function values.

The Image in Fig. 2 shows the integrated membership function values for the non-rainforest. The cells with darker colour exhibit a greater extent to which the cells belong to the non-rainforest while the cells with brighter colour present a lesser extent to which the cells belong to the non-rainforest. Both Figure 1 and Fig. 2 show consistent presence of the rainforest and the non-rainforest in the study area.

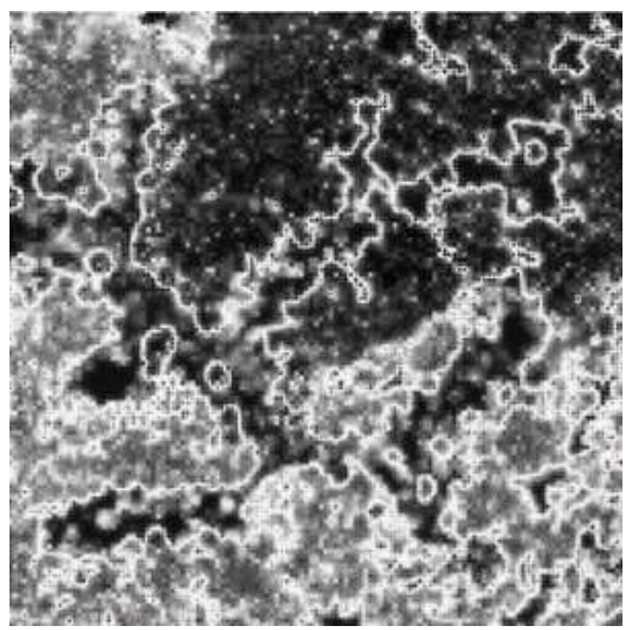

Figure 3: Confusion index values.

Fig. 3 shows the confusion index values calculated from the rainforest membership function values and the non-rainforest membership function values. Some cells exhibit very bright colour where the two membership function values are very similar. Therefore, there exist bigger confusion index values, representing the transition zone between the rainforest and the non-rainforest. The cells with the brightest tone indicate the presence of the possible boundaries.

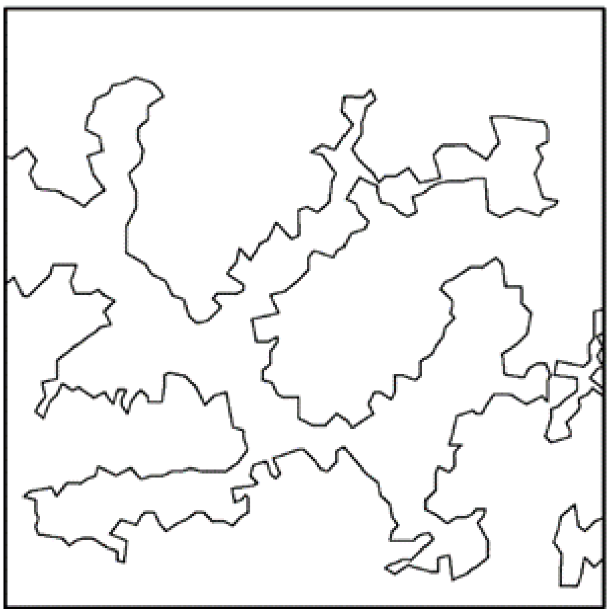

Figure 4: Boundaries between rainforest and non-rainforest.

The boundaries between the rainforest and the non-rainforest extracted from the confusion index image are shown in Fig. 4. It was compared with EVCs, which describe the spatial extent of vegetation species in the study area and were used as ground reference data in this study. Results show the boundaries determined using the way developed from this study have an overall 85% consistent with the boundary described from the EVCs.

5 DISCUSSION

A prerequisite for rainforest conservation and protection is the determination and monitoring of the boundaries between rainforest and adjacent forests. There has been increasing interest in the application of airborne LiDAR for the analysis of forest structures and related applications over the last decade. However, there have been very few studies attempting to determine the rainforest boundary using the airborne LiDAR data. Based on previous research [15], column-based LiDAR variables were derived for vertical description of the forest structure. Unlike the interpretation of airborne or satellite images which see forests only from above the top of canopy, the airborne LiDAR data provide detailed information not only from the top of the canopy, but also through the interior of the canopy and the ground because of the LiDAR penetration of forest canopies [7]. However, an effective way was required to make full use of the valuable LiDAR-derived information for effective determination of the forest boundaries.

This study demonstrated the success of using fuzzy analysis of the LiDAR data for rainforest boundary determination. In a mixed forest area, the forest types change gradually. Therefore, it is difficult to determine a crisp forest boundary. It is demonstrated here that fuzzy analysis of LiDAR-derived variables is an effective way to determine the transition zone and the boundaries between the rainforest and the non-rainforest. Furthermore, the determination of the transition zone (or ecotone) is of significant importance in ecological communities.

6 CONCLUSIONS

It has been shown that the airborne LiDAR data have advantages over passive remote sensing data in detailed description of vertical forest structure, but there is still considerable scope for developing advanced models to take maximum advantage of LiDAR-derived information for forest applications. This study developed ways of using fuzzy analysis of the airborne LiDAR data for boundary determination between the rainforest and adjacent non-rainforest. Column-based LiDAR variables were derived and used to define and calculate membership function values for both rainforest and non-rainforest. These membership values were subsequently used to create integrated membership function values for rainforest and non-rainforest respectively. The raster images from both integrated membership function values showed consistent presence of the rainforest and the non-rainforest in the study area. It is evident that derived confusion index values augmented the confusion of two integrated membership function values if two membership values are close in a location. The confusion index illustrated transition zones, indicating the presence of the boundary. This study demonstrated the success of the fuzzy analysis of the airborne LiDAR data for rainforest boundary determination. The results from this study will strongly warrant the operational adoption of the fuzzy analysis of the airborne LiDAR data in the management of forestry resources.

ACKNOWLEDGMENTS

We would like to thank HVP Plantations Pty Ltd for providing the LiDAR and EVC data for use in this study.

REFERENCES

[1] L. A. Zadeh. 1965. Fuzzy Sets. *Information and control* 8, 338-353.
[2] P. A. Burrough, R. A. McDonnell, and C. D. Lloyd. 2015. *Principles of geographical information systems*, 3rd edition., Oxford, UK: Oxford University Press.
[3] F. Wang. 1996. Fuzzy representation of geographical boundaries in GIS. *International Journal of Geographical Information Science* 10, 5, 573-590.
[4] J. J. Camarero, E. Gutiérrez, and M. J. Fortin. 2006. Spatial patterns of plant rechness across treeline ecotones in the Pyrenees reveal different locations for richness and tree cover boundaries. *Global Ecology and Biogeography* 15, 182-191.
[5] K. Hufkens, P. Scheunders, and R. Ceulemans. 2009. Ecotones in vegetation ecology: methodologies and definitions revisited. *Ecological Research* 24, 5, 977-986.
[6] G. M. Foody, and D. S. Boyd. 1999. Fuzzy mapping of tropical land cover along an environmental gradient from remotely sensed data

with an artificial neural network. *Journal of Geographical Systems* 1, 23-35.
[7] Z. Zhang, X. Liu, J. Peterson, and W. Wright. 2011. Cool temperate rainforest and adjacent forests classification using airborne LiDAR data. *Area* 43, 4, 438-448.
[8] D. Cameron. 2008. *A field guide to rainforest identification in Victoria: differential species keys for the delineation of rainforest boundaries*, Melbourne, Australia: Victorian Department of Sustainability and Environment.
[9] J. B. Davies, A. M. Oates, and A. V. Trumbull-Ward. 2002. *Ecological vegetation class mapping at 1:25000 in Gippsland*. Victorian Department of Natural Resources and Environment, Melbourne, Australia.
[10] DSE. 2005. *Native Vegetation Manage Policy*. Department of Sustainability and Environment, Melbourne, Australia.
[11] P. Adam. 1992. *Australian Rainforests*, Oxford, UK: Oxford University Press.
[12] A. S. Antonarakis, K. S. Richards, and J. Brasington. 2008. Object-based land cover classification using airborne LiDAR. *Remote Sensing of Environment* 112, 2988-2998.
[13] H. O. Ørka, E. Næsset, and O. M. Bollandsås. 2009. Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. *Remote Sensing of Environment* 113, 1163-1174.
[14] P. A. Burrough. 1996. Natural objects with indeterminate boundaries. In *Geographic objects with indeterminate boundaries*, P. A. Burrough and A. U. Frank, (Eds.). London: Taylor & Francis. 3-28.
[15] Z. Zhang, and X. Liu. 2013. WorldView-2 satellite imagery and airborne LiDAR data for object-based forest species classification in a cool temperate rainforest environment. In *Developments in multidimensional spatial data models*, A. A. Rahman, P. Boguslawski, C. Gold and M. N. Said, (Eds.). Berlin: Springer. 103-122.