AN OBIA FOR FINE-SCALE LAND COVER SPATIAL ANALYSIS OVER BROAD TERRITORIES: DEMONSTRATION THROUGH RIPARIAN CORRIDOR AND ARTIFICIAL SPRAWL STUDIES IN FRANCE

T. Tormos* a, S. Dupuy b, K. Van Looy c, E. Barbe c, P. Kosuth c

a IRSTEA, UR MALY, Pole ONEMA/IRSTEA freshwater hydroecology, Lyon, F-69336, France-thierry.tormos@onema.fr
b CIRAD, UMR TETIS, Montpellier, F-34398 France-stephane.dupuy@cirad.fr
c IRSTEA, UMR TETIS, Montpellier, F-34398 France-pascal.kosuth@teledection.fr

KEY WORDS: Top-down classification tree construction, Rules-based classification, Large mapping, Classification automation, Riparian Corridor, Artificial sprawl, Broad scale spatial analysis

ABSTRACT:
Spatial analysis using fine-scale information over broad territories is essential to define efficient restoration strategies from local to national scale. We designed an OBIA dedicated to produce operationally reliable fine-scale information over broad territories. The originality of our OBIA lies particularly in the top-down approach for the construction of the classification tree and the use of ‘knowledge-based rules’ classification technique. The implementation of this OBIA over the two study areas – (i) the Normandy region for riparian area land cover mapping (5600 km² riparian area) and (ii) fours departments over the Languedoc-Roussillon region (22644 km²) – demonstrates the operability of our approach (time-efficient, reproducible, transferable, portable). Broad scale spatial analysis conducted from resulting maps demonstrate the interest of using fine-scale information and highlight that OBIA, following our approach, will be at very short run a broadly applicable method to carry out such analysis.

1. INTRODUCTION

1.1 VHRS and OBIA

Spatial analysis is a key feature to the understanding of structures, dynamics and changes in the environment, societies and territories. Furthermore it is essential for the exchange between stakeholders in society (citizens, government, socio-economic) to define coordinated management strategies and implement them. Therefore a huge step forward in the remote sensing for Earth Observation has been seen in the last two decades for producing spatial information required to carry out spatial analysis (Turner et al., 2001). Nowadays, with the progresses of satellite and airborne imagery in High and Very High Spatial Resolution (H&VHRS), fine-scale land cover spatial analysis over broad territories could be conducted. However, due to the heterogeneity and the volume of information within H&VHSR images, extracting it efficiently and with lower costs over broad territories is quite challenging. The last five years, OBIA have revolutionized the processing of H&VHRS remote sensing data by providing effective computer-assisted classification techniques whose results come close to the quality of manual photo-interpretation, while being much faster and cheaper and much more reproducible (e.g., Durieux et al., 2007, Tiede et al., 2010). As a result, we proposed in this paper an operational OBIA procedure designing for the extraction of spatial information from H&VHRS remotely-sensed data over broad territories. Two case studies (CS) supporting public politics in France are exhibited: riparian corridor and artificial sprawl management.

1.2 CS1: Riparian corridor management

Maintaining and restoring good riparian buffer area conditions could constitute a major action to improve the freshwaters’ ecological status required by the European Water Framework Directive (WFD) by 2015. Riparian vegetation, being an interface between terrestrial and aquatic systems, influences the biodiversity and water quality of stream and river ecosystems in many specific ways (Naiman et al., 2005). As a result, the first case study (CS1) is dedicated to the fine classification of Riparian Area Land Cover (RALC). RALC maps will serve as a basis for calculating spatial indicators related to human pressures along rivers and riparian vegetation characteristics (e.g. composition, continuity and strip width). These indicators will help in better understanding and predicting mechanisms influencing river ecological status at riparian scale, in order to prioritize and design efficient river corridor restoration strategies (Gergel et al., 2002, Tormos et al., 2011). The classification must, as a minimum requirement, extract: “water surfaces” and “semi-natural bare soils” required to better delimitated the river bed for building spatial indicators; “arable areas”, “urban areas” considered to be the main causes of stream ecological status alteration (Allan, 2004); “tree vegetation”, “semi-natural herbaceous and shrub vegetation” and “permanent agricultural grassland” constituting the main natural elements of the river corridor landscape that maintain biodiversity and regulate non-point source pollution (Naiman et al., 2005).

1.3 CS2: Artificial sprawl management

In France, in the peri-urban context, artificial sprawl (i.e., sprawl of urban and other artificial land development such as roads, quarries, landfills...) dynamics are particularly strong with huge population growth as well as a land crisis. The increase and spreading of built-up areas from the city centre towards the periphery takes place to the detriment of natural and agricultural spaces. The conversion of land with agricultural potential is all the more worrying as it is usually irreversible (Pointereau et al., 2009). As a result, the second case study (CS2) is dedicated to the fine extraction of artificial objects at a given time. The production of such map on different time steps will serve as a basis for calculating spatial indicators related to

* Corresponding author.
artificial sprawl. These indicators will help in better localizing and quantifying loss of land at both local and regional scales in order to prioritize and define efficient restoration strategies dedicated to the conversion of land with agricultural potential. Two classes are required: “artificial areas” (i.e., urban and fabric areas, transport infrastructures…) and “non artificial areas”.

2. DATA AND METHODS

2.1 Study areas, datasets and pre-processing

2.1.1 For the CS1
We classified RALC on a part of the Normandy river network (25 000 km² basin; 6000 km long river network, 5600 km² riparian area) located in the North of France (see Figure 1). This area is dominated by a strong agricultural occupation, primarily focused on field crops and livestock, with specific structures in the riparian buffer, such as tree vegetation and grassland near the river. This agricultural landscape is dotted with urban zones, from small to large. Datasets were chosen according to both their availability and their cost effectiveness on the overall French territories for managers. A comparative economic analysis of the different data led to the collection of two types of H&VHSR multispectral remotely sensed data: orthophotos and SPOT 5 satellite images. The characteristics of these image and pre-processing data are summarized in table 1. Orthophotos (0.5 m pixel) provide the textural information required to detect narrow and fragmented cover types along rivers (Müller, 1997). SPOT-5 XS (10 m pixel) is also acquired in order to get information in the NIR band essential to discriminate vegetation classes (Johansen et al., 2006). In addition existing spatial thematic data, relevant for RALC classification were collected. This data gives a precise (metric precision) spatial information on artificial continuous areas (city centre), roads, hydrographic surfaces (lakes and reservoirs) and the majority culture of farms declared in the frame of the European common agricultural policy.

The European land cover database, CORINE Land Cover (CLC), was collected also to perform the spatial analysis of this case study. CLC was built by visual interpretation of both Landsat and SPOT satellite images acquired from 1990 and 2000 years. Interpretation of the images is based on transparencies overlaid on 1/100.000 hard copy prints of satellites images (Bossard et al., 2000). It is based on a standard nomenclature organized into a 3-level hierarchy containing 44 classes. CLC features, characterized by a 25 ha minimum area, are either homogenous areas or combinations of land cover types with a certain recognizable structure.

2.1.2 For the CS2
We extracted artificial objects over the four coastal departments (Aude, Gard, Hérault and Pyrénées-Orientales) of the Languedoc-Roussillon region located in the South of France. The region covers 22644 km² and had a population of 2548000 in 2007. This is one of the most productive regions in wine in France. Over the last three decades, population pressure in Languedoc-Roussillon has led to rapid and poorly managed urbanization of the coastal plain which comprises notably the most productive land in the region (Abrantes et al., 2010). This phenomenon was increased by the successive crises that occurred in the wine making sector in recent years (Jarrige et al., 2009).

![Figure 1. Localization of study areas for the two case studies](Image)

RapidEye images, HSR multi-spectral satellite images, were collected in order to extract artificial object at summer 2009. Summer period is a suitable period for this work because vegetation is in active growth on this region that limits confusion between bare soils (natural or agricultural) and artificial areas. These images are produced by the remote sensors in a five-satellite constellation in use since 2008. This constellation can quickly provide homogenous and recent data covering large areas. Images are delivered in 25 km large orthorectified blocks with a spatial resolution of 5 m. Given the difficulties involved in extracting roads from remotely sensed images, we collected this information from available French topographic data.

An earlier map informing on artificial objects in 1997 was collected also to perform the spatial analysis of this case study. This map was produced by (Dupuy et al., 2012) from indian remote sensing images (5.8 m pixel for panchromatic image and 23 m for multi-spectral image).

2.2 OBIA approach

The originality of our OBIA approach lies in (i) the employment of the thematic spatial information into the classification, (ii) the classification tree; and (iii) the definition of the classification rules.

(i) Thematic spatial data contain fully reliable information to make maps. Therefore, this is the first information that is exploited by our OBIA procedure. A first level segmentation is created from this information. The study area is segmented according to the boundaries of thematic data entities. Then, the image objects that result are affected - or not - to a thematic class using Boolean rules. Within these boundaries, other(s) segmentation level(s), suitable for extracting objects of interest, are built. The classification tree is implemented on the resulting hierarchical image object network.
(i) The construction of the classification tree can be categorized as top-down (i.e., low to high level) image interpretation. The tree starts from the classes that are the easiest to extract, according to available data sources, to the classes of interest (e.g., “water surface” / “land surface”, and then within “land surface”, “soils with high vegetation” / “soils with low vegetation”)…).

(ii) For each decision in the classification, rules are developed using fuzzy or crisp membership function based on one or several relevant spectral, spatial and contextual features selected by either expert knowledge or trial and error runs or sole visual judgment. Tree classification is implemented over the specific hierarchical image object network.

A complete description of OBIA implemented over the two cases studies can be found in (Tormos et al., 2012) for CS1 and (Dupuy et al., 2012).

2.3 Classification automation

The study zone is first divided into homogeneous mapping regions according to the image acquisition date. Then, a master ruleset is designed according to the OBIA approach on a pilot zone. Finally, the master ruleset is implemented for each mapping region in adjusting classification rules and eventually added some classes to the classification tree for taking into account regions specificities.

Master ruleset and classification rules adjustment have been developed using eCognition 8 developer. OBIA processing for each mapping region was performed with four licenses of eCognition 8 server (for running 4 CPU on server) using tiling and stitching process.

2.4 Classification validation

As a metric accuracy of boundary location was not crucial for the final aim of the studies, we did not assess planimetric accuracy of boundaries but focused on assessing the semantic quality of the classification (i.e. assessment of object nature). Obviously, the semantic quality depends on planimetric accuracy. The semantic quality assessment was performed using a confusion matrix (Foody, 2002). The confusion matrices were computed after grouping classes according to the typology targeted for each study (see 1.2 for CS1 and 1.3 for CS2). Given that the features used for classification are calculated at the object scale, objects or polygons have been chosen as sampling units for the selection of control data (Grenier et al., 2008, Tiede et al., 2006).

However, confusion matrices were computed using the area of the selected control objects (expressed as a number of 0.5 m pixels) because the land-cover maps result from the implementation of the multilevel OBIA scheme and contains objects of different sizes, from fine to large. For each CS, to select control data with the objective of a spatially and thematically well distributed sample over the studied zone, as suggested by (Congalton, 2004), the entire sampling frame is divided into N equal grid cells. Next, a stratified random sampling is performed using grid cells as geographical strata (equal area for all strata). The number of cells (N) is equal to the desired sample size, in order to have at least one object of each class per grid cell. As suggested by (Congalton, 1991), 50 samples are collected for each class in order to build the confusion matrix. Considering the size of the study areas (and mapping regions), collecting field data for the control sample would be extremely labour intensive and time consuming. As suggested by (Zhu et al., 2000), selected control objects are photo-interpreted using the image with the highest spatial resolution as control data. To maintain objectivity of photo-interpretation, the classified maps were not viewed during the process.

2.5 Spatial analysis

2.5.1 Building riparian indicators (CS1)

A spatial indicator is invariably defined by aggregating a landscape structure attribute over a delimited area (spatial scale). Some examples of structural attributes are the number of different cover types, the proportion of each cover type, the shape of patches, and the spatial arrangement and connectivity of patches (Li et al., 1995). The domain over which spatial indicators are computed at the riparian scale for the modelling is generally defined by combining a lateral distance to the river with longitudinal distances upstream and downstream from the ecological station where stream ecological status is measured (Tormos et al., 2011).

We analysed here the land cover composition (area percentage of each land cover category) according to different buffer widths (5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 m) on 3-km upstream distance from an ecological station. We used the typology targeted for this study (see 1.2). According to the definition, this makes a total of 10 spatial indicators for each land cover category. These indicators were computed from H&VHRS derived map (resulting of OBIA) and CLC database in order to highlight the gain of H&VHRS remotely sensed data for characterizing RALC. This spatial analysis in this study was conducted using ESRI GIS tools.

2.5.2 Building artificial patches (CS2)

To quantify artificial sprawl we need before to compute artificial patches. Artificial patch is defined as a continuous geographic entity of neighbouring artificial objects (Dupuy et al., 2012). Artificial patches are built generally from a mathematical morphology operation, the dilatation and erosion algorithm (Haralick et al., 1987). Distant objects up to 50 m (Euclidean distance) were considered as neighbouring objects in

<table>
<thead>
<tr>
<th>Case studies</th>
<th>Data</th>
<th>Footprint</th>
<th>Spatial Resolution</th>
<th>Spectral resolution</th>
<th>Acquisition date</th>
<th>Number of Blocks</th>
<th>Producer</th>
<th>Pre-processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS1</td>
<td>Orthophotos</td>
<td>5 km</td>
<td>0.5 m</td>
<td>B,G,R</td>
<td>Summer period, between 2003 &amp; 2006</td>
<td>2455</td>
<td>IGN®</td>
<td>mosaic</td>
</tr>
<tr>
<td></td>
<td>SPOT 5 XS</td>
<td>60 km</td>
<td>10 m</td>
<td>G,R,PiR,S WIR</td>
<td>Summer period, between 2003 &amp; 2007</td>
<td>11</td>
<td>Spotimage© Archive images</td>
<td>TOA radiometric corrections and mosaic</td>
</tr>
<tr>
<td>CS2</td>
<td>RapidEye</td>
<td>25 km</td>
<td>5 m</td>
<td>B, G, R, RE, NIR</td>
<td>Summer 2009</td>
<td>91</td>
<td>RapidEye©</td>
<td>mosaic</td>
</tr>
</tbody>
</table>

Table 1. Image data characteristics for the two case studies

(B = Blue, G = Green, R = Red, RE = Red Edge, NIR = Near InfraRed, SWIR = Shortwave Infrared)
our study. This distance was chosen by trial and error runs of this algorithm using different Euclidian distances (Dupuy et al., 2012). Artificial patches were built from 2009 artificial object maps (resulting of OBIA) and 1997 artificial object maps. The spatial analysis in this study was conducted using ERDAS imagine tools.

3. RESULT

3.1 Classification accuracy and processing time

Fine-scale maps were obtained for the two case studies (see Figure 2 RALC map). Accuracy and time-efficiency of OBIA are summarized in Table 2. A high accuracy was obtained for the two case studies with a relatively operational time-efficiency.

<table>
<thead>
<tr>
<th>Case studies</th>
<th>Total accuracy</th>
<th>threshold adjustment &amp; processing time</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS1</td>
<td>85 %</td>
<td>16 hours / 1000 km²</td>
</tr>
<tr>
<td>CS2</td>
<td>90 %</td>
<td>12 hours / 1000 km²</td>
</tr>
</tbody>
</table>

Table 2. Total accuracy (average of total accuracies obtained for each mapping region) and time-efficiency of OBIA for the two case studies: riparian corridor mapping (CS1), artificial object mapping (CS2)

Figure 2. Extract of riparian area land cover map in obtained over the Normandie region (CS1) from OBIA combining SPOT X5, orthophotos and ancillary data.

3.2 Comparison of riparian spatial indicators (CS1)

Figure 3 shows the evolution of the land cover composition (area percentage) according to the lateral distance to the river channel (buffer width from 5 m to 50 m) from H&VHSR-derived map and CLC database. Not surprisingly, the spatial resolution of land-cover maps has a strong influence on our ability to quantify landscape indicators (Lattin et al., 2004). We can see that major landscape patterns inside the riparian areas (i.e., semi-natural vegetation types) which are revealed by H&VHSR maps are totally smoothed in CLC maps. According to CLC (Figure 3B), only two land cover category are present in riparian area (grassland and urban areas) and spatial information does not significantly change with the buffer width whatever the land cover category, in contrast to the H&VHSR-derived map (Figure 3A).

Figure 3. Riparian spatial indicators: Area percentages of a given land cover type according to the lateral distance to the river channel upstream of a stream ecological station (3km upstream). In A, built from High and Very High Spatial Resolution (H&VHSR)-derived map resulting of OBIA; in B built from CORINE Land Cover (CLC) database.

3.3 Map of artificial sprawl

Figure 4 shows an extract of the artificial sprawl map between 1997 and 2007 years. It demonstrates that this phenomenon occurs in Languedoc-Roussillon region (an artificial sprawl of 18% in total was observed over the study area) localized globally around each town whatever the size.
4. DISCUSSION

Through these two case studies, we confirm OBIA have revolutionized the processing of H&VHRS remote-sensing data and demonstrated that our OBIA approach is an operational method for processing reliable spatial information in order to conduct fine-scale spatial analysis over broad territories.

Our approach appears (1) reproducible: whatever the mapping region, defined in the two case studies, the total accuracy was good to very good (see Table 2); (2) easy transferable: despite the diversity of landscapes in study areas, it has been possible to use the same parameters of segmentation and the same class hierarchy on all mapping regions, without changing classification features: the major operator task was to adjust the threshold values of the different features used to define classes. This is partly due to the top-down approach for the construction of tree classification (that divides the feature space into finer and finer units) and the use of ‘knowledge-based rules’ classification technique. Top-down approach promotes the use of simple rules that are easier to transpose to other mapping regions and facilitates the appropriation of the methodology by new operators. Moreover specific classes of a given mapping region can be integrated in the tree classification without questioning its overall construction; (3) quickly applicable over broad areas: the OBIA processing is little time consuming (see Table 2). The adjustment time of rules obviously depends on the experience of the operator and his knowledge of the study area and its land-cover diversity. As observed by (Lucas et al., 2007, Tiede et al., 2010) “knowledge-based rules” classification technique appears more flexible: all rules could be refined with the full control of the user, at any time in the classification process and, in most cases, without changing the class allocation of other objects (which is generally not the case with the “supervised” classification technique). Moreover, while the knowledge-based rules’ classification technique can be distributed easily on a cluster machine, thus drastically reducing the computation time, the ‘supervised’ classification technique cannot be distributed because it requires the processing of a mapping region as a whole in order to collect a spatially representative training sample; and (4) Portable and scalable (i.e. able to integrate new data sources): the method can easily manage information from multiple data sources by resolving conflicts between these sources using fuzzy logic. This property has been particularly valuable in the CS1 to resolve conflicts between SPOT 5 XS and orthophotos data sources. For instance, information for assigning an object to one of these two classes, ‘soil with high vegetation’ and ‘soil with low vegetation’, can be contradictory because data were not acquired in the same season on the mapping region. Thus, fuzzy logic appears indispensable to treat multisource information in this OBIA classification. Thanks to fuzzy logic, new data sources could easily be integrated and combined with the initial data sources. This reliable and fine-scale spatial information over broad territory is crucial to better understand the structures, dynamics and changes in the environment, societies and territories. Through the CS1, we can point out the interest of RALC H&VHRS map over broad territory. This map characterizes more finely and reliably riparian landscape than traditional land cover databases (see Figure 3). It provides accurate and original information on the presence and intensity of pressures close to the stream and on forested riparian strip attributes (uniformity, mean width and continuity). Given these results, new challenges emerge in view of gaining a better understanding of the relationships between land cover pressure and stream ecological status. This better understanding will offer valuable information for managers in order to achieve the WFD aims in prioritizing and designing efficient river corridor restoration strategies. Through the CS2, we can note that fine-scale artificial sprawl maps over broad territories allow to better quantify and localize from local to regional scale the loss of land (natural or agricultural) to the detriment of built-up areas (in the centre and periphery of the city). Such information is essential for French managers in order to prioritize and define efficient restoration strategies dedicated to the conversion of land with agricultural potential.

5. CONCLUSION

To conclude, through these two case studies, concerning riparian corridor and artificial sprawl management from local to national scale, this paper demonstrates the drastic need of fine-scale information for supporting action strategies and the high interest of our OBIA approach for producing this information. Our approach, based on a top-down approach for the construction of the classification tree and the use of ‘knowledge-based rules’ classification technique appears operational given the high accuracy results and relatively short processing time obtained in the two case studies. OBIA, following our approach, will be probably at very short run a broadly applicable method for producing reliable spatial information dedicated to carry out broad scale spatial analysis supporting public policies.
REFERENCES


Grenier, M., et al., 2008. Accuracy assement method for wetland object-based classification In: Geographic Object Based Image Analysis (GEOBIA 2008), ISPRS Commision IV, Calgary, Alberta, Canada.


ACKNOWLEDGMENTS

The riparian corridor mapping study has been made possible by a PhD grant from the French Ministry of Research. The work has been partly realized in the framework of a research contract between ONEMA and Irstea. Acquisition of remote-sensing data benefited from the support of the ISIS programme of the French national space agency CNES (SPOT-5 images, distribution Spot Image), while IGN’s airborne orthophotos were made available by the French Ministry of Agriculture.

The artificial sprawl mapping study has been made possible by the DRAAF Languedoc-Roussillon (regional directorate of the French Ministry of Agriculture). Acquisition of remote-sensing data (RapidEye images) benefited from the support of the GEOSUD project (http://geosud.teledetection.fr/).