TOWARDS A TYPOLOGY OF LAND-COVER/LAND-USE EVOLUTIONS USING HIGH RESOLUTION SATELLITE IMAGE TIME SERIES: APPLICATION TO THE METROPOLITAN AREA OF STRASBOURG (FRANCE)

F. N. Güttler [a,b,c] *, A. Puissant [b], P. Gançarski [c]

[a] INPT, Ecole d'Ingénieurs de Purpan, UMR 1201 DYNAFOR, 31300 Toulouse, France – fabio.nor-guttler@purpan.fr
[b] Laboratoire Image, Ville, Environnement, UMR 7362 CNRS, University of Strasbourg, 67000 Strasbourg, France – anne.puissant@live-cnrs.unistra.fr
[c] ICube Laboratory, UMR 7357 CNRS, University of Strasbourg, Strasbourg 67000, France - gancarski@unistra.fr

KEY WORDS: Temporal GEOBIA, Change Detection, Land-cover Evolution, Urban and Peri-urban Areas, Fuzzy Learning

ABSTRACT

In this work, we propose an original approach to extract structured information from long-term satellite image time series (SITS) which can be employed to construct a typology of land-cover/land-use (LCLU) evolutions. Through a semi-supervised framework we combine per-pixel and object-based image analysis with data mining techniques. Our experiments are applied to the metropolitan area of Strasbourg (France) using nine high spatial resolution (20 m) SPOT images acquired between 1992 and 2011. The framework is organized in four main steps: (a) construction of image objects, (b) multi-strategy fuzzy learning approach to predict LCLU class membership at pixel level, (c) generation of structured information at object level and (d) mining structured information across the time-series.

1. INTRODUCTION

Satellite Image Time Series (SITS) is a relevant source of information for studying evolutions in urban and peri-urban areas. However follow such complex and dynamic areas is challenging. On the one hand, high spatial resolution imagery (< 30 m) is required to identify urban spatial organization and the main types of land-cover and land-use (LCLU) at a given date. On the other hand, most of the LCLU evolutions spread over long periods, usually ranging from few months to several years. Three key issues can be associated to the spatiotemporal analysis of SITS: (1) firstly, one should be able to detect where the evolutions took place in a given landscape, (2) then, it is important to determine precisely when those evolutions arise and, (3) finally, it is essential to be able to characterize the different types of evolutions, classifying them into categories with relevant semantic meaning.

As advertised by the scientific community, methods describing multi-temporal behavior are among open challenges in GEOBIA (Blaschke et al., 2014 ; Chen et al., 2012). Chen et al. (2012) organized the techniques used for object-based change detection (OBCD) in four categories: (1) image-object change detection where multi-dates images are segmented separately and the objects are directly compared; (2) class-object change detection where a classification is performed independently for each image before comparison; (3) multitemporal-object change detection where multi-date images are combined and segmented together to produce corresponding change objects and (4) hybrid change detection which combines both object and pixel-based image analysis techniques. Still, most of those change detection techniques cannot deal with long time series (i.e. more than four or five images). Also, they usually do not satisfy the

three key issues of the spatiotemporal analysis. In this work, in opposition to the classical approaches, we propose an original approach to extract information related to those three key issues and we tested our method in a long-term SITS containing nine optical images. Our approach is hybrid and combines pixel-based and object-based image processing with data mining techniques in a semi-automatic way. The paper is organized as follows: Section 2 presents the study area and the dataset. Section 3 describes all the steps of the methodology. Experimental results are presented and discussed in section 4. Conclusions are drawn in Section 5.

2. STUDY AREA AND DATA DESCRIPTION

2.1 Study Area

The study area corresponds to the metropolitan area of Strasbourg, located in the North-East of France (Figure 1). The metropolitan area consists of 28 municipalities and has a territory of about 316 km². The population is estimated in more than a half million of inhabitants. It is built around a concentric dense city centre inherited from the Middle Ages with surroundings organised in some rings characterized by typical urban fabrics such as discontinuous urban fabric dedicated to housing, leisure activities, commercial and industrial activities, etc. The city is also covered by some urban vegetation and is bordered to the East by an important river (the Rhine). Agricultural areas and patches of forest are present in most of the surrounding municipalities. During the last decades the study area experienced a classical urbanization process of urban sprawl around the inner-city like some European cities in Western Europe.
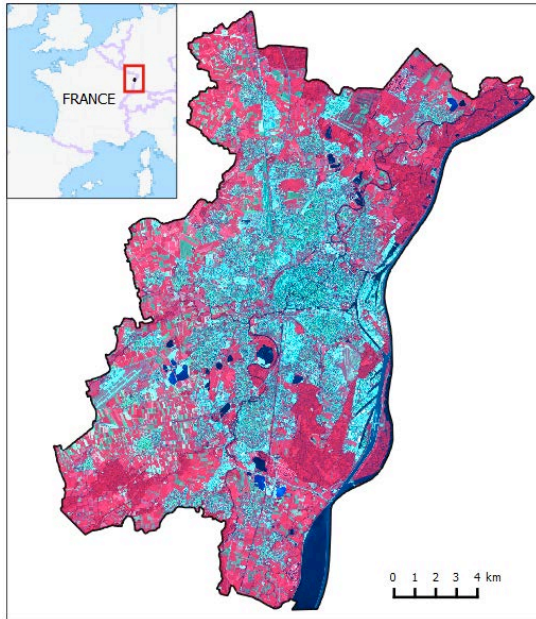
---

* Corresponding author

Figure 1. The metropolitan area of Strasbourg (Spot image ©CNES, 2011)

2.2 Data and pre-processing

The main source of data is a collection of nine high spatial resolution SPOT images acquired between 1992 and 2011 (see Table 1).

Acquisition date	Satellite	Pixel size (m)	Spectral bands*
7 Sept. 1992	SPOT 1	20 m	G, R, NIR
22 July 1996	SPOT 3	20 m	G, R, NIR
17 Aug. 1998	SPOT 2	20 m	G, R, NIR
25 July 1999	SPOT 4	20 m	G, R, NIR, SWIR
8 June 2000	SPOT 1	20 m	G, R, NIR
22 July 2001	SPOT 4	20 m	G, R, NIR, SWIR
5 Sept. 2004	SPOT 5	10 m	G, R, NIR, SWIR
5 July 2010	SPOT 5	10 m	G, R, NIR, SWIR
27 June 2011	SPOT 5	10 m	G, R, NIR, SWIR

* G = green, R = red, NIR = near infrared, SWIR = short wave infrared

Table 1. Acquisition date and main characteristics (before pre-processing) of the SPOT images used in this work.

In order to minimize changes related to the seasons (illumination conditions, phenology states, etc.), all the images were acquired during the summer (June to September) and a radiometric normalization process has been applied (Canty et al., 2004). As the original images have dissimilar spatial resolutions, some of them (those from SPOT 5) were resampled to generate a homogeneous SITS with a pixel size of 20 m. For the same reason, we removed the fourth band (short wave infrared) of the images acquired by SPOT 4 and 5. Therefore, only the remaining three bands (green, red, near infrared) were considered for the whole SITS. Finally, a spatial subset was performed to keep only the pixels inside the administrative limit of the Strasbourg metropolitan area as having valid values (see Figure 1).

3. METHODOLOGY

The methodological framework is organized in four steps. Step 1 focuses on the construction of two types of image-objects: (a) *mono-date objects* and (b) *time-series objects*. In step 2, several spectral and texture features are computed for each timestamp (or date) of the SITS. Step 3 corresponds to a fuzzy supervised learning procedure. It generates LCLU class membership at pixel level, for each timestamp of the SITS. In step 4, structured (thematic, spatial and temporal) information is obtained at object level (mono-date, bi-date and time-series). Finally, some data mining strategies (yet in development) are proposed in order to group similar trajectories and to contribute to the construction of a typology of LCLU evolutions which considers the whole SITS.

3.1 Construction of image objects

Mono-date objects: The processing chain starts with multi-scale image segmentation. Each image of the time series is segmented independently, but using strictly the same rule-set, which generates separate sets of mono-date objects.

The segmentations were obtained combining the Multiresolution Segmentation Algorithm (MSA) and the Spectral Difference Algorithm, both available in eCognition Developer 8.8.1. MSA is a bottom-up segmentation based on a pairwise region merging technique (Baatz and Schäpe, 2000). To decide whether to merge or not neighboring pixels (or objects) it uses a combination of spectral, shape and scale criteria. The following strategy was employed to tune the MSA algorithm (using two images of the SITS).

Firstly, the ESP2 tool (Drăguţ et al., 2014) was used to obtain an automated estimation of the scale parameter. As a result, the scale values 10 and 31 were selected. Then, with the scale parameter fixed at 10, we tested and visually compared all the combinations of the shape and compactness criteria (from 0.0 to 1.0 in intervals of 0.1). The most appropriated segmentation result was obtained with the following values: shape = 0.4 and compactness = 0.6. This segmentation was further enhanced by applying the Spectral Difference Algorithm (SDA value = 4), which merged spectrally similar objects previously separated by the scale parameter constraint. This segmentation result was named "fine level".

Above the fine level, a less detailed segmentation was generated using the second scale value (31) with the MSA. Again, all the combinations of the shape and compactness criteria were tested and compared. The most appropriated segmentation result was obtained with the following values: shape = 0.3 and compactness = 0.8. This result was named "medium level".

With the segmentation parameters presented above, a rule-set was created and applied over each timestamp of the SITS.

Time-series objects: Time-series objects (ts-objects) are the result of the topological overlay of all the mono-date objects of the SITS, for a given scale level. More in detail, each ts-object assembles the pixels belonging to the same sequence of mono-date objects across time, meaning they possess a similar spectral trajectory.

3.2 Computing pixel based spectral and texture features

Before starting the learning step, the optical images were processed and several spectral and textural attributes were computed at pixel level.

3.2.1 Spectral indices: In total, nine spectral indices were computed using the Orfeo Tool Box (OTB)[1]. Most of them are useful for a better characterization of the vegetation or soil. A detailed description of each index is presented in Table 2.

Index acronym and type	Index name	Formula
NDVI [Vegetation]	Normalized difference vegetation index	$\dfrac{nir - red}{nir + red}$
TSAVI [Vegetation]	Transformed soil adjusted vegetation index	$\dfrac{0.7(nir - 0.7red - 0.9)}{0.7nir + red + 0.08(1 + 0.7^2)}$
MSAVI [Vegetation]	Modified soil adjusted vegetation index	$\dfrac{(1 + L) * (nir - red)}{nir + red + L}$ where : $L = 1 - 2s * NDVI(nir - 0.4red)$
GEMI [Vegetation]	Global environment monitoring index	$\dfrac{n(1 - 0.25n) - red - 0.125}{1 - red}$ where: $n = \dfrac{2(nir^2 - red^2) + 1.5nir + 0.5red}{nir + red + 0.5}$
NDWI2 [Water]	Normalized difference water index 2	$\dfrac{green - nir}{green + nir}$
RI [Soil]	Redness index	$\dfrac{red^2}{green^3}$
CI [Soil]	Color index	$\dfrac{red - green}{red + green}$
BI [Soil]	Brightness index	$\sqrt{\dfrac{green^2 + red^2}{2}}$
BI2 [Soil]	Brightness index 2	$\sqrt{\dfrac{green^2 + red^2 + nir^2}{3}}$

Table 2. Spectral indices computed for each image of the SITS at pixel level.

3.2.2 Texture features: The whole pixel-based Haralick textures features (Haralick et al., 1973) were computed using the functions available at OTB (29 texture features in total). These features are organized in three sets: (a) Simple, (b) Advanced and (c) Higher order. After testing some correlation-based feature subset selection techniques (results not showed here), it was possible to eliminate the 14 most irrelevant and redundant features. The 15 remaining texture features are listed in Table 3. Texture features were calculated using a 5x5 sliding window (size in pixels).

Set of texture	Feature name
Simple	Correlation
Simple	Inverse Difference Moment
Simple	Cluster Shade
Simple	Cluster Prominence
Simple	Haralick's Correlation
Advanced	Mean
Advanced	Variance
Advanced	Sum Entropy
Advanced	Difference of Variances
Advanced	IC1 (Info Measures of Corr.)
Advanced	IC2 (Info Measures of Corr.)
Higher order	Grey-Level Non-uniformity
Higher order	Run Length Non-uniformity
Higher order	Run Percentage
Higher order	Low Grey-Level Run Emphasis

Table 3. Selected texture features computed for each image of the SITS at pixel level.

3.3 Computing land-cover / land-use class membership

3.3.1 Thematic labelling: Taking advantage that ts-objects group pixels sharing a similar spectral trajectory, we used them to guide the thematic labelling. In total, 95 ts-objects (from the medium scale level) were manually selected and labelled into LCLU classes (a label per timestamp). We used a two-level hierarchical nomenclature for this labelling task (see Table 4). The first level has four general LCLU classes; three of them are detailed in the second level that counts nine LCLU classes in total.

Level 1	Level 2
Bare areas	Bare agriculture Bare gravel
Build-up areas	Build dense Build industrial-commercial Build residential
Vegetation areas	Woody vegetation Herbaceous vegetation with high greenness values Herbaceous vegetation with low greenness values
Water areas	Water

Table 4. Land Cover/Land Use (LCLU) classes used in the labelling step.

Together, the 95 labelled ts-objects correspond to a total of 12,224 pixels.

3.3.2 Fuzzy learning strategy: Before starting the learning process, the original SITS was converted into a series of concatenated raster files, each of them containing 27 dimensions (i.e., the 3 former spectral bands + 9 spectral indices + 15 texture features). The collection of labeled pixels, described in the previous sub-section, was divided in two sets: (a) the training set composed of 8,623 pixels (70 ts-objects) and (b) the validation set with the remaining 3,599 pixels (25 ts-objects).

We designed a multi-strategy fuzzy learning approach that combines three machine learning algorithms: (a) Logistic Regression, (b) Random Forest and (c) Decision Tree Regressor. Coded in Python language, our approach uses the implementations of the above cited algorithms provided by the

scikit-learn library[2]. More in detail, our approach is based on the "one versus the rest" scheme, meaning that a specific model is constructed for each LCLU class at each timestamp.

The accuracy of the individual models is assessed through 10-fold cross-validation using solely the 8,623 pixels of the training set. Root Mean Squared Error (RMSE) was used as the reference metric for model assessment. This experimentation basis was primarily used for tuning the parameters of the three algorithms. Then, the best configuration of parameters per algorithm was selected; i.e. those producing the lowest RMSE considering all the timestamps. Finally, the selected models were deployed over the entire study area.

Model output is expressed as numerical predictions, at pixel level, for each LCLU class (membership values ranging from 0 to 1). In order to obtain a single membership value per class, the predictions of the three algorithms were combined by the mean of a weighted average. The weights are inversely proportional to algorithm's RMSE scores obtained for each LCLU class. (Section 4.1 presents results obtained on our dataset.)

At the end of the learning process, the averaged class membership values were exported and converted into a new series of multidimensional raster files, where the number of dimensions is equals to the number of LCLU classes.
After all, we used the validation set (3,599 labelled pixels not used during the learning process) to get an independent assessment of the quality of the results.

3.4 Generating structured information at object level

3.4.1 Mono-date objects and LCLU class membership: The fuzzy learning approach applied over the whole study area generates, for each pixel, a vector of LCLU class membership values. For the mono-date objects, the corresponding class membership vectors are computed by averaging pixel vectors. (Section 4.2 and 4.3 present results obtained on our dataset.)

After this operation, for each date, we obtain a set of mono-date objects described by their LCLU class membership. Two additional attributes are computed for each mono-date object. The first is the majority LCLU class (i.e., the label of the class presenting the highest membership value). The second is a confidence estimation of the predicted majority class. Such estimation is based on the membership value of the majority class itself and on its relative importance considering the membership values of the other classes.

3.4.2 Bi-date objects and the types of transition: Besides mono-date and time-series objects, we created also bi-date objects. They correspond to the topological overlay of the mono-date objects from two subsequent timestamps. These objects are computed to better characterize the transitions through the times series. The primary attribute computed for bi-date objects are the delta class membership values (i.e., the variation in terms of membership values) for each LCLU class, between two timestamps.
Then, we compute an additional attribute called 'type of transition', which takes into account the magnitude of the delta attributes computed for each LCLU class. Three types of

transition are considered as function of the observed magnitudes: (a) conservative, (b) semi-conservative and (c) non-conservative. Finally, another attribute, named 'major process', is computed for the bi-date objects presenting non-conservative or semi-conservative types of transition. The attribution of a major process relies on a formal definition of the possible/desired processes that one may be interested in detecting. In the present work, we formalize (as a set of conditions) the following processes: 'urbanization', 'de-urbanization', 'retrofit or replacement', 'vegetation increase', 'harvesting', 'poldering' and 'water extension'.

3.4.3 Time-series objects and the connexion of timestamps and transitions: The time-series objects can be viewed as a key element for accessing and connecting information across the time. In that light, all the information coming from mono-date and bi-date objects become connected and can be accessed as temporal sequences through ts-objects.
Each sequence consists of a list of transitions and each transition consists of a pair of mono-date objects. (Section 4.4 presents results obtained on our dataset.)

3.5 Mining structured information

The temporal sequences are structured in a database allowing multidimensional queries (e.g. one may want to map the spatial footprint of all the vegetation areas converted into residential areas before the year 2000). The objective is to be able to query the database and to generate a map with all the associated temporal and thematic information from such queries.
The next step consists in processing ts-objects sequences using specific tools of sequential pattern mining and temporal clustering. More in detail, sequences presenting similar temporal patterns may be grouped together and their corresponding frequencies should be assessed. This result will provide a synthesis of the temporal patterns based on thematic meaningful transitions. (Section 4.5 presents results obtained on our dataset.)

The paper focuses mainly on the other four steps as the mining step is yet in development.

4. RESULTS

In this section we mainly present on the results corresponding to step 3 and 4 of our methodology.

4.1 Multi-strategy fuzzy learning assessment

The accuracy assessment of the three learning algorithms (individually) and of the combined model are showed in Table 5.
As a global result, predictions from the combined model present lower RMSE for most of the LCLU classes. The best predictions were obtained for water and bare areas as well as for herbaceous vegetation (low greenness values). Intermediary quality predictions are observed for the classes build industrial-commercial, woody vegetation and herbaceous vegetation (high greenness values). The less precise predictions are related to the classes build residential and build dense.

4.2 Class membership for mono-date objects

At each timestamp of the SITS, mono-date objects are described by LCLU class membership values, ranging from 0 to 1. Figure 2 shows the spatial distribution of membership values for three LCLU classes (out of nine) in 1992.

---

[2] Python machine learning library, version 0.17.1, available at: http://scikit-learn.org/

Class label	Logistic Regression		Random Forest		Decision Tree Regressor		**Combined model**	
	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev
Bare agriculture	0.113	0.066	0.108	0.075	0.121	0.068	0.102	0.069
Bare gravel	0.100	0.033	0.089	0.027	0.113	0.030	0.089	0.026
Build dense	0.220	0.024	0.210	0.020	0.243	0.048	0.208	0.031
Build ind-com.	0.144	0.024	0.134	0.032	0.148	0.035	0.129	0.029
Build residential	0.267	0.027	0.269	0.027	0.294	0.032	0.267	0.027
Woody vegetation	0.159	0.028	0.164	0.033	0.174	0.018	0.157	0.026
Herb. veg. h-green	0.183	0.042	0.168	0.042	0.182	0.032	0.167	0.038
Herb. veg. l-green	0.114	0.033	0.080	0.038	0.100	0.038	0.087	0.035
Water	0.009	0.015	0.014	0.013	0.030	0.049	0.015	0.020
Global RMSE	0.145	0.068	0.137	0.077	0.156	0.071	0.136	0.072

Table 5. Global and per-class Root Mean Squared Error (RMSE) for the three learning algorithms (individually) and for the combined model. RMSE mean and standard deviation calculated using the validation set for the nine timestamps of the SITS.
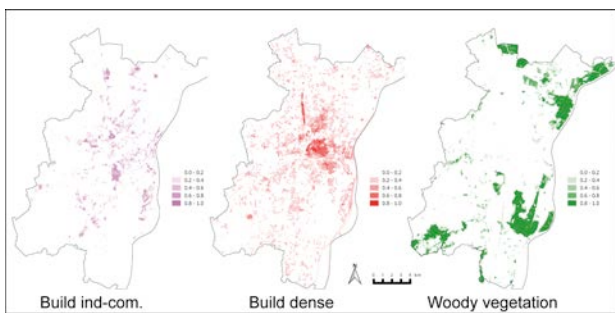

Figure 2. Spatial distribution of membership values for three land-cover/land-use (LCLU) classes: *Build industrial-commercial, Build dense, Woody vegetation*. Mono-date objects from the first timestamp of the SITS (year 1992).

In order to obtain a less fragmented view of the study area, it is useful to draw the map of the majority LCLU class predicted for each object. Figure 2(A) presents the majority class map obtained for the first image of the series (1992). The second part of the Figure 2(B) represents the confidence of the predictions.

As a general result, higher levels of confidence are observed in woody vegetation and water areas. The lowest levels of confidence are frequently associated to peri-urban areas where the edge separating build-up areas and agricultural areas is not easy to recognise.

4.3 Class membership temporal trajectories

Temporal information across the SITS can be obtained from the ts-objects. Indeed, all the attributes computed for mono-date and bi-date objects can be accessed through the ts-objects. For instance, it is possible to compute and analyse class membership trajectories for any ts-object. Figure 3 shows three different class membership trajectories obtained for ts-objects. For a better visual analysis, we plotted the trajectories using the more general level of the LCLU nomenclature (i.e. level 1 containing four general classes).
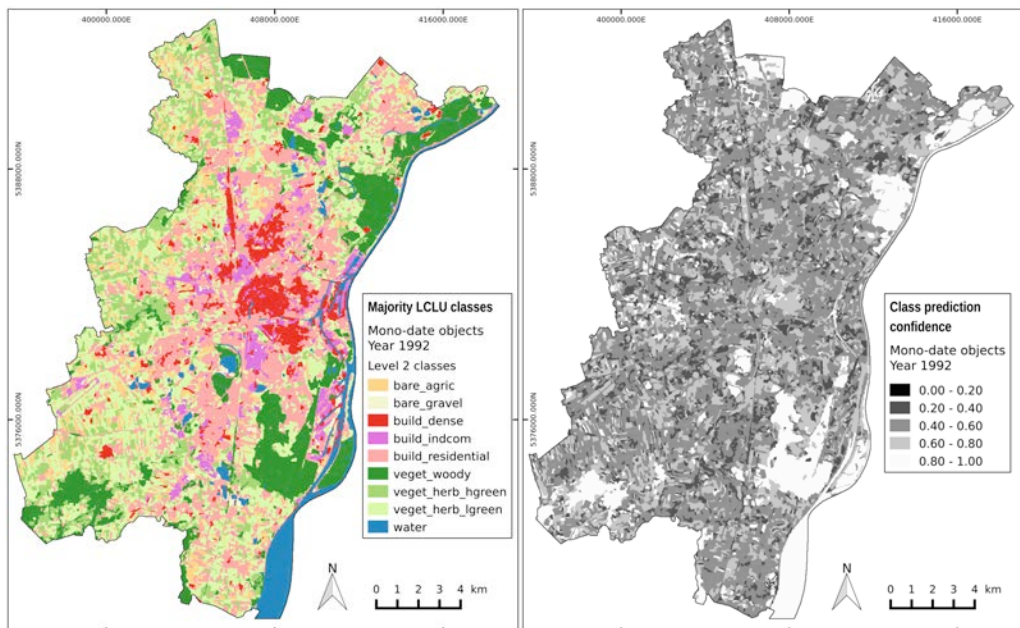

Figure 3. (A) Majority land-cover/land-use (LCLU) classes predicted from the first image of the series. (B) Majority class prediction confidence for the first image of the series.
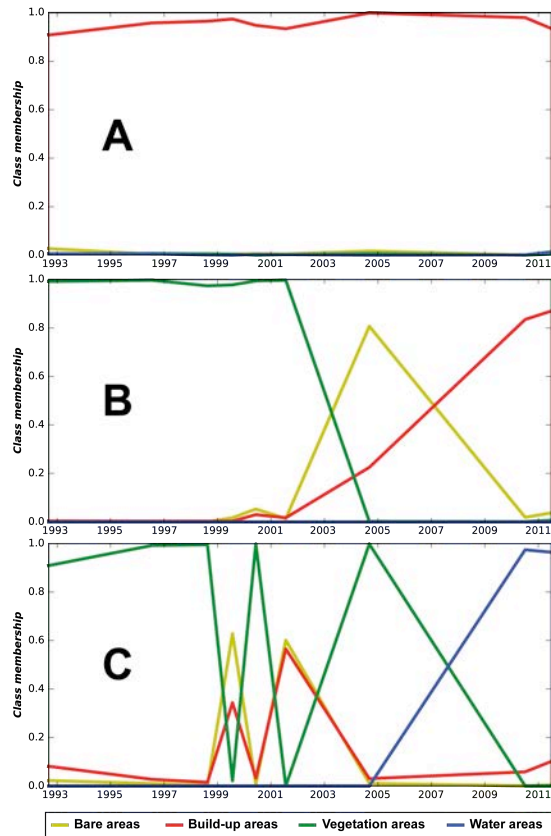
Figure 4. Four class membership trajectories for three ts-objects. (A) Stable area in the city centre of Strasbourg, (B) Vegetated area that was urbanized in the second part of the series (C) Agricultural crop area converted into a gravel extraction pond by the end of the series.

Plot A of Figure 4 shows the membership trajectories of a ts-object located in the historical city centre of Strasbourg. The class membership values are very stable during the entire time-series, with a clear dominance of the build-up class over the other three classes.

In plot B the temporal behaviour remains stable until the middle of the series, with vegetation membership values around 1. Between the timestamps of 2001 and 2004, the vegetation membership value falls to 0 while the bare class becomes dominant. However, in the next transition (2004 – 2010) the build-up class membership replaces the position of the bare class and this configuration is maintained until the final timestamp (2011). The temporal pattern observed here corresponds to the urbanisation of a meadow area that has been converted into a residential area.

Plot C presents a more complex temporal pattern. It corresponds to an agricultural crop area located next to a water pond used for gravel extraction. Until 2004, the crop field was observed whether covered by vegetation (when the image was acquired before crop harvesting) or without vegetation (when the image was acquired after crop harvesting). Between 2004 and 2010, the pond has been expanded and the water replaced the former crop field.

4.4 Transitions and cumulated intensity of change

Transitions can be analysed as function of the observed magnitudes of change in the class membership values. In order to take into account all the transitions, we computed a global score that considers the types of transition (conservative, semi-conservative or non-conservative) for of each bi-date object. This score is a cumulated sum of the non-conservative and semi-conservative transitions, where the first worth 1 and the second 0.5. Figure 5 shows the spatial distribution of the cumulated intensities of change.
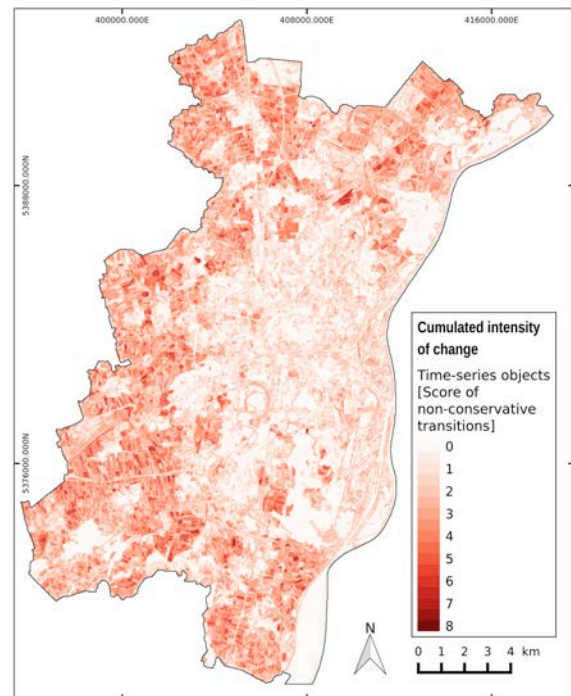

Figure 5. Cumulated intensity of change computed considering all the individual transitions of the SITS.

4.5 Long-term evolutions

In order to associate the three key issues to the spatiotemporal analysis of SITS, it is essential to able to characterize the different types of evolutions, grouping them into categories with relevant semantic meaning. First results of long-term evolution are presented in Figure 6. The most relevant is the phenomenon of urban growth from 1992 to 2011.

More precisely, we detect where urbanization took place within the study area. As this type of process was computed for bi-date objects, it is possible to determine between which pair of timestamps the urbanization has been detected. Then, we can calculate the total area becoming urbanized at each transition and compute the cumulated values for the whole SITS (see graph at the bottom of Figure 6). The analyses of other long-term evolutions are in progress, as for instance 'urban retrofit in already build-up areas' and 'expansion of water ponds related to gravel extraction'.
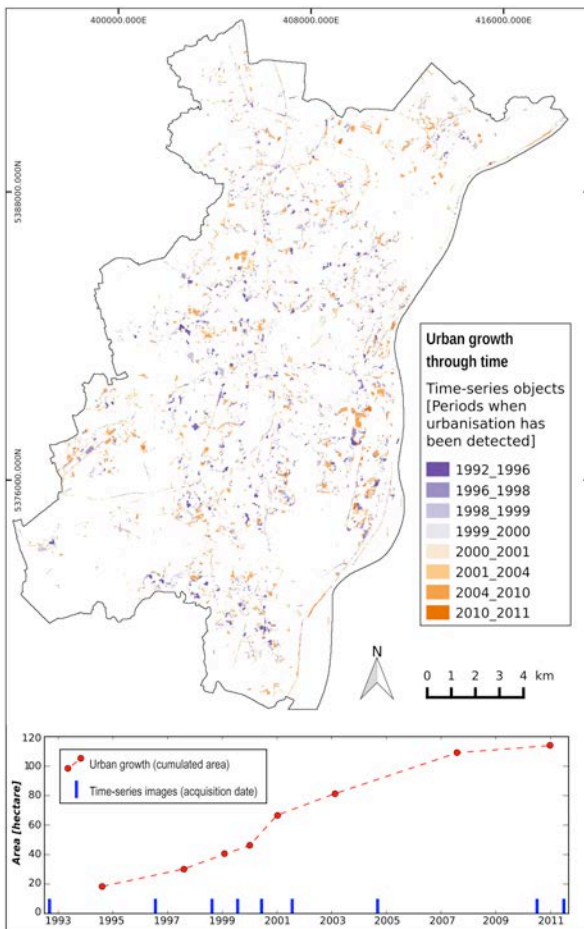
Figure 6. Urban growth from 1992 until 2011. Map presents the different periods when urbanization has been detected. Graph shows the cumulated curve of urbanized areas throughout the transitions of the SITS.

5. CONCLUSIONS

In this paper, we proposed an original and hybrid framework combining per-pixel and object-based image analysis of SITS. Information related to the three key issues of spatiotemporal analysis (where / when / what) has been extracted from a long-term SITS containing nine optical images. First results are relevant and the information extracted is structured in a database allowing multidimensional queries. Further developments are necessary in the mining step (e.g., temporal clustering and sequential pattern mining) in order to propose a typology of LCLU evolutions considering the whole SITS.

ACKNOWLEDGEMENTS

REFERENCES

Baatz, M., Schäpe, A., 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Information-sverarbeitung*, XII(58), pp. 12-23.

Baraldi, A., Boschetti, L., 2012. Operational automatic remote sensing image understanding systems: Beyond geographic object-based and object-oriented image analysis (GEOBIA/GEOOIA). Part 1: Introduction. *Remote Sensing*, 4(9), pp. 2694-2735.

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F., Tiede, D., 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing* 87, pp. 180-191.

Canty J., Nielsen A., Schmidt M., 2004. Automatic radiometric normalization of multitemporal satellite, *Remote Sensing of Environment*, 91, pp. 441–451.

Chen, G., Hay, G.J., Carvalho, L.M.T., Wulder, M.A., 2012. Object-based change detection. In: *International Journal of Remote Sensing* 33, pp. 4434-4457.

Drăguţ, L., Csillik, O., Eisank, C., & Tiede, D., 2014. Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 88, pp. 119-127.

Haralick, R. M., Shanmugam, K., 1973. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6), pp. 610-621.