Reading group

Updating land cover data: a multi-source and multi-modal approach for change qualification and land use characterization by deep learning

Martin CUBAUD

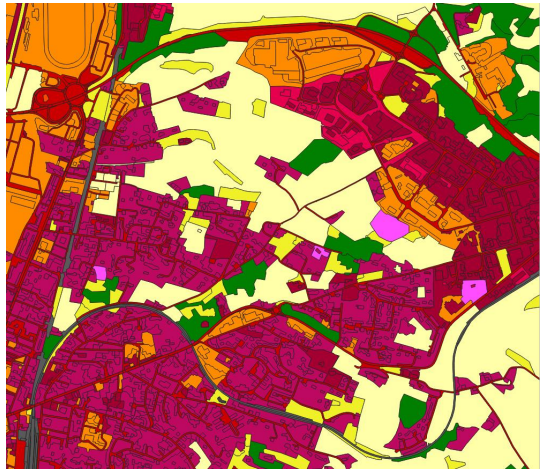
27/01/2023

PhD director: Ana-Maria Raimond; Supervisors: Laurence Jolivet, Arnaud Le Bris

Context

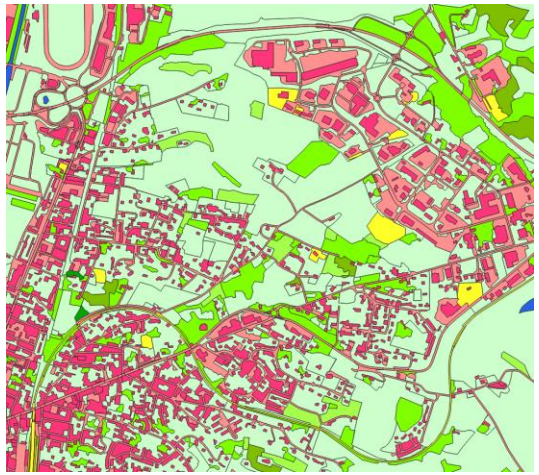
- Land Use/Land Cover, useful data for spatial planning and fight against soil artificialisation.

Orthophoto



Land Use

- Agriculture
- Sylviculture
- Activités d'extraction
- Pêche et Aquaculture
- Production secondaire
- Production secondaire tertiaire et usage résidentiel
- Production tertiaire
- Réseaux routiers
- Réseaux ferrés
- Réseaux aériens
- Réseaux de transport fluvial et maritime
- Services logistiques et de stockage
- Réseaux d'utilité publique
- Usage résidentiel
- Zones en transition
- Zones abandonnées
- Sans usage
- Usage inconnu



Land Cover

- Zones bâties
- Zones non bâties
- Zones à matériaux minéraux
- Zones à autres matériaux composites
- Sols nus
- Surfaces d'eau
- Peuplements de feuillus
- Peuplements de conifères
- Peuplements mixtes
- Formations arbustives et sous-arbrisseaux
- Autres formations ligneuses
- Formations herbacées

Context

- 2 main issues with institutional LULC maps :
  - Update frequency
  - Semantic precision
- Automatisation by machine learning :
  - From remote sensing data for land cover
  - Harder for land use

PhD objectives

To improve land use description

To detect and qualify land cover and land use changes

To update land use geometricaly and semanticaly

Construction and qualification of training and validation datasets

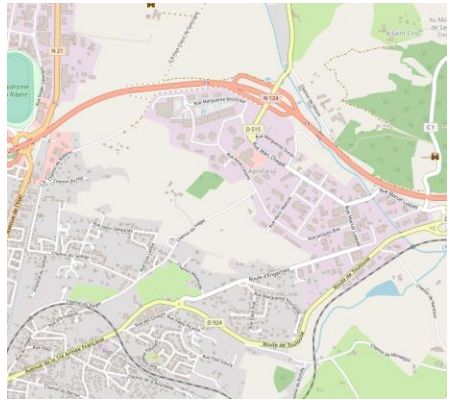
A multi-modal and multi-source approach

Raster :

- Raw or processed remote sensing data

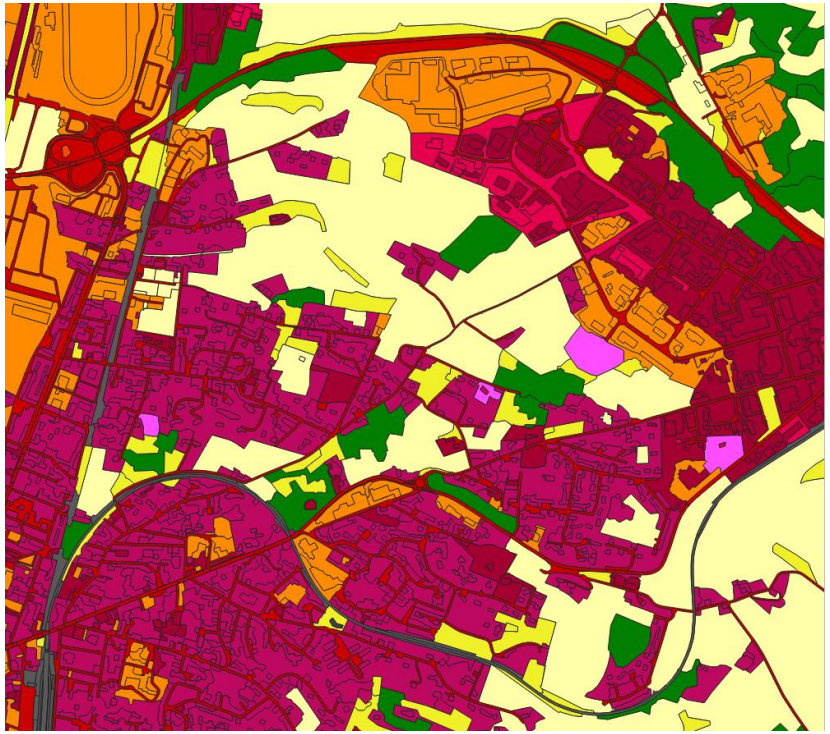
Vector :

- Authoritative databases
- Voluntary or unvoluntary geographic informations

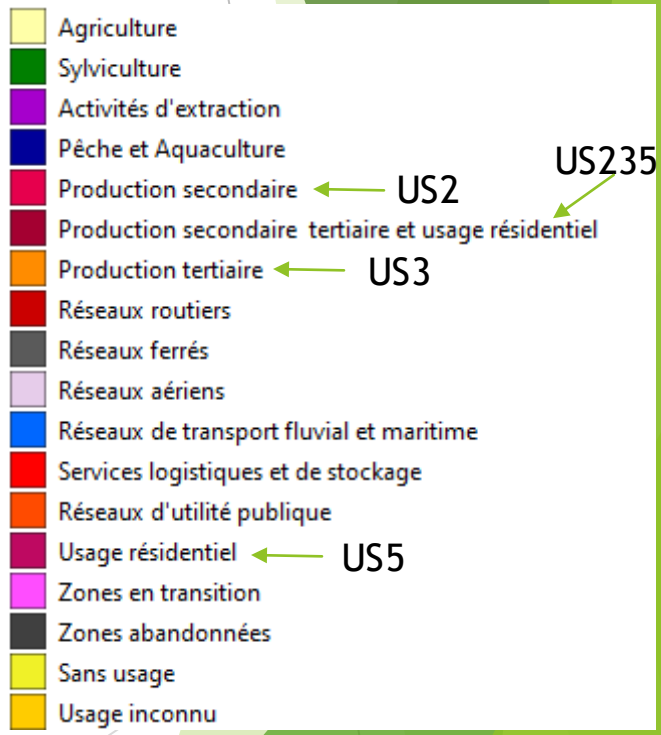


First point : land-use description improvement -> distinction US235

Orthophoto



Land Use



First point : land-use description improvement -> distinction US235

- Dataset : OCSGE polygons having land use US2, US3 or US5 in Gers in 2019
- Features : OCSGE land cover, geometric descriptors, intersection with BD TOPO, OSM, CLC, OSO, IRIS, land files, mean and std of each band of BD Ortho, and the mean of those features in the adjacent polygons.
- Approaches : Machine learning, data fusion
- Best machine learning model so far : Random forest (optimized in cross-validation)
  - Accuracy : 0,91
  - Confusion matrix :

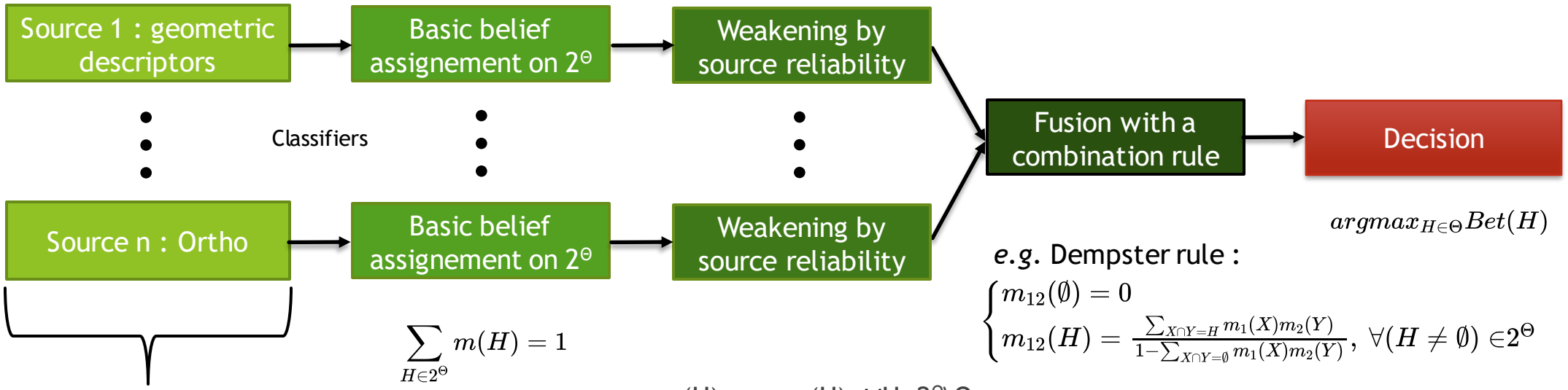
    [[171 16   3]
    [6 161 11]
    [1 10 194]]

Second approach : data fusion using Dempster-Shafer Theory

- DST : Generalisation of the Bayesian formalism which is able to deal with uncertainties

Frame of discernment
Θ = {US2, US3, US5}
Exhaustive and exclusive hypotheses

Referential of definition
2^{Θ} = {US2, US3, US5, US2 ∪ US3, US2 ∪ US5, US3 ∪ US5, US2 ∪ US3 ∪ US5 = Θ}



$$\sum_{H \in 2^{\Theta}} m(H) = 1$$

$m_{\alpha,i}(H) = \alpha_i\, m_i(H),\ \forall H \in 2^{\Theta} \setminus \Theta$
$m_{\alpha,i}(\Theta) = \alpha_i\, m_i(\Theta) + (1 - \alpha_i)$

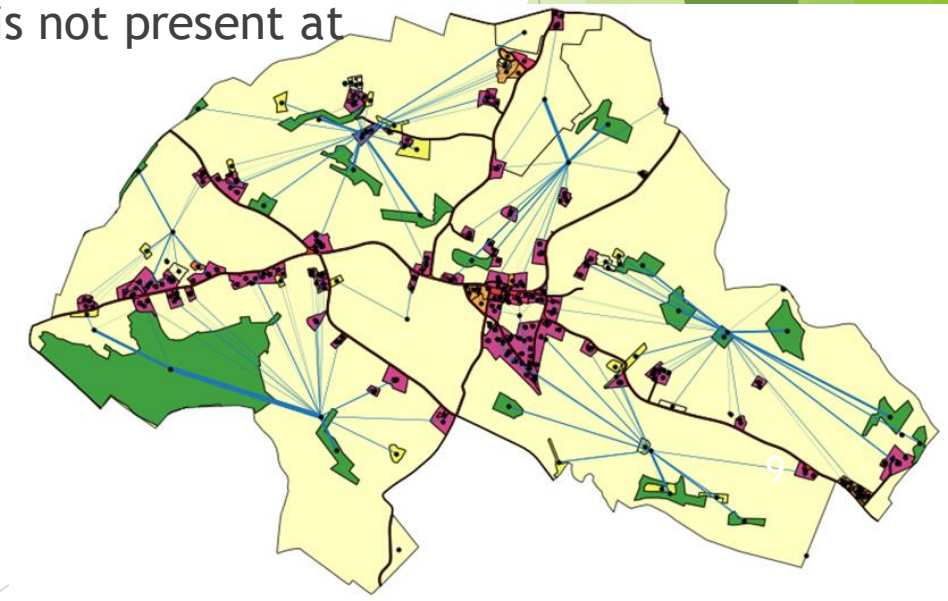
$$argmax_{H \in \Theta} Bet(H)$$

e.g. Dempster rule :

$$\begin{cases} m_{12}(\emptyset) = 0 \\ m_{12}(H) = \frac{\sum_{X \cap Y = H} m_1(X) m_2(Y)}{1 - \sum_{X \cap Y = \emptyset} m_1(X) m_2(Y)},\ \forall (H \neq \emptyset) \in 2^{\Theta} \end{cases}$$

Third approach : Graph Neural Networks

- Take into account both the spatial structure of the data and the features
- 1 graph represents 1 municipality, each node is a polygon and an edge exists between two adjacent polygons
- Land uses not in [US2, US3, US5] are present for information propagation but masked for loss backpropagation and metrics.
- I struggle with class imbalance (for most graphs the class US2 is not present at all)
- For now it doesn't work that well



Thanks for your attention