

Méthodes et outils pour la construction de la base de données MApUCE

Auteurs : E. Bocher, G. Petit

*Contact : erwan.bocher@univ-ubs.fr
gwendall.petit@univ-ubs.fr*

Laboratoire Lab-STICC (CNRS - UMR 6285) - UBS



Plan

- Introduction
- Données de référence
- Nouvelles données géographiques
- Indicateurs
- Classification du tissu urbain
- Langages et traitements
- Chaînes de traitements
- Modèle de base de données
- Architecture logicielle
- Interface de pilotage dans OrbisGIS
- Interface de publication des résultats



Mettre en place une chaîne d'analyses pour calculer les indicateurs géographiques utilisés par le modèle TEB.

- Données homogènes sur le territoire,
- Méthode générique et reproductible,
- Basée sur de nouvelles unités géographiques,
- Écosystème d'outils ouverts et interopérables.

Données de référence



- BD Topo IGN (2014)
- Carroyage de l'INSEE (2010)
- Contours IRIS + données INSEE (2011)
- Recensement de la population (2011)

Le Lab-STICC héberge les données issues des différentes équipes + les référentiels (ex : IGN, INSEE, ...)

→ 1 schéma par producteur



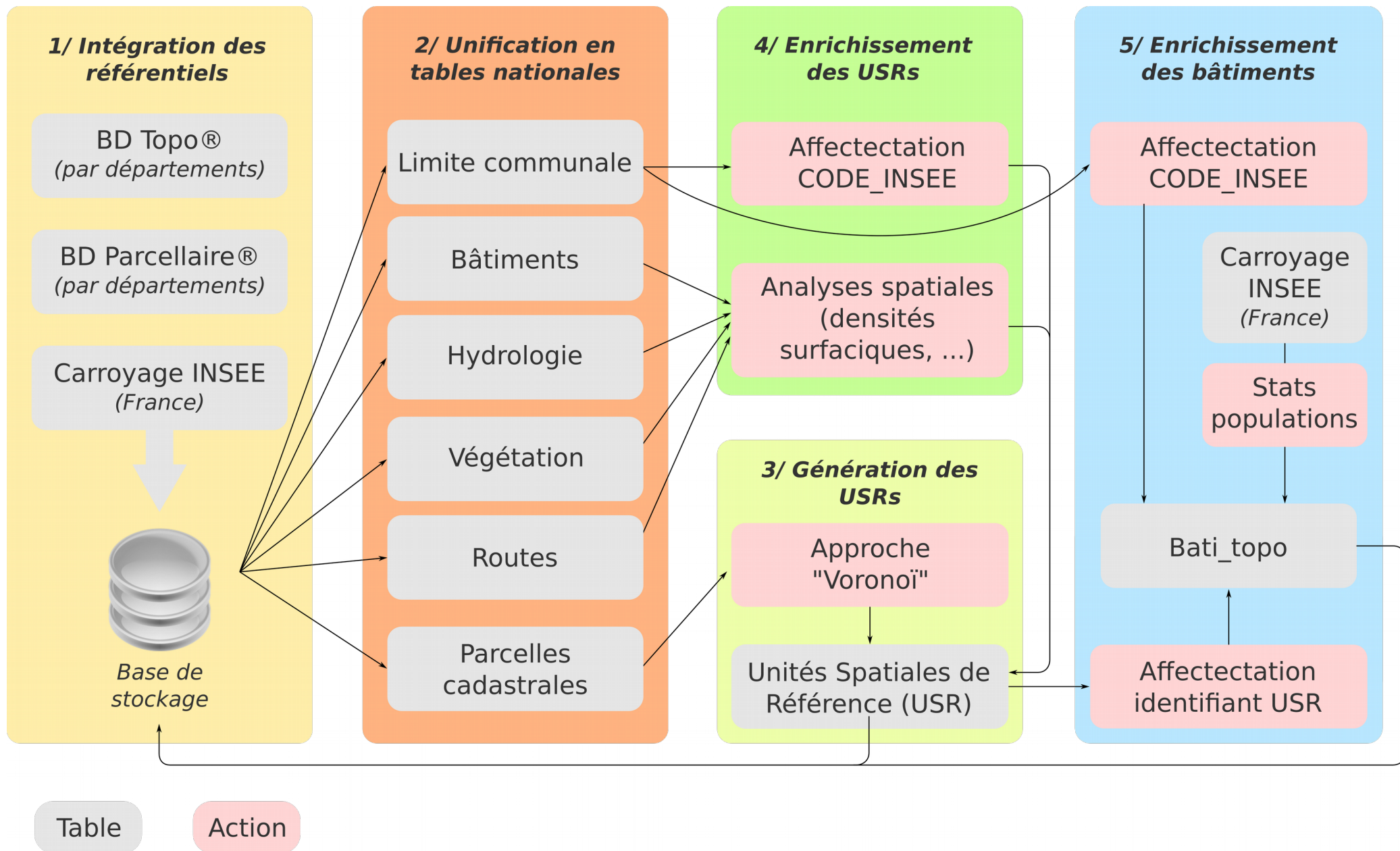
Quelques chiffres

(au 11/05/17)

Schéma	Table	Nb objets
LIENSs	Zone_etude	2238
	bati_topo	9 096 780
	usr	454 309
	unites_urbaines	80
Lab-STICC	buildings_indicators	3 726 108
	block_indicators	2 431 084
	usr_indicators	181 752

Soit environ 1/3 déjà traité ←

Nouvelles données géographiques



Nouvelles données géographiques

Table USR

Nom colonne	Type de données	Description
pk	int4	Identifiant unique pour l'USR
the_geom	geometry	Geometrie de l'USR
id	varchar	Identifiant de la BD TOPO
prec_plani	numeric	Précision planimétrique
prec_alti	numeric	Précision altimétrique
origin_bat	varchar	Source du bâtiment
hauteur	numeric	Hauteur du bâtiment
z_min	numeric	Altitude minimale du bâtiment
z_max	numeric	Altitude maximale du bâtiment
idzone	int4	Identifiant unique de la zone. Fait référence à la table zone.
idilot	int4	Unique identifier for USR
theme	varchar	Theme de la BD TOPO. Par exemple : RESEAU_ROUTIER
pai_bdtopo	varchar	Nom du PAI
pai_nature	varchar	Nature du PAI
nb_niv	int4	Nombre de niveau d'un bâtiment calculé à partir de la hauteur corrigée.
hauteur_corrige	numeric	Hauteur corrigée
insee_individus	float8	Nombre d'individus calculé à partir des données carroyées de l'INSEE

Table Bâtiments

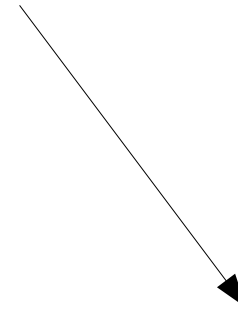
Nom colonne	Type de données	Description
route_surface	float8	Area of roads intersecting the USR. Computing is using the length of road segments intersecting the USR, and length of a clip of each segment with islet frontiers is multiplied by a buffer having the road width divided by 2. Nota: roads with fictif to true have automatically a width of 0, and 65% of secondary roads with importance=5 and fictif=false have a null width.
route_longueur	float8	Linéaire de routes qui intersectent l'USR.
route_nbpaistransport	int4	Nombre de PAI transport contenus dans une USR. Peut aider à déterminer la nature des surfaces non déterminées de l'USR.
trottoir_longueur	float8	Perimeter of the included ilot made of union of contiguous parcels. Sûrement un très mauvais indicateur. Lui préférer route_longueur.
hydro_surface	float8	Superficie des surfaces en eau (RESERVOIR_EAU + SURFACE_EAU)
hydro_longueur	float8	Linéaire des tronçons hydrographiques qui intersectent l'USR.
insee_individus	float8	Number of inhabitants computed by sum of share of INSEE grid cells intersecting buildings of the islets, proportionally to their developed area (nb_niv * area(building)), and only if the pai_nature of the building is null (which means residential a priori).
insee_menages	float8	Nombre de ménages résidant calculés par addition des parts de rectangles INSEE de 200 m intersectant l'USR
insee_men_coll	float8	Nombre de ménages en logement collectif calculés par addition des parts de rectangles INSEE de 200 m intersectant l'USR
insee_men_surf	float8	Surfaces cumulées des résidences principales en mètres carrés calculés par addition des parts de rectangles INSEE de 200 m intersectant l'USR
insee_surface_collectif	float8	Estimation de la surface de logement en collectif à partir des variables INSEE (=somme des insee_men_coll/insee_menages*insee_men_surf des surfaces de rectangles intersectés)
code_insee	varchar	Code insee de la commune

PAI = points d'activité ou d'intérêt

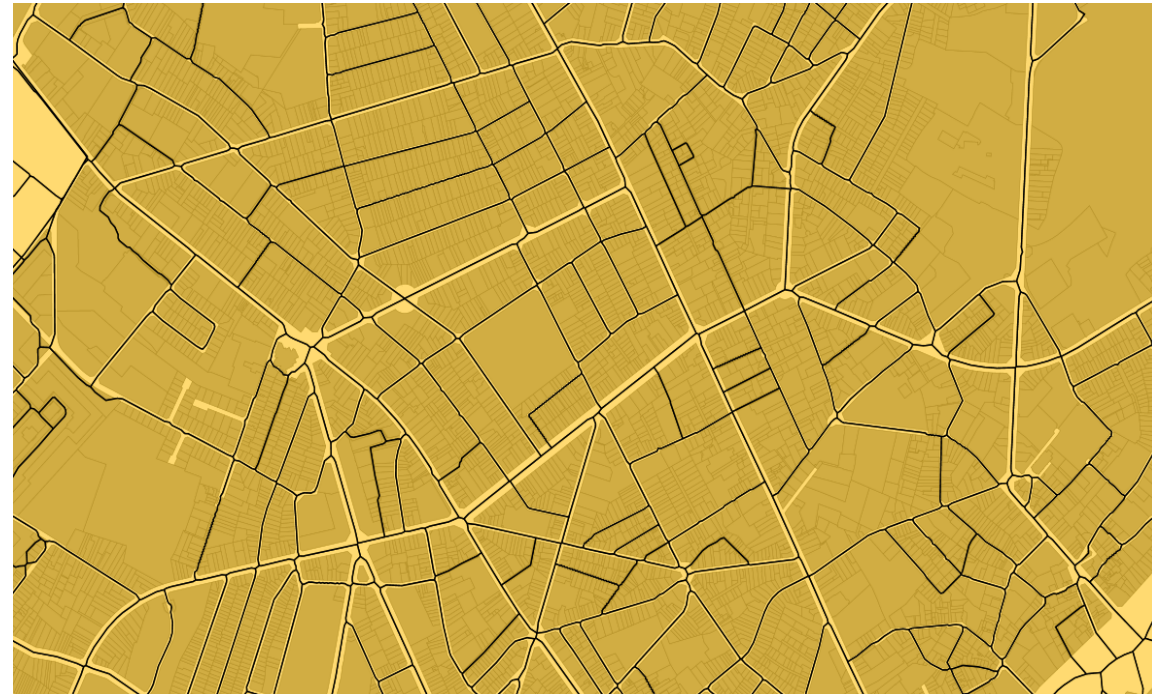
Nouvelles données géographiques



Du parcellaire cadastral

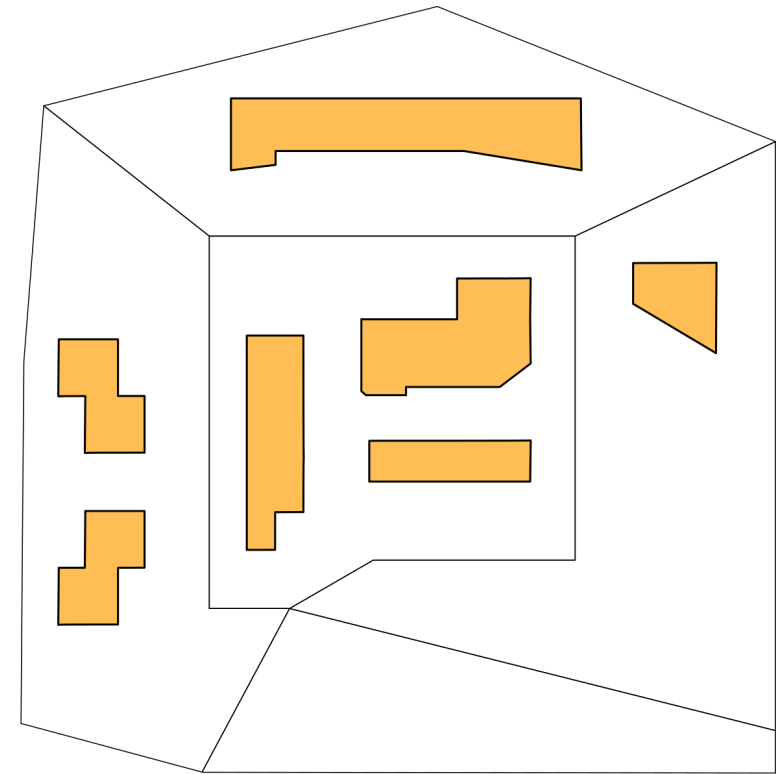
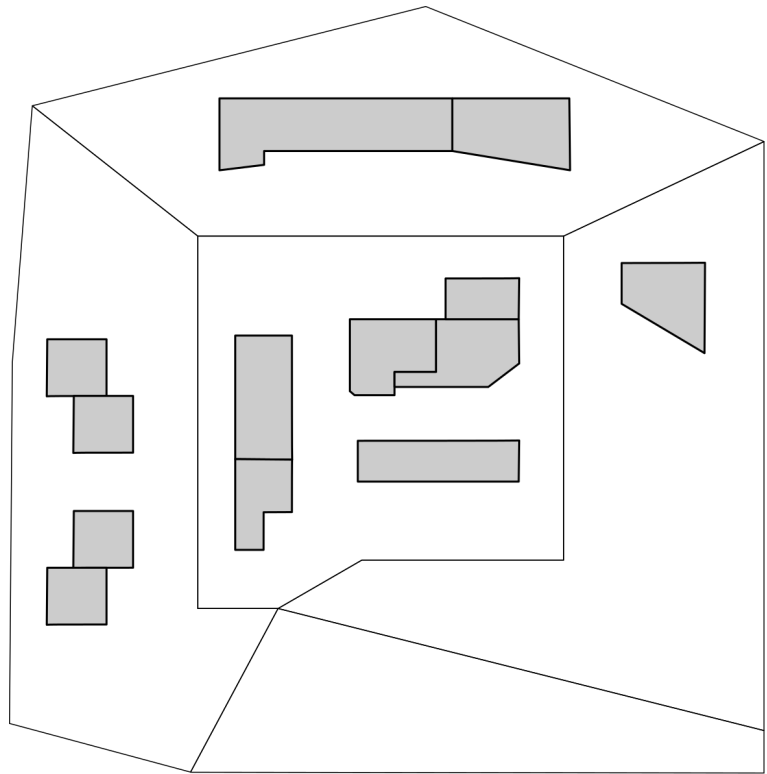


Aux USR



Indicateurs

- Plus de 50 indicateurs identifiés (étude bibliographique)
- Calculés pour 3 niveaux géographiques



■ Bâtiment

■ Bloc

□ USR

Indicateurs

... pour les bâtiments (27)

Code	Définition
area	Surface au sol (S)
floor_area	Surface de plancher (S x Nb Niv)
vol	Volume (S x H)
perimeter	Périmètre
l_cvx	Périmètre du ConvexHull
form_factor	Facteur de forme
concavity	Concavité
nb_neighbor	Nombre de voisins
b_wall_area	Surface de façade (incluant celles intérieures)
p_wall_long	Longueur de mur mitoyen
p_wall_area	Surface de mur mitoyen
free_ext_area	Surface de mur libre (non mitoyen et donc en contact avec l'air)
contiguity	Contiguïté : ratio entre la somme des surfaces libres divisée par la somme des surfaces extérieurs du bâtiment
compacity_n	Compacité (LRA) nette
compacity_r	Compacité (LRA) brute
compactness	Compacité (Gravelius)
main_dir	Direction principale
p_vol_ratio	Volume passif : ratio entre somme des surfaces de plancher à - de 6m d'une façade libre et somme des surfaces de plancher
fractal_dim	Dimension fractale
min_dist	Distances (minimale, moyenne, écart-type et maximale) entre le bâtiment et tous les autres contenus dans la même USR
num_points	Nombre de points
l_3m	Longueur de façade à moins de 3m d'une route
l_ratio	Ratio entre la longueur de façade à moins de 3m d'une route et le périmètre
l_ratio_cvx	Ratio entre la longueur de façade à moins de 3m d'une route et le périmètre du ConvexHull

Code	Définition
area	Surface au sol (S)
vol	Volume (S x H)
floor_area	Sommes des surfaces de plancher
h_mean	Hauteur moyenne
h_std	Ecart-type de hauteur
compacity	Compacité
main_dir_deg	Direction principale
holes_area	Surface de cours intérieure
holers_percent	Ratio de cours : rapport entre la sommes des surfaces de cours intérieures et la surface du bloc

Code	Définition
b_area	Somme des surfaces des bâtiments,
floor	Somme des surfaces de plancher des bâtiments,
floor_ratio	Ratio de plancher : somme des surface de plancher divisé par la surface de l'USR des bâtiments,
b_vol	Somme des volumes des bâtiments,
b_vol_m	Moyenne des volumes des bâtiments,
ext_env_area	Somme des surfaces extérieures (libres et donc non mitoyennes) de bâtiment,
compac_mean_w	Compacité moyenne (pondérée / non pondérée par la surface de l'USR) des bâtiments,
contig_mean	Contiguïté moyenne (pondérée par la surface de l'USR) des bâtiments,
contig_std	Écart-type de la contiguïté des bâtiments,
main_dir_std	Écart-type de la direction principale des bâtiments,
h_mean	Hauteur moyenne (pondérée par les surfaces de bâtiments),
h_std	Écart-type de la hauteur des bâtiments,
p_vol_ratio_mean	Moyenne des volumes passifs (pondérée par les surfaces de plancher)
build_numb	Nombre de bâtiment
min_m_dist	Moyenne des distances minimales entre les bâtiments,
mean_m_dist	Moyenne des distances moyennes entre les bâtiments,
mean_std_dist	Écart-type des écart-type de distance entre les bâtiments,
b_holes_area_mean	Moyenne des ratios de cours (pondérée par la surface des blocs)
b_std_h_mean	Écart-type des hauteurs de blocs (pondéré par la surface des blocs)
b_m_nw_compacity	Moyenne des compacités de blocs (pondérée / non pondérée par la surface des blocs)
b_std_compacity	Écart-type des compacités de blocs
dist_to_center	Distance entre les centroïdes de l'USR et celui de la ville dans lequel il s'inscrit
build_dens, ...	Densité surfacique de bâti, de route, de végétation et de surface en eau

Classification du tissu urbain

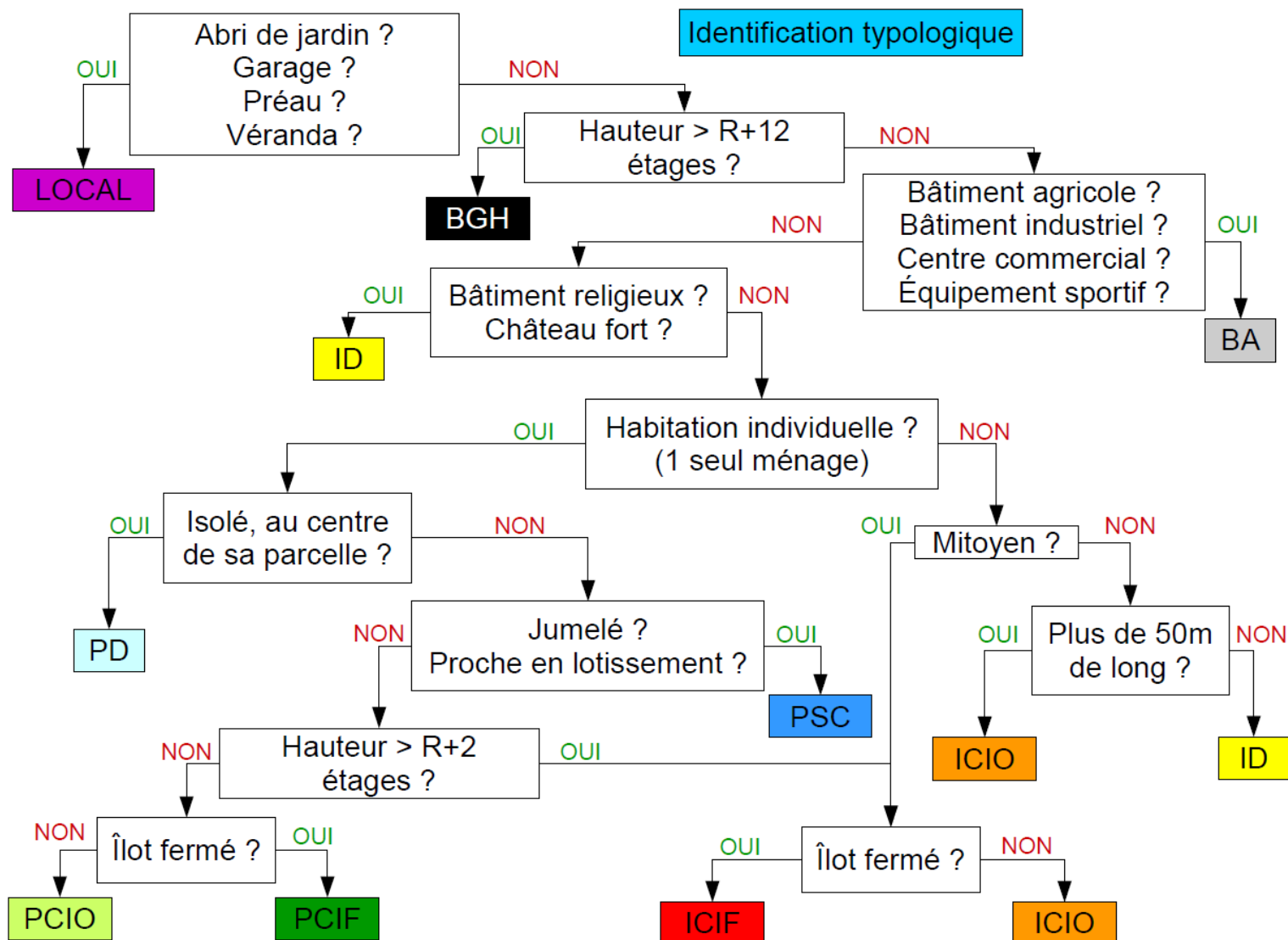


- Objectifs : identifier les grands types de tissu urbain
- Méthode
 - Apprentissage (randomforest)
 - 11 classes :
 - Bâtiment d'activité
 - Bâtiment de grande hauteur
 - Immeuble continu sur îlot ferme
 - Immeuble continu sur îlot ouvert
 - Immeuble discontinu
 - Local
 - Pavillon continu sur îlot ferme
 - Pavillon continu sur îlot ouvert
 - Pavillon discontinu
 - Pavillon semi-continu
 - Inconnu
 - Jeux de données de 18 000 bâtiments (manuel)
 - 1 classe associée aux indicateurs calculés aux 3 échelles géographiques

A legend box on the right side of the slide lists 11 urban tissue classes, each with a corresponding color swatch. The classes are: Batiment d'activite (grey), Batiment de grande hauteur (black), Immeuble continu sur ilot ferme (red), Immeuble continu sur ilot ouvert (orange), Immeuble discontinu (yellow), Local (pink), Pavillon continu sur ilot ferme (dark green), Pavillon continu sur ?lot ouvert (medium green), Pavillon discontinu (light green), Pavillon semi-continu (bright green), and Inconnu (white).

	Batiment d'activite
	Batiment de grande hauteur
	Immeuble continu sur ilot ferme
	Immeuble continu sur ilot ouvert
	Immeuble discontinu
	Local
	Pavillon continu sur ilot ferme
	Pavillon continu sur ?lot ouvert
	Pavillon discontinu
	Pavillon semi-continu
	Inconnu

Classification du tissu urbain



Source : A. Amossé, 2016

Protocole d'identification d'un bâtiment selon la typologie élaborée dans le cadre du projet MapUCE.

BA= bâtiment d'activité / BGH= bâtiment de grande hauteur / ICIF= immeuble continu sur îlot fermé /

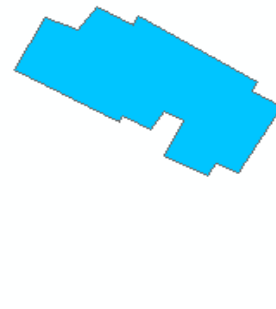
ICIO= immeuble continu sur îlot ouvert / ID= immeuble discontinu / LOCAL= local annexe / PCIF= pavillon continu sur îlot fermé /

PCIO= pavillon continu sur îlot ouvert / PD= pavillon discontinu / PSC= pavillon semi-continu

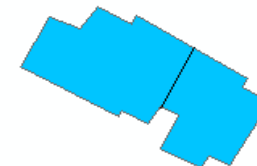
Classification du tissu urbain



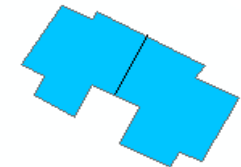
pd (bleu clair), psc (bleu foncé), local (violet)



Représentation
BDTopo



Représentation
attendue



→ Déphasage entre le terrain et les bâtiments de la BDTOPO

Source : A. Amossé, 2016

Langages et traitements

Les indicateurs sont formalisés en SQL Spatial.

Un exemple avec l'indicateur de facteur de forme (Form Factor)

$$FF_{build} = \frac{S_{build}}{L_{build}^2}$$

Where

- S_{build} is the building's area
- L_{build} is the building's length (perimeter)

```
DROP TABLE IF EXISTS BUILD_FORM_FACTOR ;
CREATE TABLE BUILD_FORM_FACTOR (PK integer primary key,
    FORM_FACTOR double)
AS SELECT PK , ST_AREA (THE_GEOM) / POWER (ST_LENGTH (
    THE_GEOM) ,2) AS FORM_FACTOR
FROM BATI ;
```

Langages et traitements

La classification est
réalisée avec un script R



```
1  ### Loading packages
2  library(randomForest)
3  library(RH2GIS)
4
5
6  ### Loading the model based on the morphological train data
7  TUFA=get(load(model_path))
8
9  ### Creating data to predict
10 dbSendQuery(con,"DROP TABLE IF EXISTS buildings_to_predict; CREATE TABLE buildings_to_predict
11
12
13 data_predict = dbGetQuery(con, "SELECT * FROM buildings_to_predict");
14
15 #Convert column names in lowercase as required by the R model
16 data_predict = setNames(data_predict, tolower(names(data_predict)))
17
18 ### Remove column identifiers
19 var_pred_com=data_predict[,-2]
20
21 ### Compute typologies
22 set.seed(7)
23 typologie=predict(TUFA,var_pred_com,type="class")
24 ### Matching typologies with building identifiers
25 tab_typo_bati=cbind.data.frame(data_predict[,1],typologie)
26 ##Export typologies
27 dbWriteTable(con, "TMP_TYPO_BUILDINGS_MAPUCE", tab_typo_bati, append=TRUE, row.names=FALSE)
28
29 ### Compute floor percent
30 pct_bati=ifelse(data_predict$u_floor!=0, data_predict$i_floor/data_predict$u_floor*100,0)
31
32 ### Compute % of typologies by USR
33 typo_USR=tapply(pct_bati,list(data_predict$pk_usr,typologie),sum)
34 typo_USR=ifelse(is.na(typo_USR),0.00,round(typo_USR,2))
35
36
37 ### Extract first max and second max typologies
38 #Extract first max
39 extract_typo_maj=function(ligne)
40 {
41   val_maj=which.max(ligne)
42   return(colnames(typo_USR)[val_maj])
43 }
44
45 #Extract second max
46 extract_typo_sec=function(ligne)
47 {
48   val_sec=order(ligne)[length(ligne)-1]
49   return(ifelse(ligne[val_sec]==0,"na",colnames(typo_USR)[val_sec]))
50 }
```

Source

Chaîne de traitements ...

... pour les bâtiments



Entry data



USR	
THE_GEOM	polygon
PK	int

BUILDINGS	
THE_GEOM	polygon
PK	int
HAUTEUR	double

1/ Preprocess

BUILD_USR_MATRIX	
PK_BUILD	int
PK_USR	int



Roads

BUILD_NEXT_ROAD	
THE_GEOM	polygon
PK	int
L_TOT	double
L_CVX	double
L_3M	double
L_RATIO	double
L_RATIO_CVX	double



2/ Building's level

BUILD_AREA_VOL	
PK	int
AREA	double
FLOOR_AREA	double
VOL	double

BUILD_FORM_FACTOR	
PK	int
FORM_FACTOR	double

BUILD_USR	
THE_GEOM	polygon
PK	int
HAUTEUR	double
PK_USR	int
NB_NIV	int

BUILD_CONTIGUITY	
PK	int
B_FLOOR_LONG	double
B_FLOOR_AREA	double
B_WALL_AREA	double
P_WALL_LONG	double
P_WALL_AREA	double
NB_NEIGHBOR	int
FREE_P_WALL_LONG	double
CONTIGUITY	double

BUILD_P_VOL_RATIO	
PK	int
THE_GEOM	polygon
P_VOL_RATIO	double

BUILD_DIST	
PK	int
MIN_DIST	double
MEAN_DIST	double
MAX_DIST	double
STD_DIST	double
PK_USR	int

BUILD_NUM_POINTS	
PK	int
NUM_POINTS	int

BUILD_COMPACTNESS	
PK	int
COMPACTNESS	double

BUILD_CONCAVITY	
PK	int
CONCAVITY	double

BUILD_COMPACITY	
PK	int
COMPACITY_R	double
COMPACITY_N	double

FRACTAL_DIMENSION	
PK	int
FRACTAL_DIM	double

BUILD_MAIN_DIR	
PK	int
MAIN_DIR_DEG	double

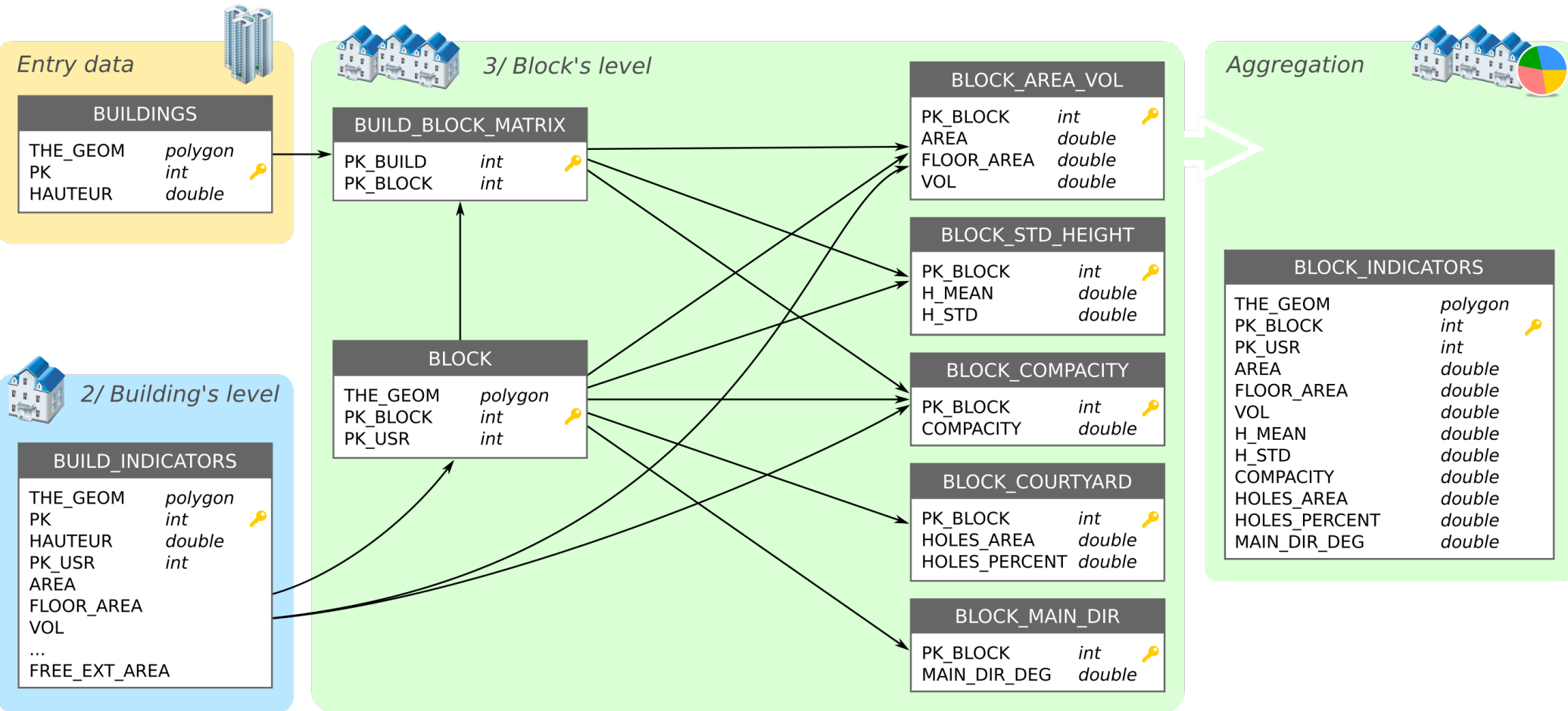
Aggregation

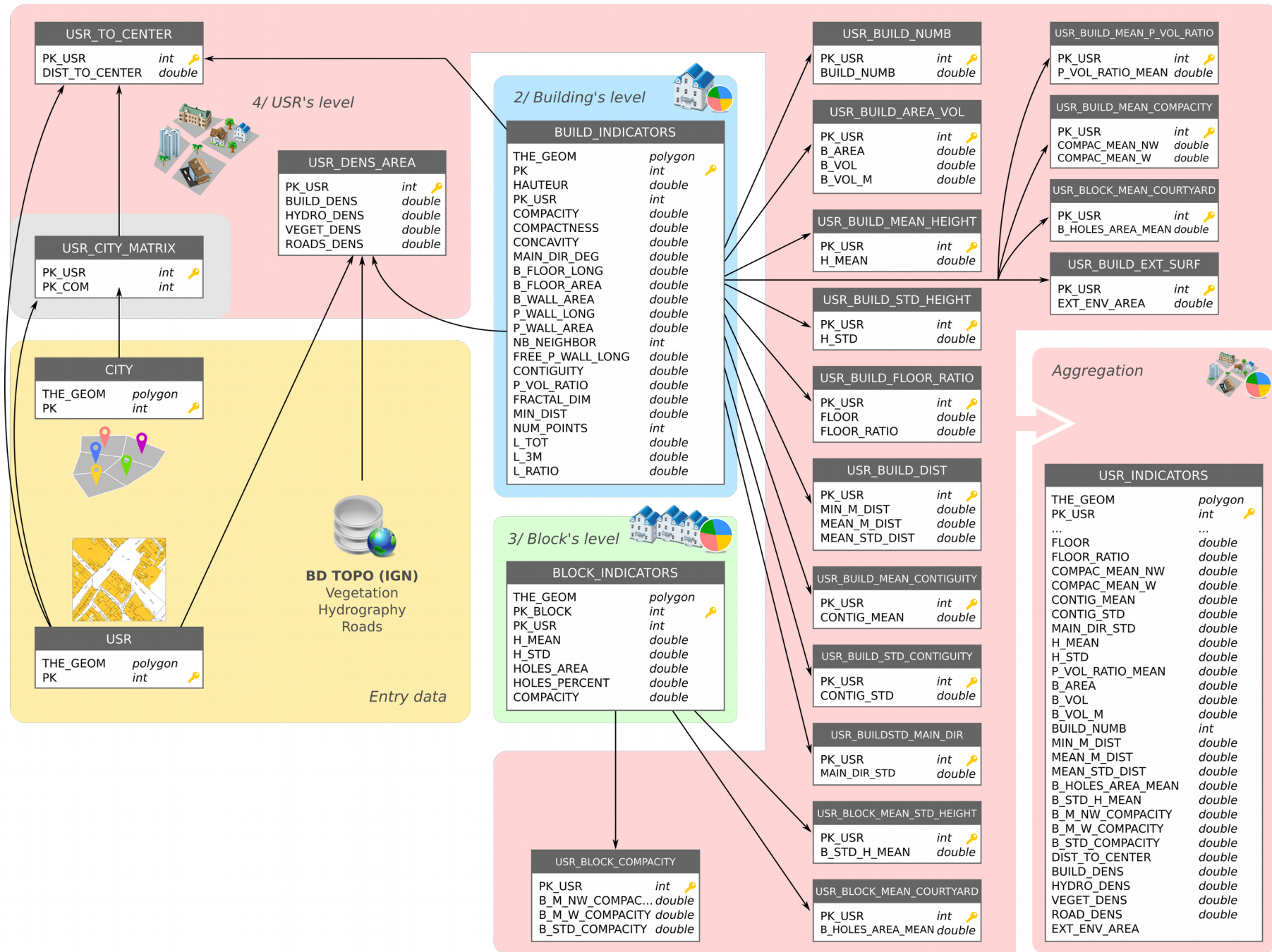


BUILD_INDICATORS	
THE_GEOM	polygon
PK	int
HAUTEUR	double
NB_NIV	int
PK_USR	int
AREA	double
FLOOR_AREA	double
VOL	double
COMPACITY_R	double
COMPACITY_N	double
COMPACTNESS	double
FORM_FACTOR	double
CONCAVITY	double
MAIN_DIR_DEG	double
B_FLOOR_LONG	double
B_FLOOR_AREA	double
B_WALL_AREA	double
P_WALL_LONG	double
P_WALL_AREA	double
NB_NEIGHBOR	int
FREE_P_WALL_LONG	double
CONTIGUITY	double
P_VOL_RATIO	double
FRACTAL_DIM	double
MIN_DIST	double
MEAN_DIST	double
MAX_DIST	double
STD_DIST	double
NUM_POINTS	int
L_TOT	double
L_CVX	double
L_3M	double
L_RATIO	double
L_RATIO_CVX	double

Chaîne de traitements ...

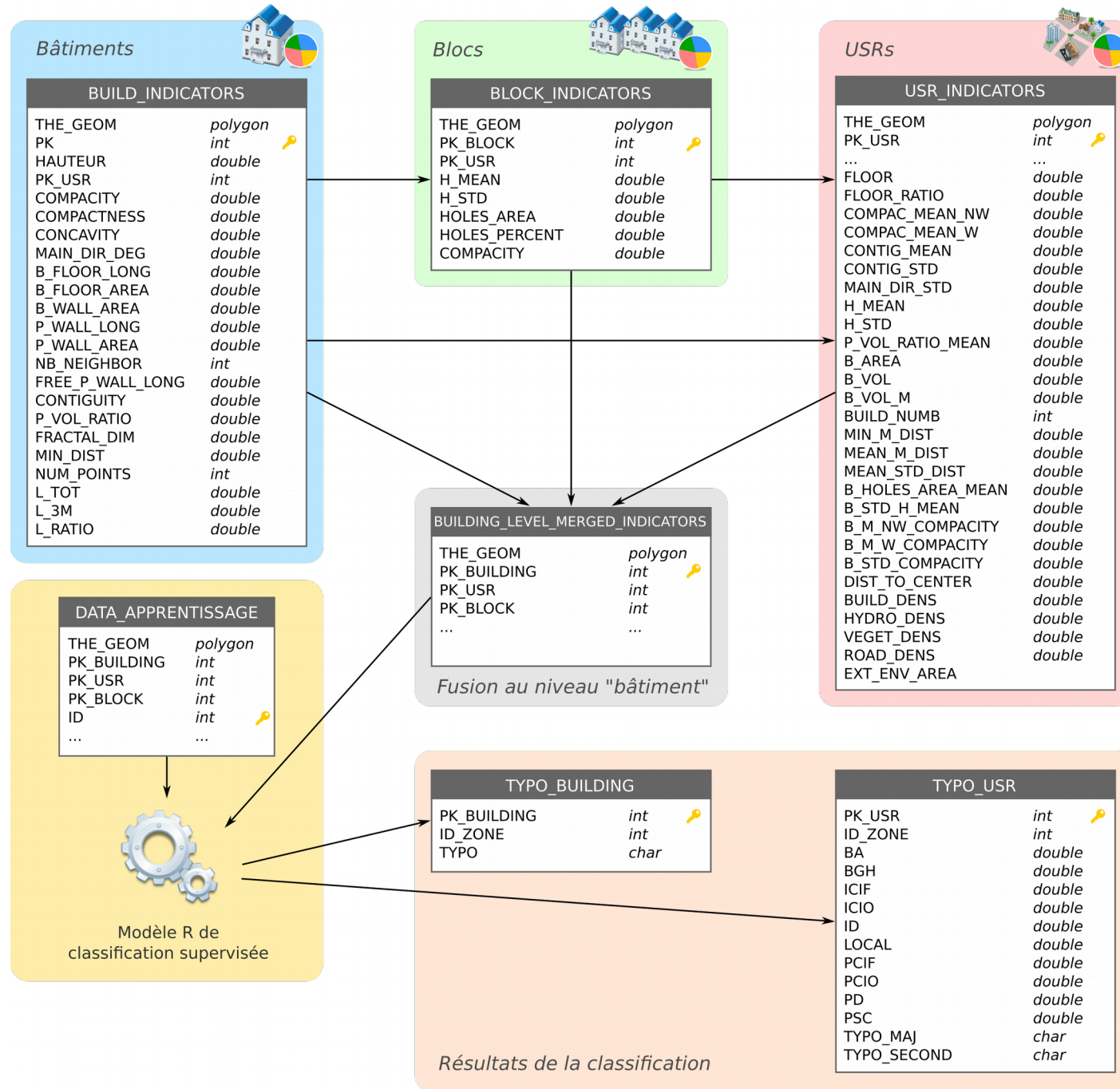
... pour les blocs





Chaîne de traitements ...

... pour la classification



Modèle de la base de données



Schéma "labsticc"

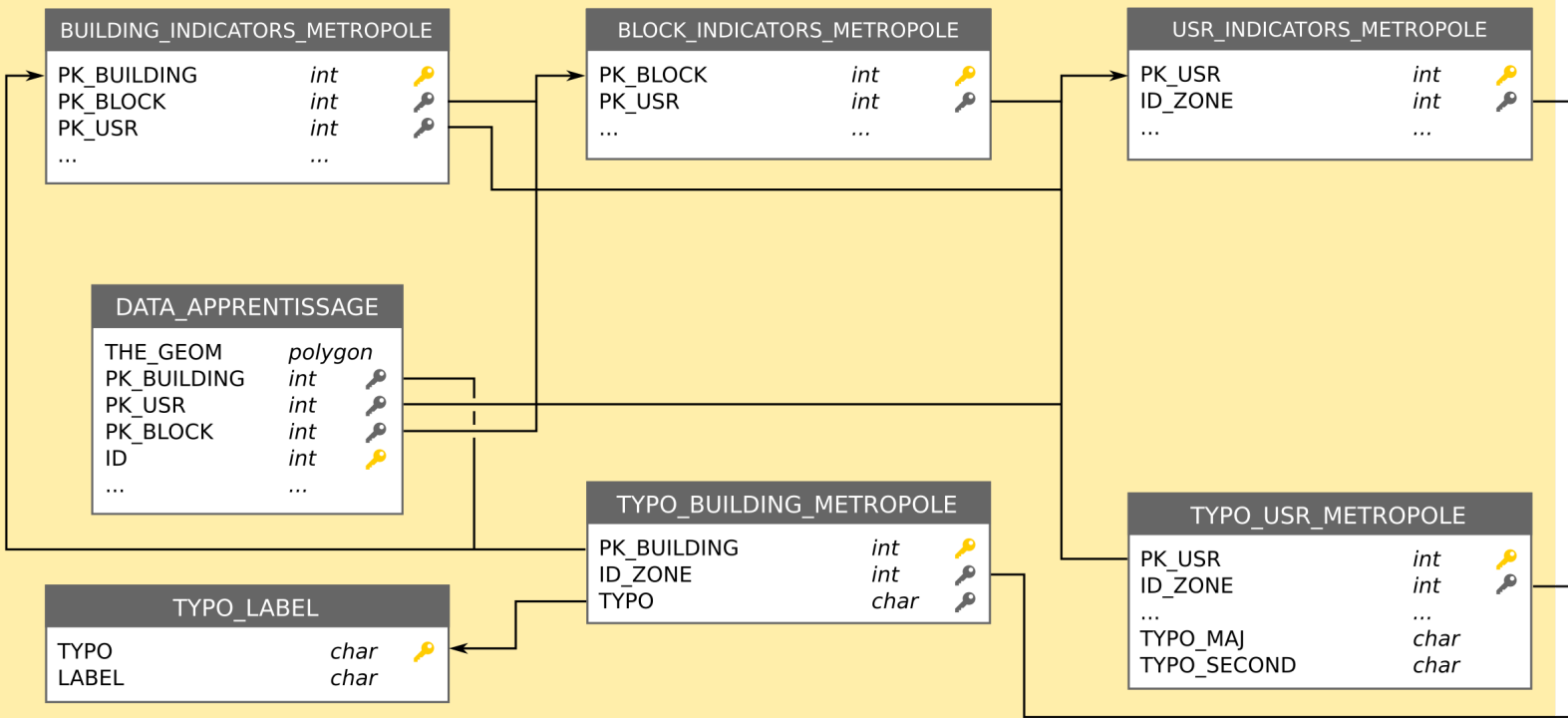
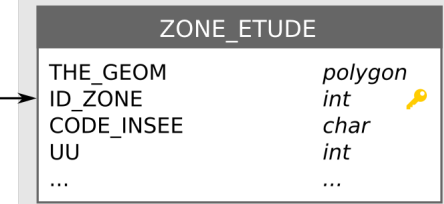


Schéma "liens"

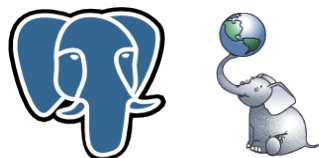


Architecture logicielle

Utilisation exclusive de solutions open-source



Spatial DB (store)



Spatial DB (process)

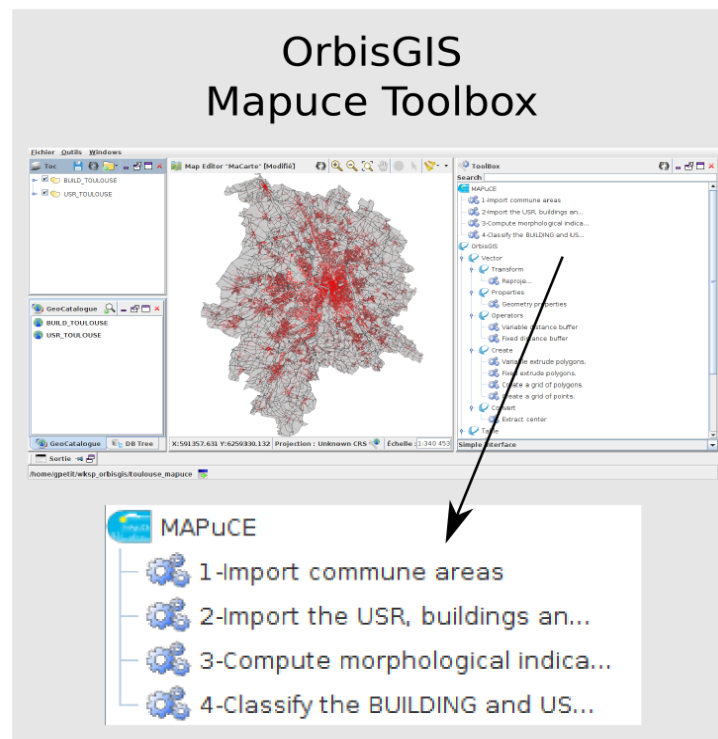
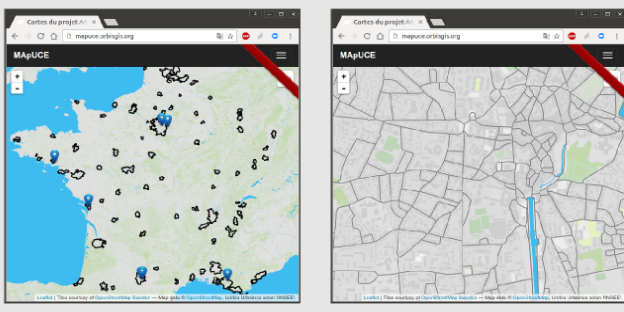


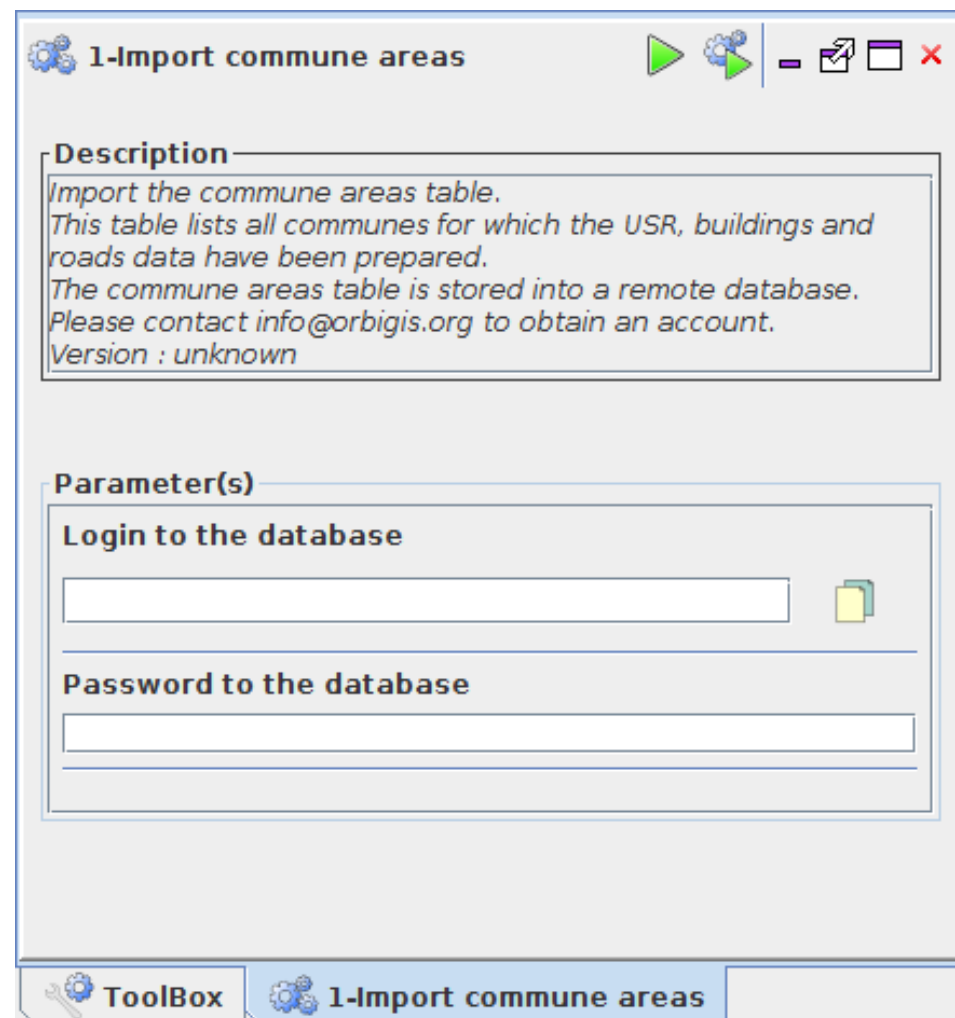
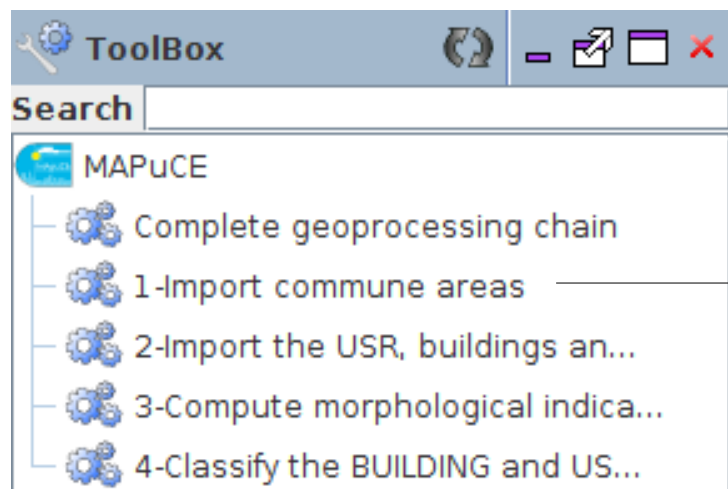
WPS Service



OrbisGIS
Mapuce Toolbox

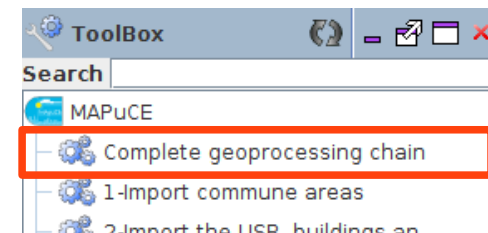
Web carto
mapuce.orbisgis.org



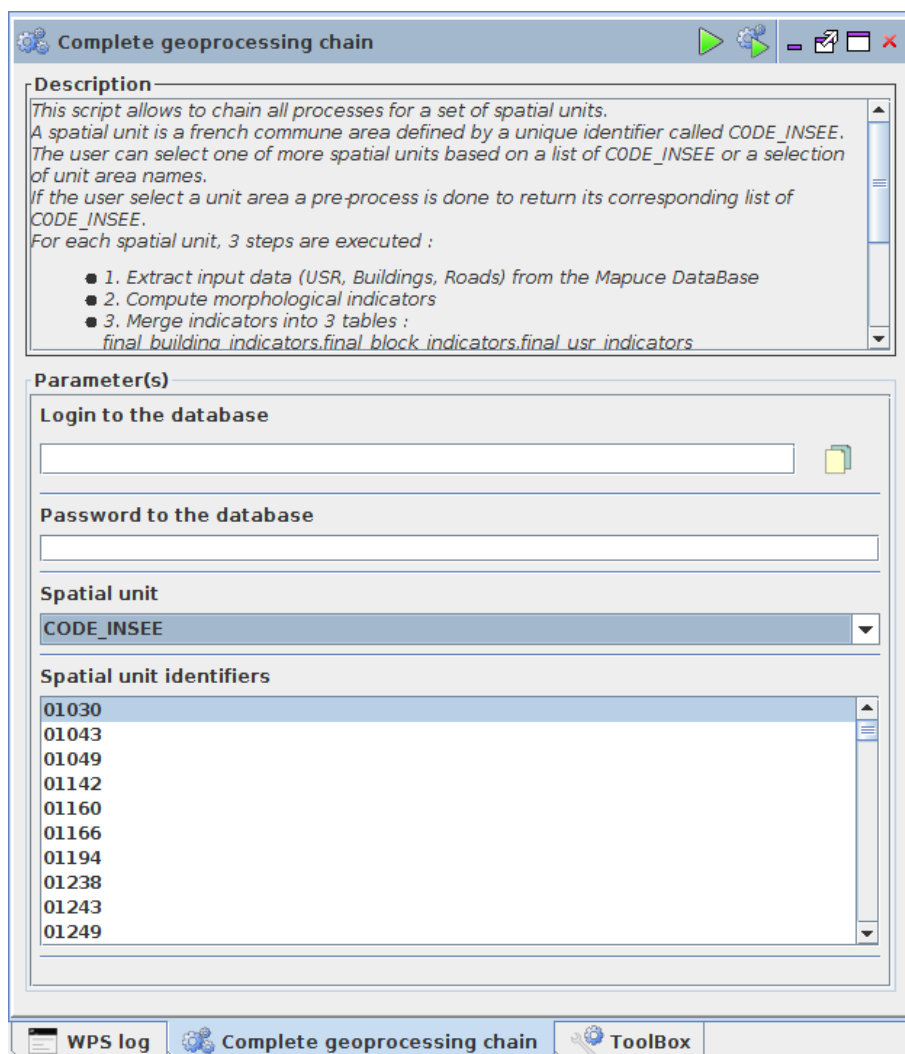


- 5 scripts présents dans la Toolbox MAPUCE
- « *1- Import commune areas* » doit être exécuté en 1^{er}. C'est lui qui va importer la liste des communes disponibles au traitement
- Si on souhaite faire du pas à pas, il faut utiliser les scripts 2, 3 et 4
- Sinon, le script « *Complet geoprocessing chain* » permet de tout faire en une seule fois (excepté le script 1)

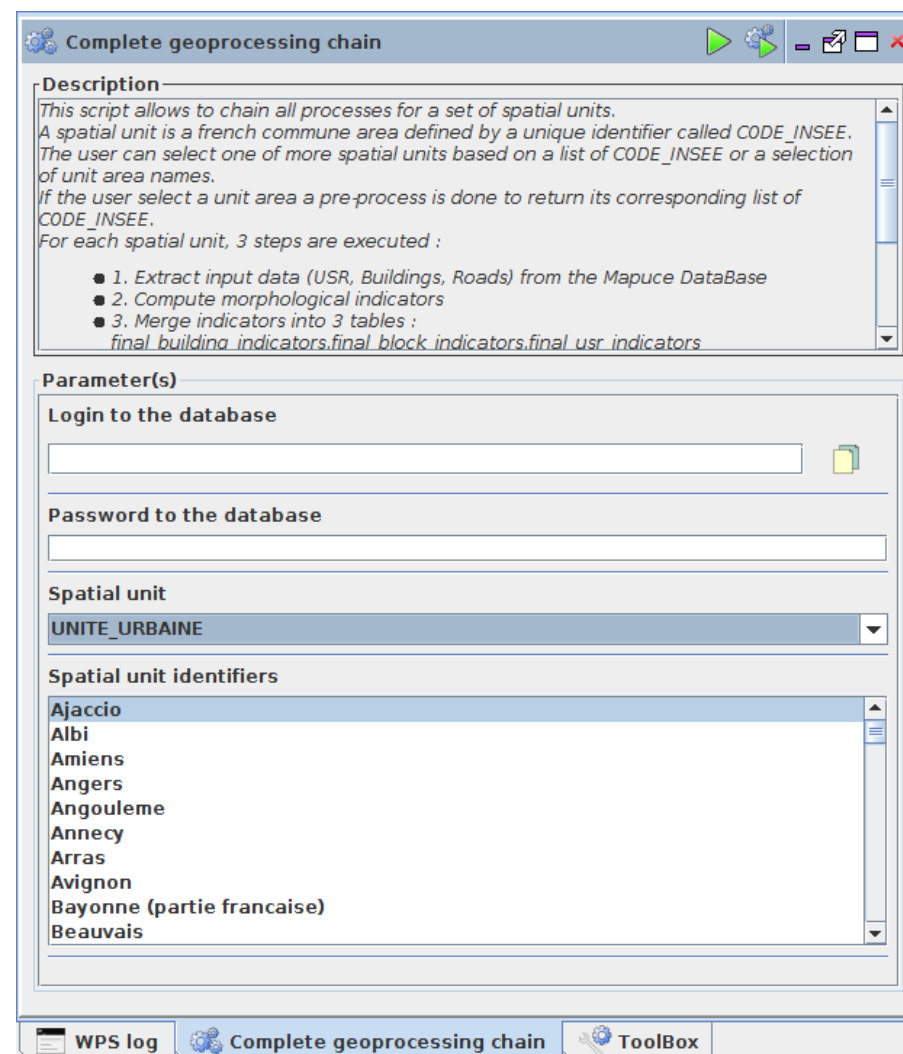
Interface de pilotage dans OrbisGIS



Traitement sur la base de une ou plusieurs commune (code insee)

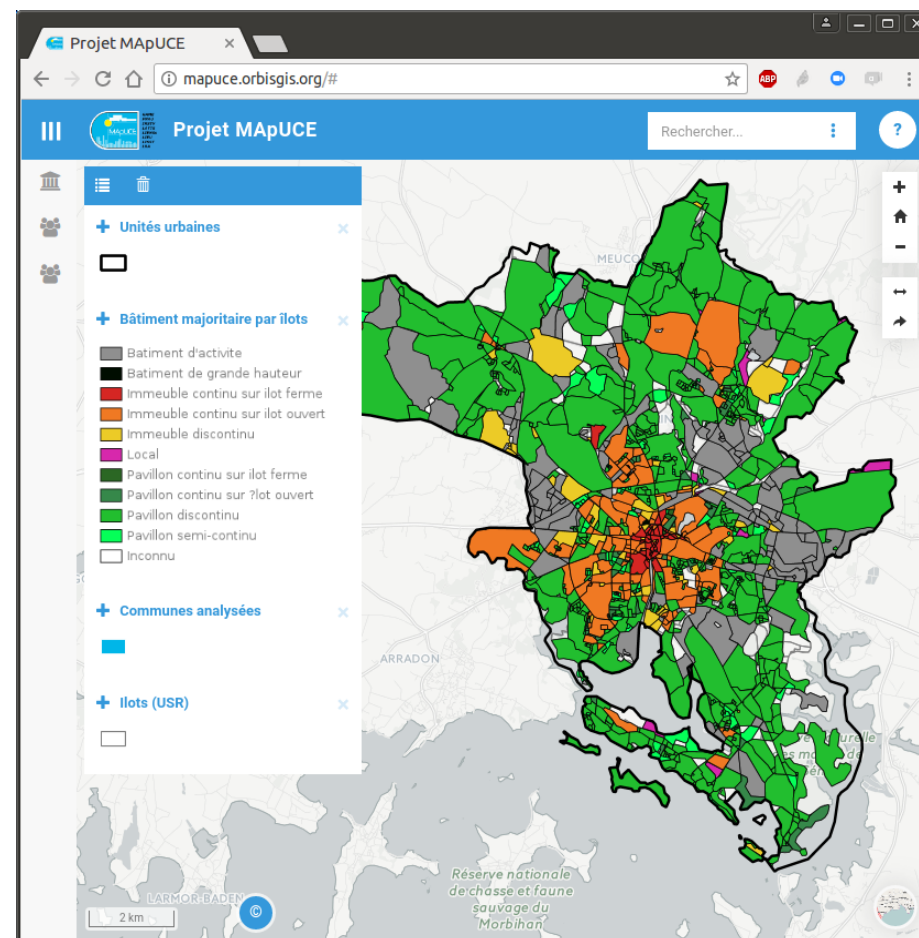
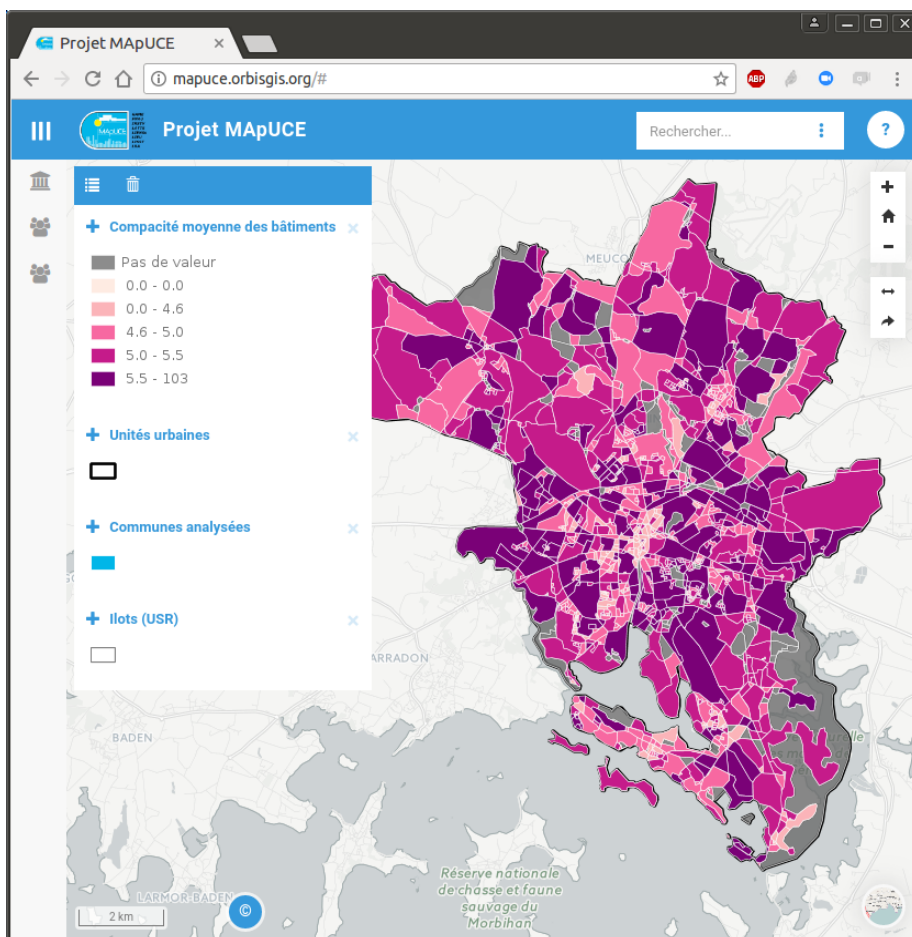
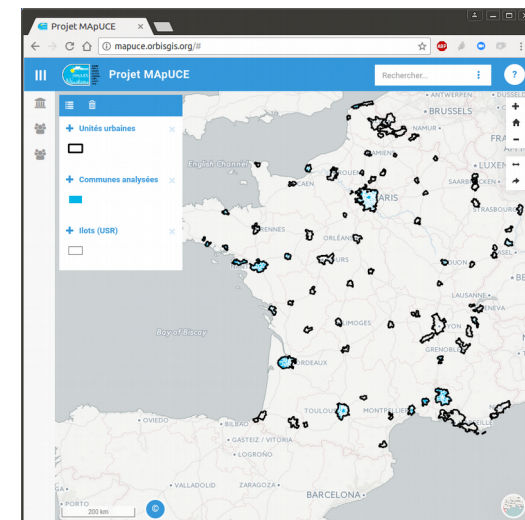


Traitement sur la base de une ou plusieurs unité urbaine



Interface de publication des résultats

- Portail cartographique ouvert → <http://mapuce.orbisgis.org/>
- Couches : Unités urbaines, Zones traitées, USR, Indicateurs par USR, classification des USR



Merci pour votre attention

Place aux questions

Télécharger la présentation



Ou bien : <https://tinyurl.com/khezvde>

Liens

- Portail cartographique Mapuce → <http://mapuce.orbisgis.org/>
- OrbisGIS : <http://orbisgis.org/>
- H2GIS : <http://www.h2gis.org/>
- Dépôt codes sources : <https://github.com/orbisgis>