Cellular Automata and Urban Development Simulation: Transition Rules Creation Process Based on Statistical Analysis.
A methodological approach

Valentine Judge [1,2] , Jean-Philippe Antoni [1]

[1] ThéMA (UMR 6049 CNRS - Franche-Comte University)
&
[2] LISER

Conférence ThéoQuant 21-05-2015
Supported by the National Research Fund, Luxembourg (Project INTER/CNRS/12/02)

Sommaire

Land use simulation and cellular automata

Cellular Automata have been widely used for modeling land use change in geography *(i.e. Tobler, 1970; Couclelis, 1985; Phipps, 1989; White, 1997; Torrens, 2002, etc.)* Main attractive features: simplicity, interaction between neighboring land units, emergence.

Cellular Automata : *"System of spatially located and interconnected automata." (Beneson & Torrens, 2004)*

Automata : *"Processing mechanism with characteristics that changed over time based on its internal characteristics, rules and external inputs." (Beneson & Torrens, 2004)*

$$A \sim (S, T, R)$$

A : Automata
S : Sate
T : Transition rules
R : Neighborhood

Tool: LucSim *(developed from scratch by: Vuidel & Antoni)*

General objective

Urban development study using cellular automata: A transition rules calibration approach

Hypothesis:

- The transition of a cell from one land use type to another is driven by its initial land use and by the land use of cells from its neighborhood.

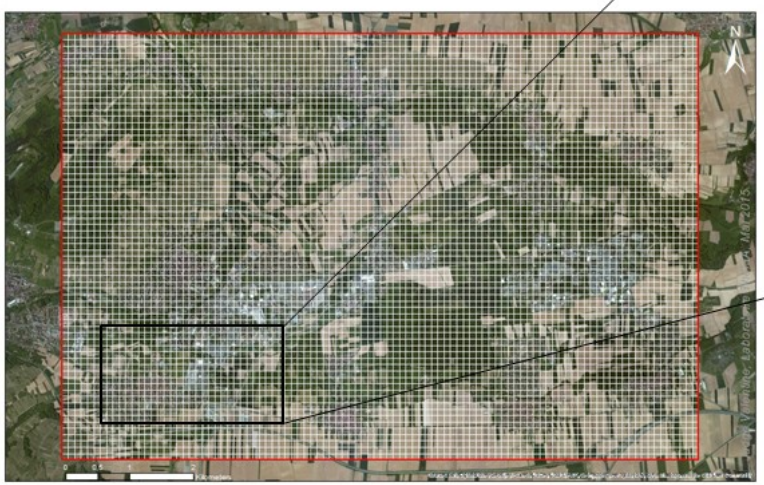
Questions:

- How can transition be defined?
- How can transition be written as rules?

Study area

Extend of the study area:
6700 Cells.
Cell dimension:
1 Ha (100x100m cell)

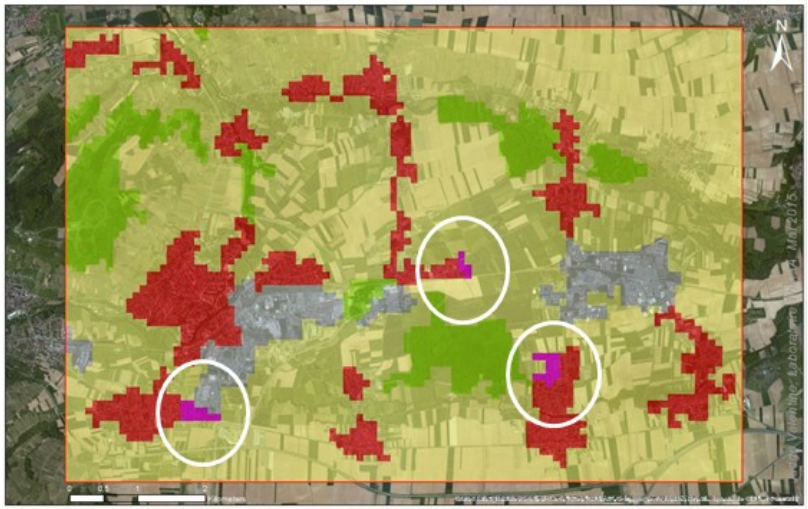
Extend of the study area:
67 km^2 (6700 Ha).



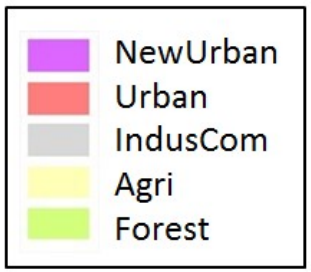
Cellular approach

Land use cover (1990)

Urban Changes (1990-2006)



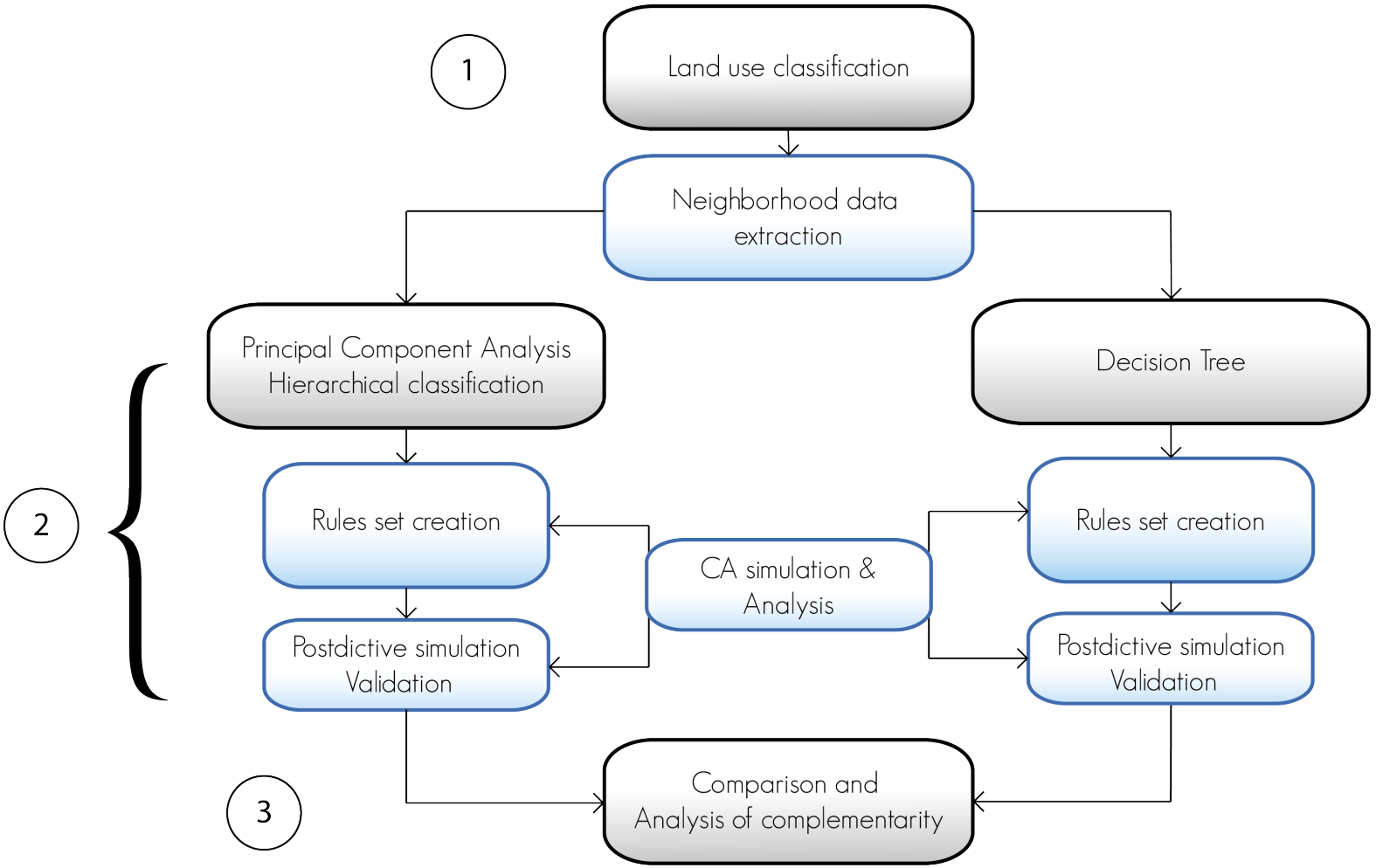
Re-classified CorineLandCover



		2006					1990-2006
		Urban	IndusCom	Agri	Forest	Loss	Change
1990	Urban	810	0	1	0	-1	29
	IndusCom	12	309	1	0	-13	46
	Agri	18	59	4780	23	-100	-50
	Forest	0	0	48	572	-48	-25
	Gain	30	59	50	23		

Sommaire

Methodological scheme



Rules definition proposal

Assumption CA based: Land use of cells settled in the neighborhood can explain on a potential urban transition of a cell.

$$A \sim (S, T, R)$$

Objective: Create a set or transition rules based on the assumption.

$$S_{i(t)} \rightarrow S_{f(t+1)} : R_{Sa}\{N_{Sa}(D_a; n_a), N_{Sb}(D_b; n_b), N_{Sc}(D_c; n_c), \dots\}$$

Questions :

- Which states?
- Which distances?
- How many cells?

$S_{i(t)}$: Initial sate
$S_{f(t+1)}$: Final state
R_{Sa}: Neighborhood
N_{Sa}: Neighbor characterized by State "a"
n_a: Number of State "a" neighbor
D_a: Distance where is found State "a" neighbor

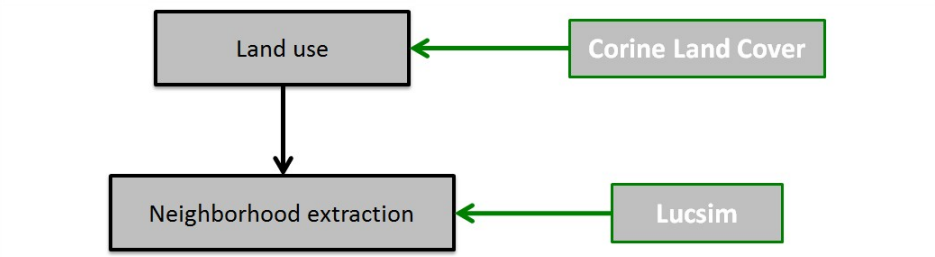
Proposition : Create transition rules based on statistical analysis of the neighborhood of cells that carried out specific transition.

LUCSIM LAND USE CELLULAR AUTOMATA SIMULATION

Dataset

Datasets : Ernolsheim-Bruche area

- Tables {Individual, initial state, final state, a 5 pixels neighborhood}.
- Data: 1990-2006.
- Variable shows the number of cells of a land use type at a specific distance.



IdFixed	State_CLC_00	State_CLC_90	Urban1px	IndusCom1px	Agri1px	Forest1px	Water1px	Urban2px
0	4	4	0	0	0	3	0	0
1	4	4	0	0	0	5	0	0
2	4	4	0	0	0	5	0	0
3	4	4	0	0	0	5	0	0
4	4	4	0	0	0	5	0	0
5	4	4	0	0	0	5	0	0
6	4	4	0	0	0	5	0	0
7	4	4	0	0	0	5	0	0
8	4	4	0	0	0	5	0	0
9	4	4	0	0	0	5	0	0
10	4	4	0	0	0	5	0	0
11	4	4	0	0	0	5	0	0
12	4	4	0	0	0	5	0	0
13	4	4	0	0	0	5	0	0

Principal component analysis and hierarchical classification

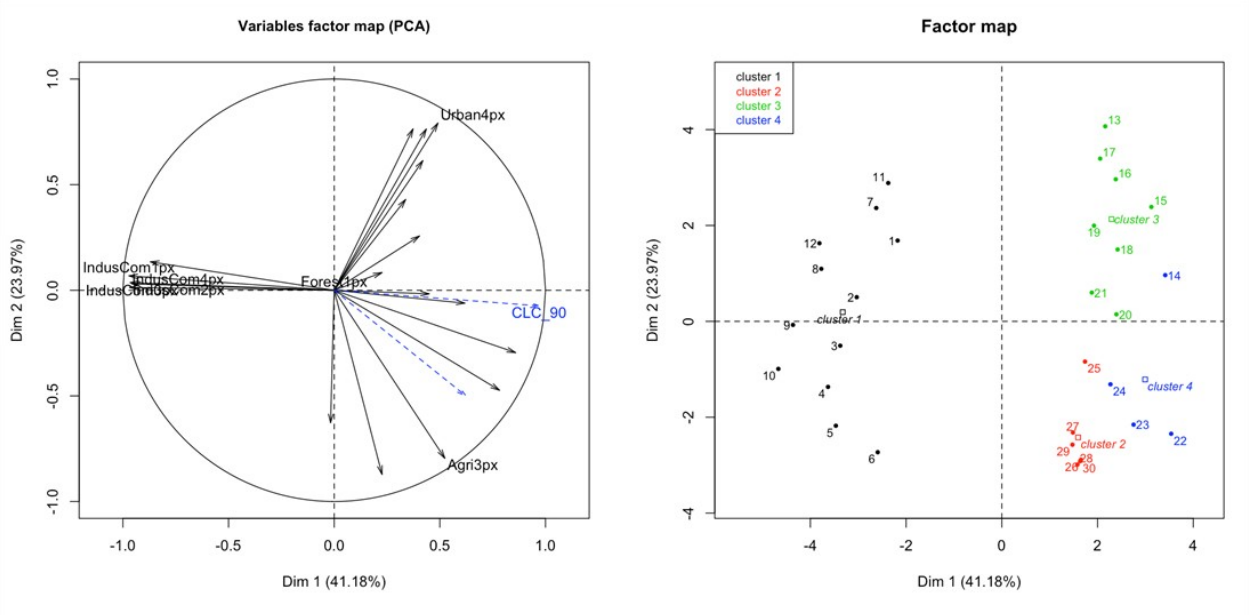
Objectives :

- Target variables summarizing the most of information contained in the dataset being explored.
- Create cluster of individuals (cells) considering variables defining them.
- Analyze the neighborhood of individuals from the same cluster in order to get land use influence threshold (number of cells within a specific distance).

Dataset :

- Focused on cells that carried out a transition from a λ state to an urban state (30 individuals).

Set of rules creation



Method :

- Variable that summarize the most of the dataset information are used as a base for the rules set (i.e. IndusCom4px, Urban4px, Agri4px, Forest5px).
- A minimum and maximum threshold is set from the analysis of individuals or clusters from the hierarchical classification(i.e. $4 <$ IndusCom4px <7, $0 <$ Urban4px < 6, etc.).

Decision tree

Objective :

- Sequential partition of the dataset using a top down approach. The aim is to target the variable that best split the different cluster of data left at each step.

Dataset :

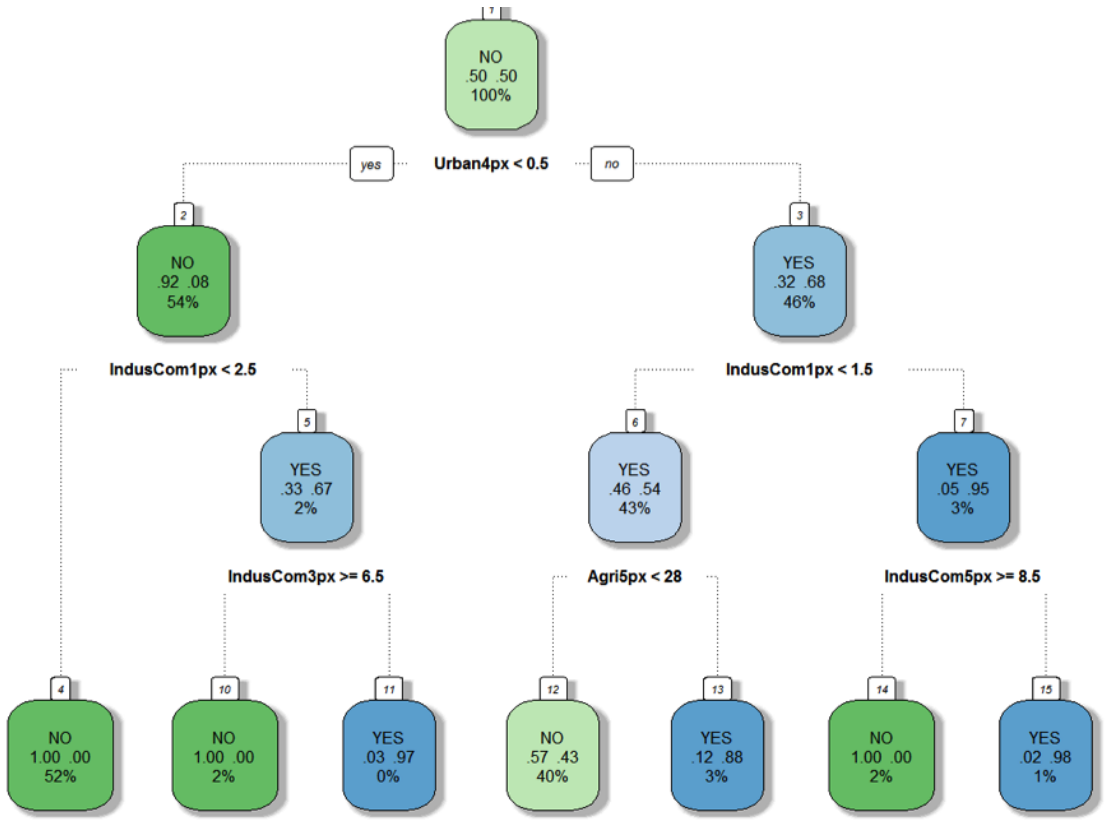
- Every cells regardless of their transition.
- Adding a binary categorical variable to characterize an urban transition or not.

Set of rules creation

Decision tree

Method:

- Rules can be extracted directly from the tree analysis. Each "leave" corresponding to the transition into urban land use is considered and transcribe as rules.



LAND USE CELLULAR AUTOMATA SIMULATION

Rules set analysis

Why do we analyse rules sets produced?

- Are the same variables extracted with both methods?
- If yes: Are the threshold similar?
- If not: what is the extend of the dissimilarity?

Principal Component Analysis based rules set

$$Agri\ or\ IndusCom \rightarrow Urban :$$

$$0 \leq Urban4px \leq 6\ And\ 4 \leq IndusCom\ 4px \leq 7\ And\ 3 \leq Agri3px \leq 11\ And\ 1 \leq Forest5px \leq 4$$

Decisions tree based rules set

$$Agri\ or\ IndusCom \rightarrow Urban :$$

$$Urban4px > 0\ And\ IndusCom\ 1px \geq 1\ And\ Induscom5px < 9\ And\ Agri5px < 28\ And\ Forest5px > 0$$

Sommaire

Validation

Objective :

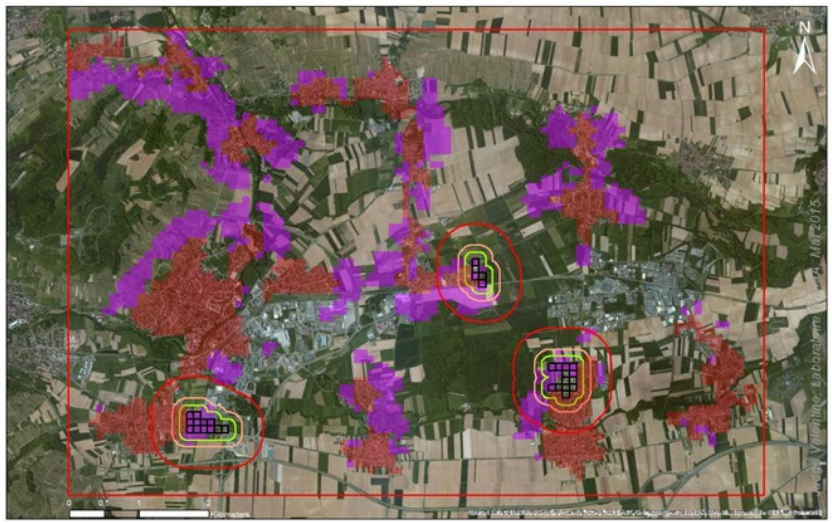
- Evaluation of the results quality comparing with data from 2006.
- Compare result produced by the PCA/HC method and the decision tree method.

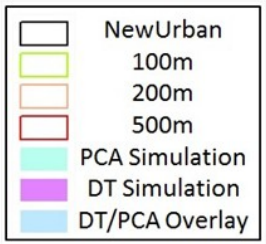
Validation method: GIS based

- Pixel to pixel validation
- buffer 100m, 200m, 500m

Rules sets simulation and analysis

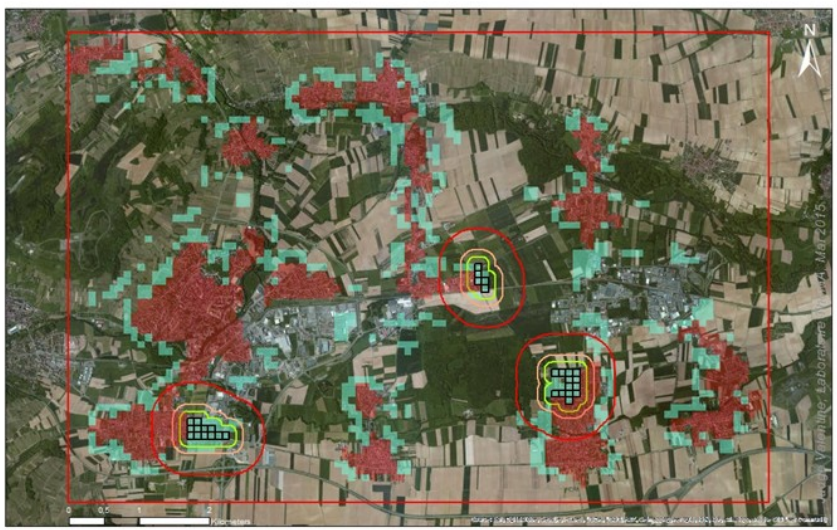
Case: Simulation matching the most real NewUrban cells.





PCA Simulation

DT Simulation

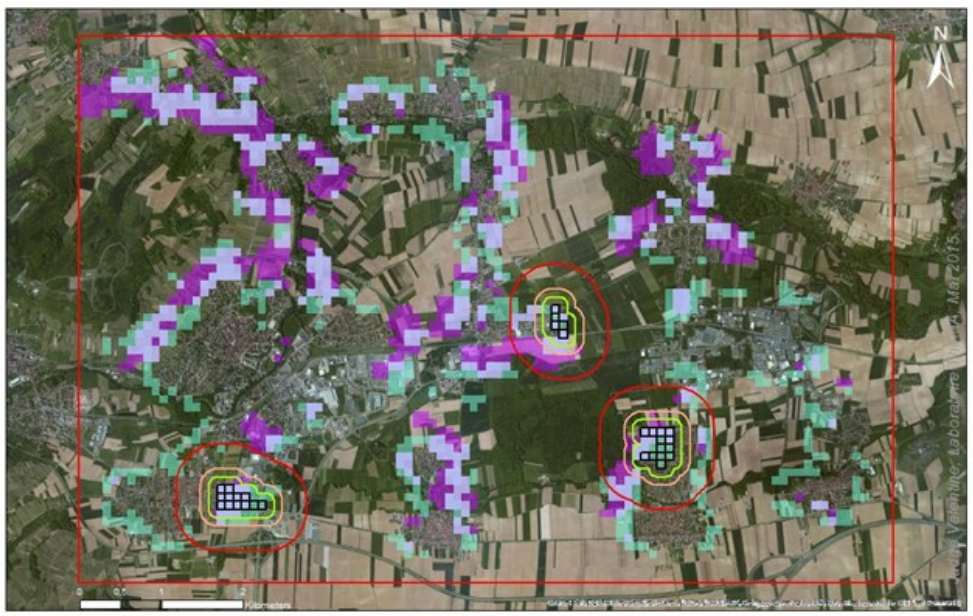


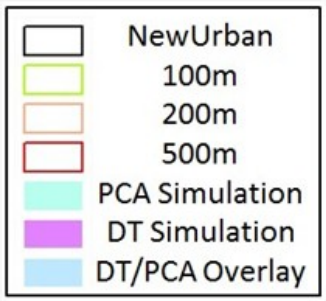
	PCA Simulation	DT Simulation	Similarity
Unites	Number of cell		Percentage
NewUrban	604	746	
Px vs Px	30	18	49
100m	19	13	72
200m	13	45	82
500m	38	50	95

LAND USE CELLULAR AUTOMATA SIMULATION

Rules sets simulation and analysis

Case: Simulation matching the most real NewUrban cells.





	PCA Simulation	DT Simulation	Similarity
Unites	Number of cell		Percentage
NewUrban	604	746	
Px vs Px	30	18	49
100m	19	13	72
200m	13	45	82
500m	38	50	95

Comments

Rules sets analysis:

- Some identical variables pop up from both analysis (i.e.Urban4px, Forest5px)
- Not all variables that pop up are identical (distances can differ i.e. IndusCom5px vs IndusCom4px)

Simulation analysis:

- PCA simulation is more efficient (localize 100% of the real 2006 NewUrban cells)
- DT simulation is more compact than PCA simulation
- Both simulation seems to spatially complement on another having a 72% similarity when a 100m accuracy is taken into account.

General comment: Both set of rules are partly satisfying or interesting to explore but miss accuracy.

Sommaire

Conclusion

Urban development study using cellular automata: A transition rules calibration approach

First results:

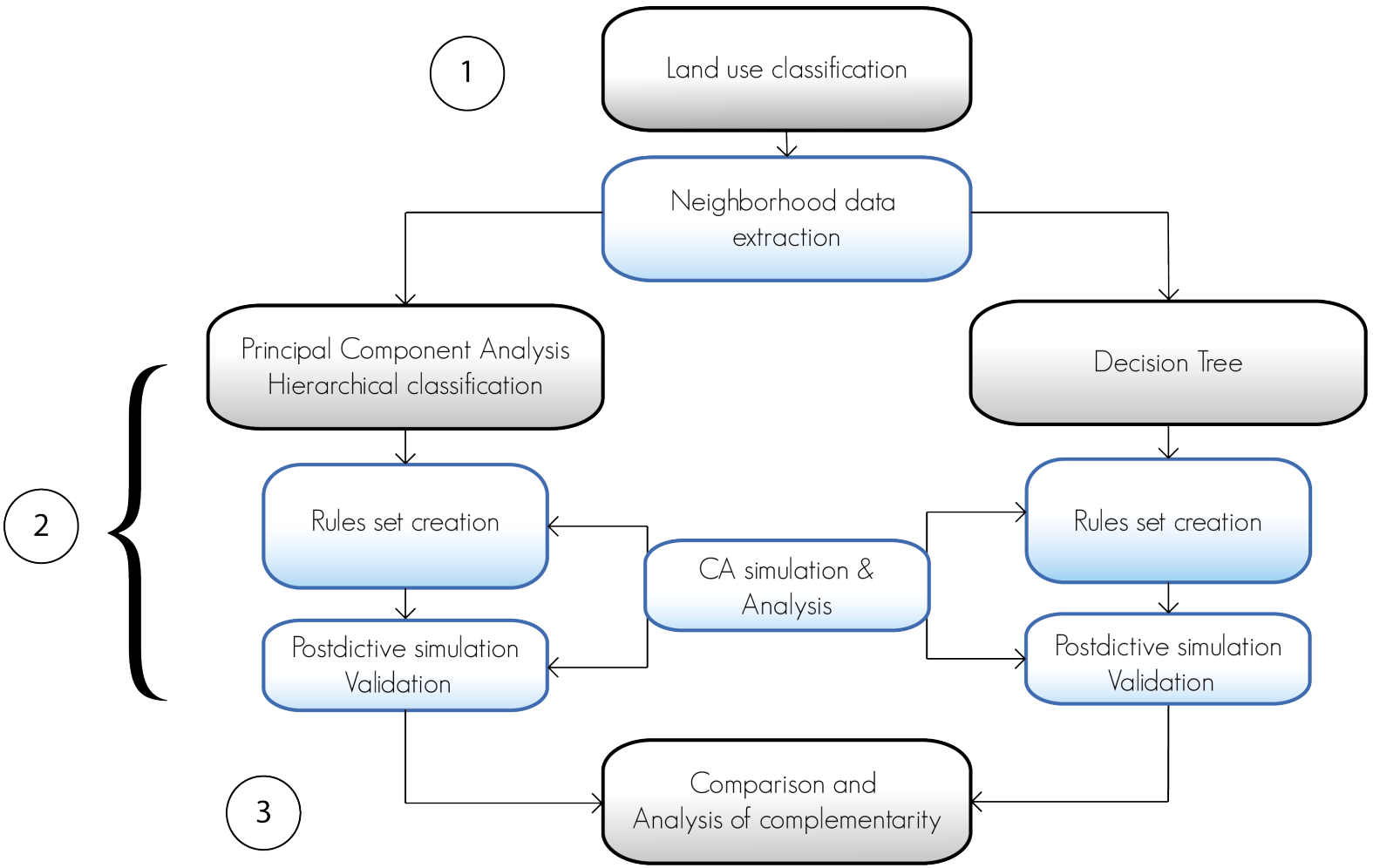
- Rules can be written based on the analysis of the neighborhood.
- Several solution can be use, two have been tested:
  - Principal component analysis with hierarchical classification
  - Decision trees
- Both set of rules are partly satisfying according to the validation but miss accuracy

Perspective

First results : perspectives ?

- Improve rules sets according to both methodologies
- Improve the accuracy of the rules (i.e. add new variables ? accessibility ?)
- Apply the methodology developed here on a larger area (reduce the bias from the data)
- Explore results more deeply from a thematic point of view

Methodological scheme



LAND USE CELLULAR AUTOMATA SIMULATION

Cellular Automata and Urban Development Simulation: Transition Rules Creation Process Based on Statistical Analysis.
A methodological approach

Valentine Judge [1,2] , Jean-Philippe Antoni [1]

[1] ThéMA (UMR 6049 CNRS - Franche-Comte University)
&
[2] LISER

Conférence ThéoQuant 21-05-2015
Supported by the National Research Fund, Luxembourg (Project INTER/CNRS/12/02)