

# Neighbourhood coefficients of cellular automata for research on land use changes with map algebra<sup>1</sup>

## Abstract

The dependency of land use and the neighbouring land cover patches is related to the existing neighbourhood or the predominant land use types in a given region, which stimulates further land changes of the lot being observed. Land use changes can be considered a complex and (to an extent) random process. The complexity of interactions means that the conventional deductive models are constrained, particularly in the case of the spatio-temporal phenomena, and implies the application of cellular automata to spatial research. Cellular automata are similar to the map algebra models in GIS. The concept of using neighbourhood coefficients in the analysis of land use changes is based on the combination of map algebra with two-dimensional cellular automata.

The aims of this research included formulating the theoretical structure of neighbourhood coefficients, analysing their operationalization and testing their practical application. The verification of research and procedures included maps and statistics of simulations of land use changes in Poland.

## Keywords

Spatio-temporal modelling • land use • cellular automata • map algebra • neighbourhood coefficients

© University of Warsaw – Faculty of Geography and Regional Studies

## Piotr A. Werner

Faculty of Geography and Regional Studies,  
University of Warsaw  
e-mail: peter@uw.edu.pl

Received: 13 February 2012

Accepted: 22 June 2012

## Introduction

Land use is defined as the spatial distribution of forms of land cover patches that are utilized or unutilized by humans within the framework of spatial and mutual relationships. The term refers to a given terrain's functional character and is identified with a socio-economic description of the surface (Ciolkosz & Bielecka, 2005, after Stamp, 1960, Kostrowicki, 1959, Kozubek, 2002). The observed dependency of land use and neighbouring land cover patches is discussed in many geographical and spatial economy publications i.e. the economic utilization of an observed lot (or patch) has significantly fewer implications for its future utilisation than the existing land use in its neighbourhood. In either case, factors stimulating further land use changes include the existing neighbouring land use or the predominant land use type in a given region (Hagoort, 2006).

## Models of land use changes

The analysis of land use changes involves a wide set of tools and methods including statistical and econometric models, spatial interaction models, models of (non)linear programming and economic models - utility maximization models or multi-criteria and multipurpose decision support systems. The interdisciplinary character land use changes forces researchers into the application of integrated, simulation-based tools. Further analytic methods include input-output models (Hagoort, 2006) as well as fuzzy logic (Bilozor, 2003).

The advancement of geographic information systems (GIS) has facilitated the simulation and geovisualisation of results. The majority of analyses confronted two categories of factors that influenced the observed land use changes and which, at their core, involved an evaluation of spatial changes that resulted from the impact of neighbouring and consecutive forms of land use in a given area. The types of land use classes differed depending

<sup>1</sup> The first part of the paper involves a summary of the previous results of the author's research that was needed for the further extension of the method (Werner 2009) and supplemented by calculated Kappa coefficients.

on the aim, the scale and the area of an individual geographical study (Werner, 2009).

Yet another aspect of research into land use changes involved spatial externalities (Hagoort, 2006), defined as the unexpected, indirect results of leverage of human activity that cause changes of costs or values for others - without visible operation (Mishan, 1969). Spatial externalities of land use changes are usually disclosed through intensity (rate) and spatial range. Spatial externalities can be classified as positive or negative. The intensity of spatial externalities related to a given spatial land use class (form) and the probability of change of the land use class of a given spatial unit are positively correlated (Tobler, 1979). The aggregate impact of the land cover's spatial externalities (net externalities) can be analysed using multi-variables and spatial or conditional spatial models of land use changes.

The majority of land use models are geography-based. The recognition of the main changes in land use enabled a set of model changes in land use to be distinguished (Hagoort, 2006, Tobler, 1979) including: the Independent, the Dependent, the Historical, the Multivariate, the Geographical and the Future Geographical models.

In everyday practice, research and spatial planning often involve applying mixed forms of the aforementioned models which are also construed using Spatial Decision Support Systems, accounting for transportation facilities, population, employment etc.

As geographers' attention has, in most cases, centred on regions of urbanisation, the majority of research projects have been devoted to studying urban land use changes. The dissimilarity of their conclusions could be attributed to the individual and diverse character of the range of the studied geographical locations but the findings also identified a number of repeatable processes. Studies of the urban area determined that individual elements of landscapes change at various speeds (Hagoort 2006, after Whitehand, 1967) and some of them are more susceptible to change. It was further established that the rate of land use changes differs depending on relative distance from the city centre (Hagoort, 2006, after Wilder, 1985). The various temporary profiles of urban land use changes are nonetheless much longer than the average length of a human life and longer than a spatial planning perspective (Hagoort, 2006 after Wegener, 1995). On the other hand, main streets and their courses remain the most durable landscape elements (Hagoort, 2006, after Whitehand, 1967) and specific types of land use forms are change-resistant, while sometimes the successions are not observed at all (Hagoort, 2006, after Wilder, 1985).

The geographical model is a model of spatial changes. It is assumed that the future land use of a particular place depends on the present land use of the neighbouring land and can be described graphically (Tobler, 1979, Fig. 1).

The analysis is based on a one-to-one relation between each unique configuration of Moore neighbourhood land use classes ( $3 \times 3$  cells) and the calculated Neighbourhood Coefficient which enables the obtaining of an intermediate map of neighbourhoods' coefficients. Thereafter, it facilitates the performance of spatial

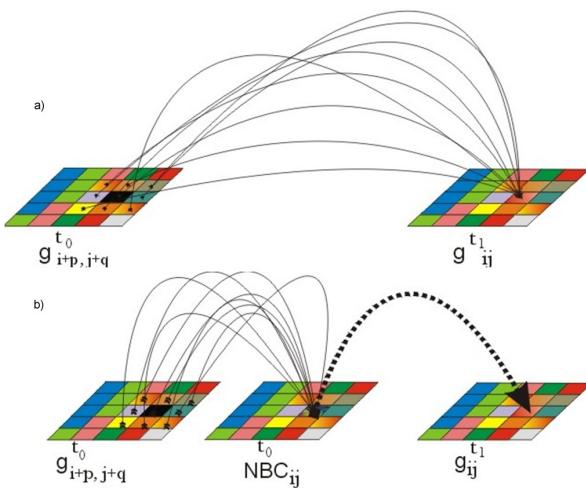


Fig 1. a) Geographical Model of land use changes (Tobler 1979)  
b) Proposed modification of the Geographical Model

comparisons of the new intermediate map of NBCs and the destination map of land use classes and most importantly, it enables analysis of the successions of neighbourhoods (Werner, 2009).

A similar method can be defined using the mixed geographic-historical model.

#### Cellular automata and map algebra (CA & MA)

The changes of land use can be treated as a complex and (to a certain extent) random process. The complexity of interactions implies that conventional deductive models are constrained, especially in the case of the spatio-temporal phenomena (Wu, 1999). Probabilistic models are also effectively applied and they make use of the concepts of random development of reality, the notions of statistical events, probability, random variables and of empirical distributions of probabilities (Ratajczak, 1999). According to Fotheringham (et al. 2000), one of the purposes of spatial and visual analysis of maps is the indirect understanding of phenomena by analysing spatial patterns and inferring the spatio-temporal processes. Wolfram (2002, p. 299) defined three possible mechanisms accounting for the complexity and randomness: '(...) there is random input from the environment at every step (...) there is random input only in initial conditions and in the third case, there is effectively no random input at all (...). The first mechanism is captured in stochastic models, the second has its roots in chaos theory and the third mechanism is formed on the basis of the behaviour of simple programs (ibid). This last approach is strictly related to the application of cellular automata in land use change research (Werner, 2005).

The application of cellular automata in geographical research is becoming increasingly popular (Takeyama & Couclelis, 1997). 'Cellular automata may be considered as (parallel-processing) computers, in which the initial configuration encodes the program and input data, and time evolution yields the final output' (Wolfram, 1982).

Cellular automata can be defined as mathematical models for complex natural systems containing large numbers of simple identical components with local interactions. They consist of a lattice of sites, each with a finite set of possible values. The

sites' value evolves synchronously in discrete time steps following identical rules. The value of a particular site is determined by previous values of the neighbouring sites. Cellular automata, alongside the agent-based model, are frequently applied in studies concerning the dynamics of complex spatial systems (Batty, 2008). White and Engelen (1997) represent the many authors of operational versions of CA models developed for the purpose of practical policy applications for the European Commission.

The cellular automata are similar to map algebra models in GIS. Numerous approaches have been undertaken with the aim of creating a uniform methodology binding both ideas (Takeyama & Couclelis, 1997). Map algebra is a cartographical modelling scheme which describes all the possible transformations of raster maps, grouping them into four basic classes which include: local, focal, global and zonal operations (Takeyama & Couclelis, 1997, Longley et al., 2006, after Tomlin, 1983).

- Local operations examine raster maps cell by cell, comparing a chosen layer's cell values with the values of the same cells in the other layers;
- Focal operations compare the value of each cell with the values of the neighbouring cells – this approach commonly involves comparisons with eight neighbouring cells;
- Global operations produce results which are true for the entire layer e.g. the layer's mean value;
- Zonal operations are used to compute results for blocks of contiguous cells that share the same value, e.g. for the calculation of shape of contiguous areas characterised with the same type of land use; they facilitate the assigning of results to all of the cells in each contiguous block.

### Neighbourhood coefficients (NBCs)

The concept of applying neighbourhood coefficients to the analysis of land use changes is based on the combination of map algebra with two-dimensional cellular automata. The research aims included the formulation of a theoretical structure of the neighbourhood coefficients (Werner, 2009), analysing their operationalization and verifying their practical application.

The NBC is calculated on the basis of a mathematical formula (Eq. 1) which contains the numbers describing land use classes and the consecutive numbers of the cells in the Moore neighbourhood. The operationalization of a research procedure requires the construction and application of a special filter to the original input map of the numbered land use classes.

The NBC is reversible. It is therefore possible to reconstruct (recalculate) the original input land use classes in the Moore neighbourhood (their nominal numbers) on the basis of the value of the central cell's NBC. This procedure allows modifications and extends the Geographical Model of land use (Hagoort, 2006, after Tobler, 1979) to enable the definition of rules of succession of land use classes when comparing maps concerning two or more years.

The mathematical foundation of NBC rests with modular arithmetic. The numerical operationalization of the proof is carried out by a spreadsheet application (using Eq. 1).

$$NBC_c = \sum_{i=0}^8 k_i * n^i \quad \text{where } k \in \{0, 1, \dots, n\}, \quad (\text{Eq. 1})$$

where:

$i$  – the consecutive number of a cell in the Moore neighbourhood,

$k$  – the consecutive number describing the class of land use in an ordered nominal scale,

$n$  – the total number of land use classes.

### Operationalization of the analysis and simulations

Rather than applying the mathematical formula directly (Eq. 1), the operationalization of the modified geographical model requires constructing a spatial linear filter which can be applied to the input maps of land use (where the land use classes are consecutively numbered). The results are identical. The aim is to obtain intermediate NBCs maps and retrieve a set of transition rules for the subsequent simulations (see Fig. 2).

The spatial filter is construed by applying the formula to each cell in the Moore neighbourhood (for a description see Eq. 1). For the two-dimensional, two-state cellular automata (i.e. 1 describes

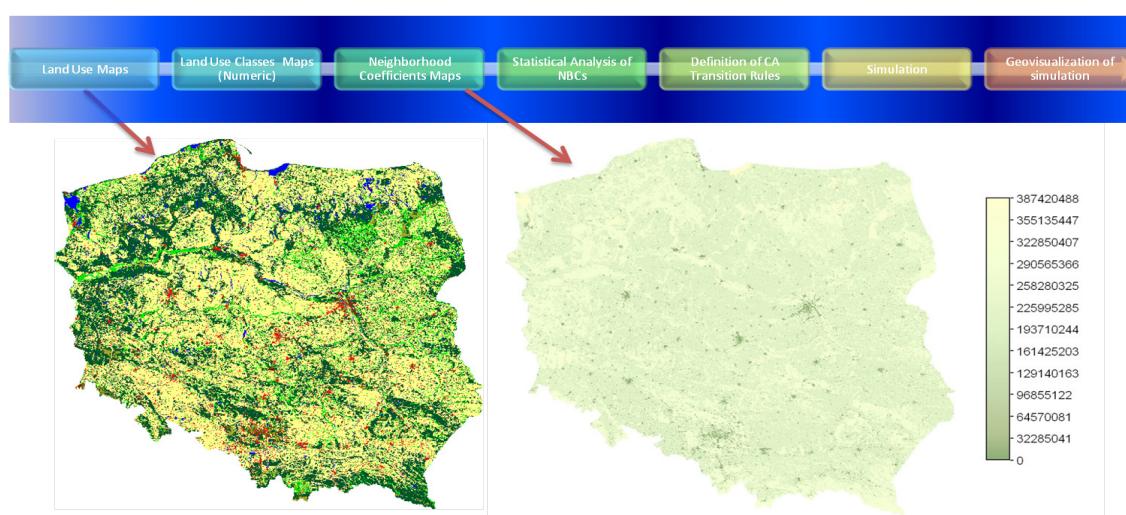


Fig 2. A flowchart of procedures

the urbanized areas and 0 the non-urbanized areas) the kernel of a spatial filter takes the form presented in Table 1.

Table 1. The kernel of linear spatial filter for the calculation of neighbourhood coefficients for two-dimensional two-state cellular automata

$2^0=1$	$2^1=2$	$2^2=4$
$2^3=8$	$2^4=16$	$2^5=32$
$2^6=64$	$2^7=128$	$2^8=256$

All of the 512 possible configurations of Moore Neighbourhoods of the urbanized and non-urbanized cells can be enumerated.

To illustrate: using real data, for example the spatial extent of the Warsaw agglomeration including its suburban areas in 1969 (excerpt from a map by F. Uhorcza, 1969), the application of the aforementioned spatial filter and the definition of the transition rules in comparison to the destination map of Warsaw, i.e. the generalized Corine Land Cover land use map (1996), demonstrated 79 defined transition rules (Werner, 2009).

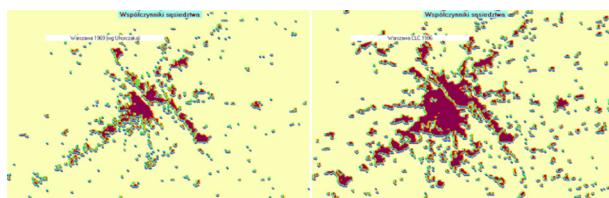


Fig 3. Neighbourhood Coefficients of urban areas – the Warsaw agglomeration in 1969 (left) and 1996 (right); ( $0 \leq NBC \leq 511$ )

The analysis of the statistical graph of successions between the neighbourhood coefficients in 1969 and 1996 demonstrates its visual similarity to the classic fractal (the Sierpinski triangle). The visual graph analysis is performed in order to identify the occurrence of transition rules.

The simulated Warsaw urban area in 1996 was then compared with the real extent in 1996. The Kappa Coefficient of similarity demonstrated a substantial agreement value.

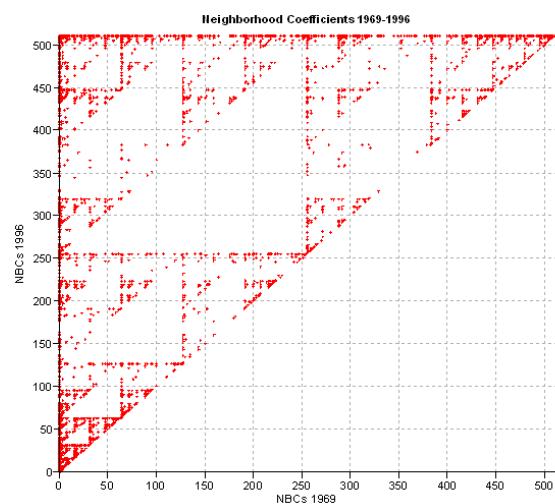


Fig 4. The succession of NBCs of the Warsaw urban area 1969-1996

Table 2. Statistics and Kappa Coefficient for simulated and real extent of Warsaw in 1996

Share (kappa=0.641)	Area of simulated urban extent (%)	Area out of simulated urban extent (%)	total %
Area of real urban extent (%)	7.53%	3.84%	11.37%
Area out of real urban extent (%)	3.22%	85.40%	88.63%
Total %	10.76%	89.24%	100.00%

The prognosis of the Warsaw urban area's spatial extent in 2023 was simulated on the basis of the real extent map in 1996.



Fig 5. The simulated spatial extent of the Warsaw urban area in 2023, using NBCs

#### Definition of transition rules of the cellular automata

At present, the stage of defining the transition rules is still not fully automated. The procedure includes the analysis of the descriptive statistics of the area of particular NBCs, the rejection of those that are less significant and tabular decomposition of the remaining NBCs into the original land use classes.

The calculation of the reverse algorithm of reconstruction of the original Moore neighbourhood can be performed using the following procedure: firstly, the modulo division of the NBC value (using the total number of land use classes as the divider) produces the original land use class in the NW (North-West) cell; secondly, the modulo division of the quotient of the previous calculation produces the value of the original land use class in the N (North) cell... and so on.

The only assumption made consisted of a hypothesis about the identical sequence of cells in NBC and the reverse neighbourhood cell value calculations. The reverse algorithm of NBC entailed the set of divisions as shown in Eq. 2.

$$k_i = (NBC_c / n^i) \bmod n, \quad \text{where } k \in \{0, 1, \dots, n\}, \quad (\text{Eq. 2})$$

where:

$i$  – the consecutive number of a cell in Moore neighbourhood

$k$  – the consecutive number describing the class of land use in an ordered nominal scale

$n$  – the total number of land use classes.

Both equations are true for any number of studied land use classes and can also be extended into the other types of neighborhoods. Computer resources and limits of computation with very large (huge) numbers are the only existing barriers for the procedures described.

Tabular manipulations of the original land use classes' values and NBCs, as well as their horizontal and vertical sorting, allow the transition rules to be defined in a conditional form: *if the starting centre cell is defined and the starting neighbourhood is defined and (some other conditions inferred...) then the destination centre cell takes the defined value, otherwise it remains intact.* The pilot project indicated the need for a more advanced research (Werner, 2009).

#### Application of neighbourhood coefficients for the prediction of land use changes - results and verification

The verification of the research and the procedures discussed is presented in the form of geovisualisations with maps and statistics of simulations of land use changes in Poland (1990-2006). The results of simulations performed with NBCs were then compared with analogous research using (SDSS) Metronamica™ Single Layer model for the Warsaw Metropolitan Area in Poland<sup>2</sup>.

The simulation of land use changes in Poland were performed on the basis of the Corine Land Cover maps (1990, 2000 and 2006). The following nine classes of land use were defined and numbered: (0) urban fabric, (1) industrial, commercial or transport areas, (2) mineral extraction, dump or construction sites, (3) green urban areas, (4) permanent crops or mosaic farmlands, (5) pastures, (6) standing forests, (7) natural & semi natural vegetation and (8) water courses and bodies. The NBCs maps were calculated for all three studied years and are presented in Fig. 6.

As the values of NBCs are very large (even massive), only the Moore Neighbourhoods were taken into account. Again, the shape of the graph of the NBCs succession resembled the Si-

erpinski triangle fractal. The performance of simulated land use in Poland in 2006 involved completing an analysis of changes which took place between 1990 and 2000 and constructing 45 transition rules which were then applied to the map of land use in 2000. The comparison of results is shown in Table 3.

The final aim of this research was to perform a simulation of land use in Poland in 2012 based on the above analysis, which can hopefully be verified shortly.

#### Concluding remarks and future applications

NBCs appear to prove their value for the purpose of land use changes simulation. In practice, the transitional rules are the subset of all changes that took place during the calibration years. They were not intended to cover all of the transitions of the geographic model. Although the research method is still at the inception stage, it encapsulates the complexity of factors and enables indirect inference to be made regarding spatial patterns of land use changes, bridging the idiographic and causality reasoning that concerns spatial processes.

The anticipated direct future application of NBCs includes the extension of Tobler's Historical Model of land use changes for example. NBCs could find further direct application in the preparation of maps of the neighbourhoods of land use (see Fig. 7). The only thematic data on the map represents urban neighbourhoods and the colours applied inform of the prevailing land use surrounding a given land unit. The NBCs could be employed in the analysis of spatial differentiation of complex territorial units, providing information on their environment.

This research project was conducted using the assembled open source and commercial GIS software: Iwisi, Map Comparison Kit™, ArcGIS™, Idrisi™, although the pivotal simulations and procedures were performed on Iwisi as this GIS system enables direct manipulation and transformation of raster cells using the scripting language.

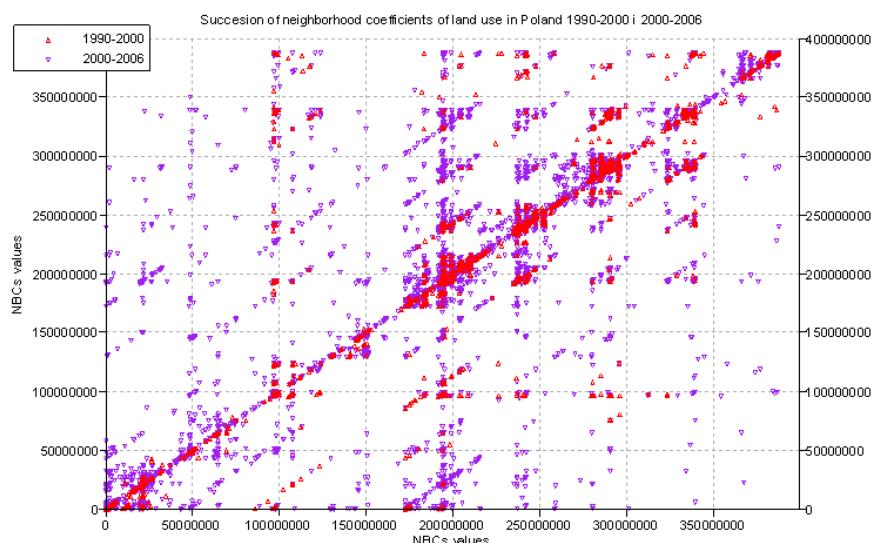


Fig 6. Neighbourhood coefficients of land use in Poland 1990-2000 and 2000-2006

Table 3. A comparison of real and simulated land use in Poland in 2006

Land use in Poland 2006	Real 2006	Simulation 2006	Kappa coefficients
	km <sup>2</sup>	km <sup>2</sup>	
Urban fabric	9984	10981	0.72
Industrial, commercial or transport areas	1448	1444	0.87
Mineral extraction, dump or construction sites	499	517	0.07
Urban green areas	673	512	0.64
Permanent crops or mosaic farmlands	169519	167121	0.92
Pastures	27357	31508	0.87
Standing forests	94985	92920	0.97
Natural & semi-natural vegetation	4855	4427	0.73
Water courses and bodies (including lagoons)	5392	5281	0.98

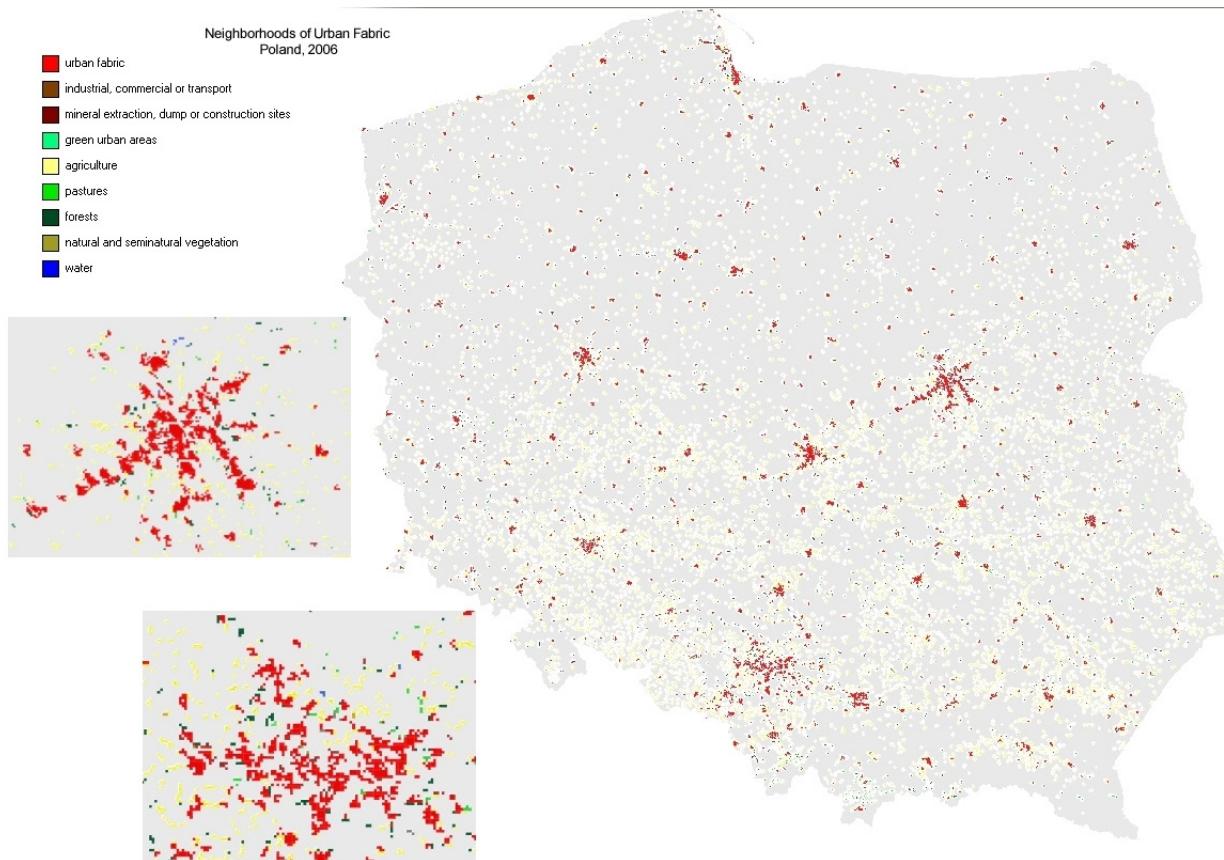


Fig 7. Neighbourhoods of the Urban Fabric in Poland (2006). Insets: upper - Warsaw agglomeration, lower - Upper Silesia conurbation

### Acknowledgements

Elements of this research which are related to the simulation of land use changes in Poland (Werner & Kozubek 2011) were conducted in cooperation with the Institute of Geodesy & Cartography, Warsaw, Poland, with the support of a research grant from the Ministry of Science and Higher Education.

## References

- Batty, M 2008, 'Fifty Years of Urban Modeling: Macro\_Statics to Micro-Dynamics' in S Albeverio, et al. *The dynamics of complex urban systems. A disciplinary approach*, Physica-Verlag HD, New York, pp. 1-20.
- Biłozor, A. 2003, 'Zastosowanie logiki rozmytej do delimitacji strefy przejściowej miasta i wsi' in: ed H Rogacki & T Czyż, *Problemy interpretacji wyników metod badawczych stosowanych w geografii społeczno-ekonomicznej i gospodarce przestrzennej*, Bogucki Wyd. Naukowe, Poznań, pp. 215-223. (In Polish)
- Ciolkosz, A. & Bielecka, E 2005, Pokrycie terenu w Polsce. *Bazy danych CORINE Land Cover*. Inspekcja Ochrony Środowiska, Warszawa. (In Polish)
- Fotheringham, SA, et al. 2000, *Quantitative geography: perspectives on spatial data analysis*. SAGE Publications Ltd, London.
- Hagoort, MJ 2006, *The Neighborhood Rules. Land use interactions, urban dynamics & cellular automata*. Netherlands Geographical Studies, Utrecht.
- Korcelli, P, et al. 2011, *Warsaw – Spatial Growth with Limited Control*, IGiPZ PAN, Warsaw [2010], /typescript/.
- Kozubek, E 2002, *Zmiany użytkowania ziemi w regionie tarnobrzeskim pod wpływem uprzemysłowienia...* IGiPZ PAN, Warszawa. (In Polish)
- Longley, PA, et al. 2006, GIS. *Teoria i praktyka*, PWN, Warszawa. (In Polish)
- Mishan, EJ 1969, *Welfare economics: Ten introductory essays*. Random House, New York.
- Ratajczak, W 1999, *Modelowanie sieci transportowych*, Wyd. UAM, Poznań. (In Polish)
- RIKS BV 2010, *Metronamica Documentation*, The Netherlands: s.n., Maastricht. Available from <<http://www.riks.nl>>.
- Takeyama, M & Couclelis, H 1997, Map dynamics integrating cellular automata and GIS through Geo-Algebra. *Int. J. Geographical Information Science*. pp. 73-91.
- Werner, P 2009, Application of Cellular Automata and Map Algebra in Studies of Land Use Changes. The Neighbourhood Coefficients Method. *Geoinformatica Polonica*, vol. 9, pp. 7-20.
- Werner, P & Kozubek E 2011, Visualization of land use changes in Poland based on neighbourhood analysis (1990-2006), *Proceedings of XXVth International Cartographic Conference (ICC)*, Paris, France, 3-8 July 2011, International Cartographic Association, Available from <[http://icaci.org/documents/ICC\\_proceedings/ICC2011/](http://icaci.org/documents/ICC_proceedings/ICC2011/)>
- Wu, F 1999, GIS-based simulation as an exploratory analysis for space-time processes. *Journal of Geographical Systems*. vol. 1, pp. 199-218.