A Fuzzy Cellular Automata Based Shell for Modeling Urban Growth – A Pilot Application in Mesogia Area

Eleftherios A. Mandelas[1], Thomas Hatzichristos[2], Poulicos Prastacos[3]

[1] Regional Analysis Division, Institute of Applied and Computational Mathematics, Foundation for Research and Technology, Vasilika Voutwn, GR 71110, Heraclion Crete, Greece
mantelas@iacm.forth.gr

[2] Department of Geography and Regional Planning, National Technical University of Athens - Hellas, I.Politechniou 9, GR 15786, Zografou Athens, Greece
thomasx@survey.ntua.gr

[3] Regional Analysis Division, Institute of Applied and Computational Mathematics, Foundation for Research and Technology, Vasilika Voutwn, GR 71110, Heraclion Crete, Greece
poulicos@iacm.forth.gr

Keywords: Urban Growth, Rule-based Modeling, Cellular Automata, Fuzzy Logic.

1 INTRODUCTION

Urban growth is a global phenomenon and one of the most important reforming processes effecting both natural and human environment through many ecological and socioeconomic aspects. Given the recent growth rates and the fact that urban society's needs in space, services, facilities and energy, grow as it's population increases, it is of major importance for present and future societies that urban growth occurs in an optimal way, maximizing the benefits for urban population while minimizing both economical and environmental cost.

The term 'modeling' refers to creating a strictly defined analog of real world by subtraction, considering only the most important parameters (Koutsopoulos, 2002), yet there is no rigorous framework for modeling such a spatio-temporal phenomenon, as urban growth since there lies great inherent spatial, temporal and decision-making heterogeneity, which results from socio-economic-ecological heterogeneity itself (Cheng, 2003). There have been proposed many methods attempting to model urban growth; recent trends among other include cellular automata (CA).

There is a great variety of different crisp approaches concerning CA based urban growth modeling, such as the stochastic approach of Mulianat *et al* (2004), the object-oriented approaches of CAGE (Blecic, 2004) and OBEUS (Benenson, 2006), the approach proposed by Morshed (2002) and the environment LAUDE that combine CA and Genetic Algorithms and the widely applied models, SLEUTH (Dietzel, 2004) and GEONAMICA. Recently a few fuzzy approaches have been proposed, such as the approach of Vancheri *et al* (2004) which introduces fuzzy data management and the approaches proposed by Dragicevic (2004) and Liu

(2001) that use fuzzy-crisp rule based systems to determine the evolution of the CA. This paper presents an urban growth modeling framework that in order to simulate urban growth's mechanisms incorporates a cellular automata technique while it adopts a fully adjustable fuzzy logic inference engine.

2 CELLULAR AUTOMATA

Cellular Automata (CA) are a computational method, capable of simulating growth process by describing a complex system through a set of simple rules (Krawczyk, 2003). Each cell is interacting with other cells in predefined neighborhoods while interactions among cells take place in discrete time steps. The cell's next status is calculated considering its neighborhood's status, while this process takes place for all cells and repeats itself. This is the inherent mechanism according to which CA is evolving in space, through time, propagating information in a self-reproductive way with no external interference. CA is a bottom-up technique that accesses the global behavior of a system by local interactions. It is of this linkage between micro-macro approaches that CA consist an advisable technique for complex phenomena simulation and has been applied to various science fields, such as numerical analysis, fluid dynamics, simulation of biological and ecological systems, traffic analysis and urban growth modeling.

3 FUZZY LOGIC

Fuzzy logic was originally proposed in 1965 by Lotfi A. Zadeh as a generalization of binary logic (Kirschfink, 1999); is used to model imprecision, vagueness and uncertainty in real world. In fuzzy logic, variables consist of partially overlapping fuzzy sets, which form qualitative groups of values within given ranges of values. A linguistic variable is assigned to each fuzzy set in order to describe its properties. In order to convert crisp numerical variables into fuzzy, fuzzy sets are fully defined by membership functions, which return a membership value (μ) within [0,1] for a given crisp object in the fuzzy set. Linguistic variables allow of defining new sets, based on the existing ones by applying fuzzy hedges, such as 'more or less', 'not', 'very' etc; each hedge is joined to a numerical expression.

The knowledge base is represented as "IF…THEN" rules, connecting hypotheses to conclusions through a certainty factor. Fuzzy inference is divided into the stages of aggregation, implication and accumulation (Kirschfink 1999, Hatzichristos 2001). Aggregation returns the fulfillment of hypothesis for every rule individually, implication combines aggregation's result to the rule's certainty factor (CF) resulting to the degree of fulfillment for each rule's conclusion, while accumulation corresponds to compromising different individual conclusions into a final fuzzy result.

4 MODELING SHELL STRUCTURE

The scope of this work is the development and the application of a methodological framework which incorporates fuzzy cellular automata techniques in order to simulate urban growth. Till now we have mainly focused on the Fuzzy Cellular Automata Module. The 2-D CA lies upon a normal grid formed over raster data being able to use various neighborhoods.

Cells' states are defined as memberships in fuzzy sets while transitions are determined by a fuzzy system built within that manages spatio-temporal rules. Variability over time and space is gained by incorporating spatio-temporal parameters within each rule's CF. The fuzzy transition system incorporated, allows of specialized logical connection of premises and uses various numerical operators and linguistic hedges, hence it can deal with almost any given form of knowledge base.

To achieve optimized computational behavior there are actually used two separate fuzzy systems dealing with separate cases. The first one uses a subset of the time-stationary data in order to conclude over an inner fuzzy variable named 'propensity' which is used as an input to the second fuzzy system. While the first system runs only ones, this system is iterative and each time-step's output is the input for the next step.

Though the main idea of combining CA and fuzzy logic exists in previous approaches, the herein presented approach contains some advantages. Liu (Liu, 2001) uses fuzzy-crisp rules; fuzzy hypothesis lead to predefined crisp patterns of growth expressed as logistic curves, while Dragicevic (Dragicevic, 2004) uses modified membership functions to model the transitions. Our approach combines some attributes of the above approaches, discards some others while brings forward some new ones:

- more than one, fuzzy sets are used to describe fuzzy urban status which allows of qualitative analysis and the use of urban cover sub-classes,
- transitions are not limited among adjacent sets, yet when using ordered sets, transitions are only allowed towards a higher state
- fuzzy hedges are enabled both in rule's hypothesis and conclusion,
- a complete fuzzy algebra is incorporated
- spatio-temporal parameters are taken into consideration within each rule's certainty; more specifically, spatial variability is gained by defining a spatial (2-D) fuzzy variable which expresses the relative location of a cell within the study area (figure 1).

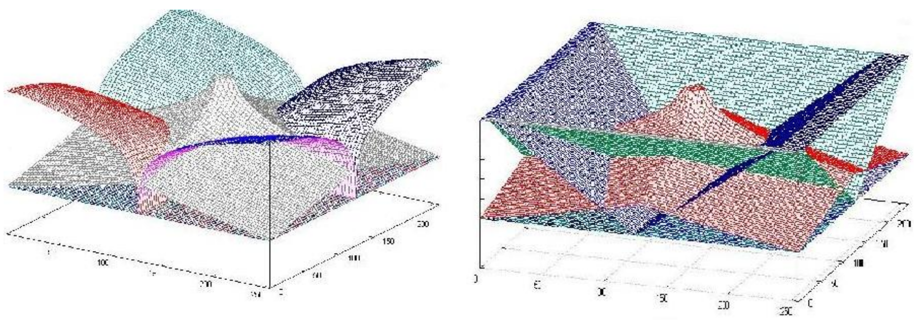

Figure 1: **A** Graph of the 2-D spatial fuzzy sets corresponding to the center and the 4 primary wind directions (left) and to the center and the coupled combinations of the 4 primary wind directions (right).

5 PILOT APPLICATION

The Fuzzy Cellular Module and the primary rule-extraction module were applied in Mesogia area which lies in the eastern space of Attica basin and has presented intense urban growth. Available data were urban cover for the years 1994, 2001 and 2004, Corine land cover for the years 1994 and 2004 and road network. Road network data were separated in primary and secondary and layers of euclidian distance and line density for each cell were derived for both primary and secondary using ArcToolbox. Corine based classification was regrouped according to each class's proportion that got urbanized during the given periods. Furthermore, urban cover's borders got fuzzified to provide graduation between urban and non-urban areas.

5.1 Knowledge Base Extraction

For a fuzzy knowledge base to be both efficient and comprehensible, fuzzy sets should be carefully defined and the number of the sets should be the least that adequately expresses qualitative diversifications of the objects. After various settings were tested, most fuzzy variables were decided to consist of 3 linear sets corresponding to the traditional form of 'low', 'average' and 'high'. Corine based suitability consists of a sole linear set, while fuzzy urban cover consists of 3 non-linear sets (as shown in figure 2). Numerical data where fuzzified in cell level and aggregate fuzzy sets were calculated for Moore neighborhoods of radius 1 to 7 cells, which led to an explosive growth of fuzzy data. Fortunately, further analysis revealed that using neighborhoods of radius 1 led to a loss of statistic significance 0.9%. Table 2 shows the R-criterion value for the full regression model using various neighborhoods.

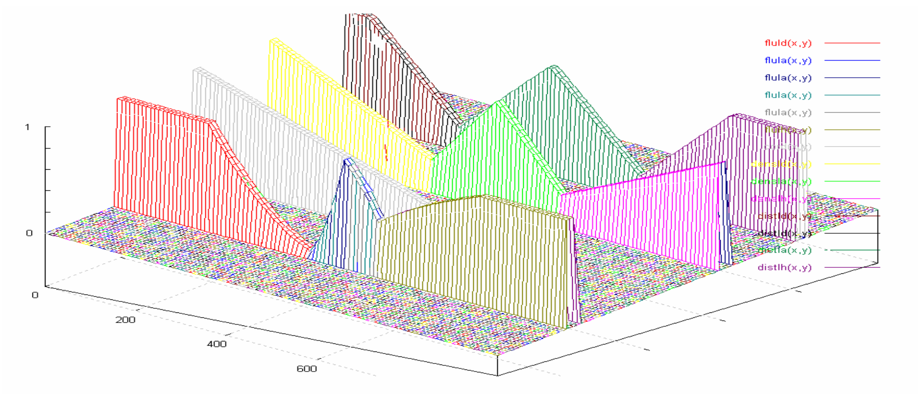

Figure 2: plots of fuzzy sets used, in the front lies fuzzy urban cover and as we move backwards, Corine based suitability, road density and road distance.

Neighborhoods	1	1, 2	3, 4	5, 6	1, 6	1, 5	3, 5	1-6
R^2_{adj} %	75,4	76,1	74,7	72,7	75,8	75,8	74,5	76,3

Table 1: Statistical significance of full data set used in various neighborhoods.

The multiple regression models composed rules using as implication operator the weighted sum operator. Moreover variables whose presence drove to the same conclusion and appeared not to be highly correlated to each other formed complex premises connected using the 'AND' operator while those that appeared to be highly correlated exist in the same premise connected by the 'OR' operator. Each rule's CF was calculated taking into consideration the R-value of the regression equation. Discarding the effect of temporal variables' presence a new set of rules was formed. Those rules were used in the first fuzzy system of the model and instead of concluding over urban cover status they concluded over the inner fuzzy variable named urban propensity. By replacing the effect of static variables in the primary set of rules with the inner variable 'propensity', the knowledge base of the second fuzzy system was extracted.

First results implied a spatial heterogeneity in error and a general urban underestimation tendency. In order to deal with the latter, we introduced the use of the fuzzy hedges "more", "most" and "extremely" whose numerical expressions are shown in Table 2 while figure 3 illustrates an example of the effect of the fuzzy hedge "extremely". Moreover in order to a posteriori enhance the model's local behavior in different sub areas within the study area, we introduced the use of the spatial fuzzy variable within each rule's CF. Under the final settings of the system, each time-step corresponds to an actual year; a part or the final knowledge base used is given in tables 3 and 4.

Fuzzy Hedge	Numerical Interpretation
extremely	$\mu_{out} = \mu_{in}^{2+5(0.8-\mu_{in})}$, if $\mu_{in} \geq 0.8$ $\mu_{out} = \mu_{in}^{2+40(0.8-\mu_{in})}$, if $\mu_{in} \leq 0.8$
more	$\mu_{out} = \mu_{in}^{0.88}$
most	$\mu_{out} = \mu_{in}^{0.66}$

Table 2: Fuzzy hedges applied in the system and corresponding numerical expressions.

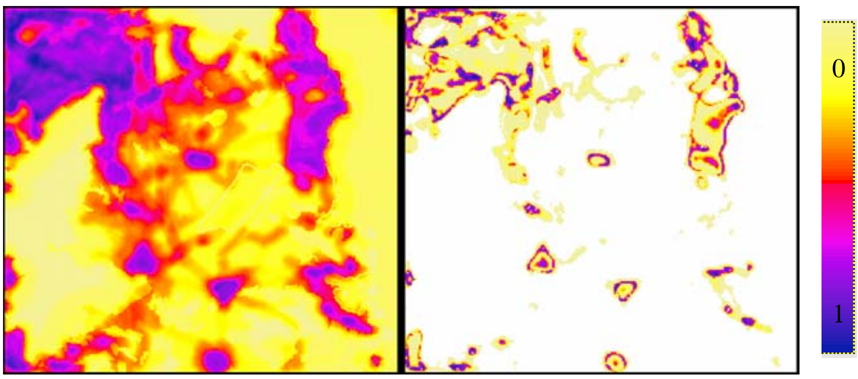

Figure 3: The effect of the fuzzy hedge 'extremely' within the study area when applied to the fuzzy set 'propensity high'.

IF	{HYPOTHESIS}		{CONCLUSION}	CF
	propensity low	=>	urban cover low	0.8
	urban cover low	=>	urban cover low	0.7
	urban cover average AND propensity extremely high	=>	urban cover average	0.7
	propensity extremely high	=>	urban cover average	0.7
	urban cover high AND propensity high	=>	urban cover higher	1.0
	urban cover high AND propensity extremely high	=>	urban cover highest	1.0
	urban cover average AND propensity high	=>	urban cover higher	0.5

Table 3: Knowledge base concluding over the fuzzy variable 'urban cover'.

IF	{HYPOTHESIS}		{CONCLUSION}	CF
	Corine not proper	=>	propensity low	1.0
	main road distance high	=>	propensity low	1.0
	secondary road density low	=>	propensity low	0.5
	secondary road density in MOORE-1 low	=>	propensity low	0.9
	main road density low	=>	propensity low	0.4
	main road density in MOORE-1 low	=>	propensity low	0.1
	secondary road distance low	=>	propensity high	0.9
	secondary road density in MOORE-1 high OR secondary road density average in MOORE-1	=>	propensity high	0.7 +SE
	main road density high OR main road density average	=>	propensity high	NE

Table 4: Knowledge base concluding over the intermediate fuzzy variable 'propensity'.

5.2 Model's Validation

In order to validate the model's behavior there were defined 4 indicators testing the fitting of the results to real data, which were calculated for the cases of forecasting the urban cover for the years 2001 and 2004. The indicators that were defined are:

- I-1 : the percentage of cell's that were correctly estimated to be of low urban cover status,
- I-2 : the percentage of cell's that were correctly estimated to be of high urban cover status,
- I-3 : the percentage of cell's whose urban cover status (either 'low' or 'high') was correctly estimated,
- I-4 : L-1 norm of the error.

The values of the above defined fitting indicators for each one of the 3 cases studied are given below in table 5.

forecasting	I-1	I-2	I-3	I-4
2004 based on 2001 (step3)	0.977	0.946	0.970	0.128
2001 based on 1994 (step7)	0.981	0.793	0.938	0.162
2004 based on 1994 (step10)	0.974	0.761	0.918	0.167

Table 5: Fitting Indicators.

6 CONCLUSIONS AND FUTURE WORK

Under the assumption that spatial patterns under which urban growth occurred in the area of Mesogia during the last decade won't get dramatically altered during the next one, our model simulates in a satisfactory way the evolution of urban growth (as shown in figure 4.a). According to the fitting indicators, the model simulates efficiently the qualitative patterns under which urban growth occurs for short periods; moving further into time though, error accumulation due to iterations is becoming significant. This is partially caused by the poor input data that in some cases do not allow of efficient qualitative discreteness (figure 4.b).

Technically, cellular automata and fuzzy logic are combined in a mostly satisfactory way. The pilot application indicates that a well tuned CA is capable of simulating the spread of urban cover; it appears less efficient though when it comes to "fresh" urban areas. Fuzziness is an efficient way of dealing with insufficient and vague data since it enhances the potential qualitative discrimination of the model. Fuzzy hedges used, appear to be most effective while spatial rules are improving the model's behavior, yet till now they are subjectively set. Though the model supports temporal rules, the rules' extraction technique that is used in the pilot application does not.

Much work still needs to be done concerning automated rule-extraction processes. At this time a technique using optimization theory has been designed in order to extract qualitative relations between fuzzy variables while it determines at the same time the optimum number of fuzzy sets for each variable and the form of their membership functions.

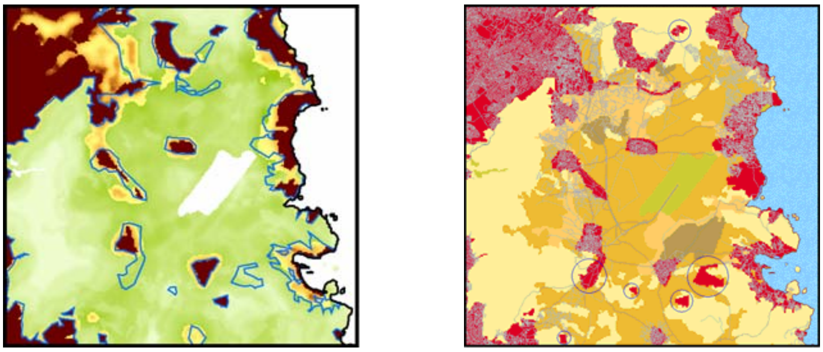

Figure 4: **a)** Forecasting 2004 based on 1994, the yellow to brown graduation indicates high urban cover status estimation while within blue lines lie the actual urban areas (left), **b)** areas that slip from the discreteness of available data (right) are indicated within blue circles.

BIBLIOGRAPHY

Beauregard R. A., “History in urban theory" Journal of urban history”, vol. 30, no.4, p.627-635, 2004

Benenson I., Kharbash V., “Geographic Automata Systems and the OBEUS software for their implementation’ in “Complex Artificial Environments”, p. 137-153, Springer Berlin Heidelberg, 2006

Blecic I., Cecchini A., Prastacos P., Trunfio G.A., Verigos E., “Modelling Urban Dynamics with Cellular Automata: A Model of the City of Heraclion”, 7th AGILE Conference on Geographic Information Science, 2004

Cheng J., Masser, I., “Understanding urban growth system: theories and methods”, 8th International Conference on Computers in Urban Planning and Urban Management, 2003

Cox E.: “The Fuzzy System Handbook”, Academic Press, Inc. 1994

Dietzel Ch., Clarke K. C., “Replication of Spatio-temporal Land Use Patterns at three Levels of Aggregation by an Urban Cellular Automata”, Lecture Notes in Computer Science, vol. 3305, pp. 523-532, 2004.

Three Levels of Aggregation by an Urban Cellular Automata”, Lecture Notes in Computer Science, vol. 3305, 2004

Dragicevic S., “Coupling fuzzy sets theory and GIS-based cellular automata for land-use change modeling”, 2004

Ganguly N., Sikdar B.K, Deutsch A., Canright G, Chaudhuri P.P., 2003, “A Survey on Cellular Automata”, Technical Report, Centre for High Performance Computing, Dresden University of Technology.

Gar A., Yeh O., “Digital Urban Planning - The Use of GIS in Urban Planning”, International symposium and workshop on “managing heritage environment in Asia”, 2003

Hall L., “Rule Chaining in Fuzzy Expert Systems”, IEEE Transactions on Fuzzy Systems, V. 9, No. 6, pp. 822-827, 2001

Hatzichristos Th., “GIS and Fuzzy Logic in Spatial Analysis”, Educational notes, NTUA 2001

Kirschfink H., Lieven K., “Basic Tools for Fuzzy Modeling”, Tutorial on Intelligent Traffic Management Models in Helsinki, 1999

Koutsopoulos K., “Geographic Information Systems and Spatial Analysis”, Papasotiriou 2002

Krawczyk R.J., “Architectural Interpretation of Cellular Automata”, 2003

Liu Y., Phinn S.R., “Developing a Cellular Automaton Model of Urban Growth Incorporating Fuzzy Set Approaches”, Proceedings of the 6th International Conference on GeoComputation, 2001

Longley P., Higgs G., Martin D., “The predictive use of GIS to model property valuations”, International Journal of Geographical Information Systems 8, 1994

Morshed A., “Land Use Change Dynamics: a Dynamic Spatial Simulation” , PhD Thesis, 2002

Mulianat I., Hariadi Y., “Urban Area Development in Stochastic Cellular Automata”, Urban/Regional 0412001, EconWPA, 2004

Pullar D., “Rule-Based Modelling In GIS”, Proceedings of GeoComputation ‘97 & SIRC ’97, p. 15-23, 1997

Sarkar P., “A Brief History of Cellular Automata” , ACM Computing Surveys (CSUR), vol. 32, Issue 1, p. 80 – 107, 2000

Stefano B., Fuks H., Lawniczak A.T., “Application of Fuzzy Logic in CA/LGCA as a Way of Dealing with Imprecise and Vague Data” , Canadian Conference on Electrical and Computer Engineering, vol. 1, p. 212-217, 2000

Stocks Ch. E., Wise St., “The Role of GIS in Environmental Modelling”, Geographical & Environmental Modelling, vol. 4, no.2, p.219-235, 2000

Torren P. M., Benenson I., “Geographic Automata Systems”, Geographical Information Science, vol. 19, no.4, p.385–412

Vancheri A., Giordano P., Andrey D., Albeverio S., “A model for urban growth processes with continuum state cellular automata and related differential equations” , 2004

Waddell P., Ulfarsson G.F., “Introduction to Urban Simulation: Design and Development of Operational Models”, In Handbook in Transport, Volume 5: Transport Geography and Spatial Systems, Stopher, Button, Kingsley, Hensher eds. Pergamon Press, 2004, pages 203-236.