Advanced Geosimulation Models

Editors

Danielle J. Marceau
University of Calgary
Canada

&

Itzhak Benenson
Tel-Aviv University
Israel

CONTENT

FOREWORD

Geosimulation is the term coined by Benenson and Torrens (2005) in their book of the same name to describe a class of simulation models that focus exclusively on systems whose structure and function is dominated by geographical (or spatial) representations and processes. In fact, this term might cover all models of spatial systems but the focus is restricted to representations characterized by elemental objects ranging from individuals, in the literal sense, in human systems, to automata, in the virtual sense, in computer systems. Furthermore, these objects, which are often called agents or cells, embody explicit behaviors that need to be simulated. In this sense, they intrinsically reflect the notion of change or dynamics in their systems of interest, often translating into movements and flows that define system structure and function. These kinds of models are part of a much wider trend in simulation which is referred to generically as agent-based modeling (ABM) of which cellular automata (CA) modeling represents a particular case. It is these models that comprise those presented in this book.

In one sense, geosimulation is part of the sea change that has occurred over the last 25 or so years of modeling in both the physical and social sciences. In the mid 20th century, scientists tended to articulate their systems of interest as though they were organized and constructed from the top down. Systems theory and the systems approach that grew out of engineering and biology before and during the Second World War, was predicated on the analogy with systems that were closed and controlled by negative feedback and thus, quite well-defined and well-behaved. Generally these were systems that could be conceived of as being in equilibrium. Although deviations from the equilibrium state could occur, processes of negative feedback were always able to restore or move the system to a new equilibrium in smooth and tractable ways.

This model of the world which came to structure much simulation in the 1950s and 1960s, was found to be wanting in many ways. Spatial systems such as those of concern in geosimulation, tend not to be in equilibrium in the same way that closed mechanical systems, the analogues of the systems approach, were. Cities, for example, are places of vibrant change, great diversity and heterogeneity. As they have become more complex as the wealth of their inhabitants has loosed constraints on social and spatial behavior, top-down equilibrium seeking models that represent systems at one cross-section in time, have appeared increasingly out of touch with both understanding and policy analysis. Moreover, these earlier models, particularly those used in city and transportation planning, were often highly aggregate in their treatment of populations, and inevitably the quest began to disaggregate them to a level where at least individual groups could be defined that represented more consistent average behaviors. For example, aggregate spatial interaction models built around analogues with classical physics, often referred to as social physics, gave way to discrete choice and thence activity models of travel behavior while demographic processes were disaggregated to the point where techniques of micro-simulation became their *modus operandi.*

Despite these trends towards finer and finer disaggregation, the notion of developing dynamic behaviors of relevant population groups has forced the field to consider models with purposive behavior. As population aggregates are fractured into ever greater detail, there comes a point where they reach some elemental level which is, in itself, self-contained with respect to its behaviors. In simulating mobility, for example, this is at the level of the pedestrian or the vehicle whereas for spatial objects such as the space itself, then it is the cell that can take on this relative independence from the aggregate. The sea change that has occurred in how we perceive spatial systems is reflected in this switch from top down to bottom up. Agents rather than their aggregates now constitute the key elements in model representation with the processes that engender their change through time and over space being the focus of simulation. Agent-based models have thus come firmly onto the agenda with variants such as cellular automata and more relaxed versions of these in the form of cell-space and cell-state models now forming part of the conventional wisdom.

These new models that constitute the heartland of geosimulation, are much richer than their more aggregate, cross-sectional equilibrium-seeking counterparts. As such, they are much harder to validate against data in that they embody many processes that are plausible but for which data is rarely available. Indeed, process-based models are extremely difficult to fit unambiguously for usually data about key processes has to be assumed or is not available and somewhat indirect methods are thus required to validate the model against whatever data and assumptions are made explicit. This book develops many of these challenges presented first in chapter by Marceau and Benenson which deals with key issues facing the field. In essence, these challenges relate directly to the 'geo' in simulation, focusing on spatial representation, model validation, visualization and dynamics which are dealt with in the sequence of seven chapters that constitute the bulk of this book. In fact, in a number of the chapters, all these various themes are interwoven but in the first two chapters, the focus is on spatial representation. CA models invariably define spaces based on regular tessellations such as grids in 2-dimensions, occasionally 3 (for urban and environmental systems) but in the strict applications of these automata, neighborhoods around cells are defined in the most limited sense. Indeed emergence in such models as in the evolution of fractal patterns and forms can only be guaranteed if neighborhoods based on nearest neighbors are used. In spatial systems however, there is a need to overlap neighborhoods, to relax their extent to cover more than their nearest neighbors and to deal with hierarchies of cells.

In a model for the Dublin region, White, Shahumyan, and Uljee develop what they refer to as a variable grid model that seeks to deal effectively with these concerns. In a related chapter which follows, Moore demonstrates how the regularity and homogeneity assumptions of the grid (which is largely fashioned after raster representations in GIS) can be relaxed to non-regular shapes. In this, rasters are replaced by vectors and he demonstrates how such irregularities can be produced in visualizing von Thunen's model of agricultural land use where concentric rings are replaced by irregular zones that require iteration in the solution of the model. This is an intriguing demonstration of the effect of agents modifying their own geometric surroundings and it opens the door to the notion that geosimulation models might also be developed so that agents actually define the spatial system on which they operate rather than simply engage in processes that take place on a fixed representation.

Issues of visualization are always central to geosimulation for the focus on space and its complexity is intrinsically visual. In complex models, Mandelbrot's long standing dictum that 'the fit of the model is in the seeing' is a benchmark that guides research and development. To this end, there are three chapters in this book that follow this discussion on representation and these focus on geo-visualization, particularly on the construction of geographic virtual environments, on 3-d representations, and on the use of multimedia in both the scientific development and the dissemination of outcomes. Crooks, Hudson Smith, and Patel present a portfolio of model applications involving multi-agent systems that are urban and spatial in intent, focusing largely on agents that move in urban space such as pedestrians and vehicular traffic. They show how 3-D representation and motion can be used as building blocks for embedding such models into virtual exhibition spaces and show how relatively straightforward it is to build models in scripting languages such as **Repast** and **Netlogo** which produce outputs that can be immediately visualized in interactive virtual worlds. A nice feature of these models is that users can enter the scene as avatars and mix with virtual agents, thus involving real time interaction of real with modeled behavior, opening up possibilities of models that are mixtures of the real and the virtual. Mekni, Moulin, and Paris ground these ideas in more structured contexts by building semantically-enhanced virtual geographic environments. This is a rather different approach to agent-based modeling than the usual one which seeks to visualize outcomes rather than exploit visualization capabilities for the very behaviors that are being simulated. In short, embedding virtual geographic environments within the decisions that agents make with respect to their spatial behaviors turns the problem on its head: visualization is being used here to extend the behavioral abilities of agents as well as inform the model-builder as to agent behaviors and behavioral outcomes. In some senses, this is the wave of the future where visualization of the cellular space in which agents behave becomes the focus of the simulation itself.

In the next chapter, Heckbert and Bishop change the focus to the empirical calibration of agent-based models but embed these in various visual environments, thus picking up on ideas from the previous two chapters. In fact visualization is one of the key ways in which validation is developed in these models. Three data-based approaches to validating agent model are developed for landscape problems, with three models being developed, the first using discrete choice survey to estimate the parameters of an ABM from survey data, the second using participatory approaches to reveal preferences for locational activity, and the third using data extracted from an experimental economics-like setting which is used to tune similar kinds of model. In all these examples, the key problem of reconciling observed data on outcomes with implied data on processes is identified, showing how important it is to get to grips with the problem of validation in such models. The implication of their work is that models should be tuned to available data, not the other way around and that model specification should be flexible to the validation process: changing the model is likely to be as important a strategy in geosimulation where the goal of parsimony in model design will always be problematic, as finding data that represents as many aspects of the model as possible.

This emphasis on validation is then continued in more conventional terms by Hatna, and Benenson who extend their work on residential segregation in ethnically diverse cities using the distribution of incomes. In essence, they extend and enrich Schelling's (1969) model of residential segregation, departing somewhat from Schelling's demonstration that high tolerance levels for dissimilar populations in their immediate neighborhoods can lead to serious segregation and a low level of diversity. In their examples for Israeli cities, they find the opposite and thus they operationalize the model using ideas about income distributions. One of the key features of their approach is that amongst many ABMs and much geosimulation, their work stands out as being empirically robust in that they have extensive data on outcomes, if not on the processes that lead to these outcomes. Once again they show the way in terms of model calibration and validation. The last chapter by Straatman and Marceau takes geosimulation to a new realm: they develop models of what they call 'open-ended' agent-based economic evolution where the focus is entirely on processes that do not close on final outcomes. Models of an evolving economy are used and made tangible through input-output structures that enable new technologies to be produced and diffused. In some respects, the traditions that draw on might be called *econo-simulation* rather than *geosimulation* and although their models are a-spatial at best, they do open the way to enriching the field of geosimulation with ideas from agent-based computational economics and, like Heckbert and Bishop before, to ideas from experimental economics. This, in many ways, is a fitting conclusion for it throws wide open the idea of geosimulation and impresses upon the reader the dynamism and curiosity that is central to the ideas developed here.

These interesting and informative contributions mark out the frontier in geosimulation in particular, and in cellular automata and agent-based modeling in general. Readers will gain much from what follows and the editors are to be applauded for collecting together a set of chapters that inform and extend the state of the art. The online presence of these contributions will ensure that these ideas will be available to a potentially very wide set of interests.

Michael Batty
CASA
University College London

REFERENCES

Benenson, I., and Torrens, P. M. (2004) **Geosimulation: Automata-Based Modeling of Urban Phenomena**, John Wiley, London.

Schelling, T. C. (1969) Models of Segregation, **American Economic Review**, **59**, 488-493.

PREFACE

The desire to understand our world and foresee its future is embedded deeply in the human mind. Starting from the 1960s, our forecasts are based on computer simulations. New computing facilities enforce the development of novel methods of scientific inquiry and enable the investigation of problems that only a decade ago were considered overcomplicated.

As part of this computer-based evolution of science, *Geosimulation* has recently emerged at the intersection of Geographic Information Science, Complex Systems Theory and Computer Science. Geosimulation aims at understanding the dynamics of complex human-driven spatial systems through the use of spatially explicit computer simulation. Geosimulation sees dynamic spatial systems, such as cities, as consisting of, and driven by, the human agents; the system's dynamics becomes, thus, the synergetic output of agents' actions and interactions. Numerous human agents behave in urban space pursuing their goals and the model simulates their spatial, social and economic activities. In this way, Geosimulation model becomes a laboratory for exploring complex human systems and their possible paths of development and evolution.

The perspectives of Geosimulation are defined by the advances in representation of geographic space and analysis of spatio-temporal dynamics of human-driven systems including emergence, path dependence, hierarchy and multi-scale dependency. The approaches and tools for validating Geosimulation models are especially important for understanding their complex and spatially heterogeneous outcomes.

This book originates from a kind invitation of *Bentham Science Publishers* and their new publication venture of electronic books. The *Advanced Geosimulation Models* pursues Itzhak Benenson's and Paul Torrens' 2004 book *Geosimulation – Automata-Based Modeling of Urban Phenomena* in which they defined the field of Geosimulation and provided a thorough assessment of its potential contributions for the urban and related phenomena. Considering the growing influence of Geosimulation on several disciplines, this book aims at presenting to the scientific community the recent conceptual and methodological advances achieved in the field.

The book targets scientists and graduate students working in the fields of Complex Systems Modeling, Geocomputation, GIScience, Geography, Regional Science, Computer Science, Artificial Intelligence, Environment Simulation and Modeling, and Environmental Engineering who are developing geosimulation models and who want to learn about the recent developments in the field. It will also be of interest to practitioners attracted by the potential of Geosimulation for managing complex urban and environmental problems.

We want to thank all of those who have contributed to the realization of this book. We are grateful to all the authors who have responded enthusiastically to the invitation of providing a high-quality manuscript. We deeply appreciate the external reviewers for their deep and constructive comments. Finally we thank the people at Bentham Science Publishers, and, especially, Ms Sara Moqeet, Assistant Manager of the E-books Publications Department, for her patience and support during the production process of this book.

List of Contributors

Danielle J. Marceau

Department of Geomatics Engineering
University of Calgary, Calgary, Alberta, Canada
2500 University Dr NW, Calgary, AB
T2N 1N4, CANADA
Phone: 403-220-5314
E-mail: dmarceau@ucalgary.ca

Itzhak Benenson

Department of Geography and Human Environment
Tel Aviv University, Israel
Ramat Aviv, 69978, Tel Aviv
Tel: +97236409896
E-mail: bennya@post.tau.ac.il

Mehdi Mekni

Institut national de la recherche scientifique (INRS), Centre Eau-Terre-Environnement
490, rue de la Couronne, Québec (Québec)
G1K 9A9, CANADA
Tel: 418-262-7601
E-mail: mmekni@gmail.com

Bernard Moulin

Université Laval, Département d'informatique et de génie logiciel
Pavillon Pouliot, 1065, rue de la Médecine
Québec, QC G1V 0A6, Canada
Tel: (+1) 418-656-5580
Fax: (+1) 418-656-2324
E-mail: bernard.moulin@ift.ulaval.ca

Paris Sebastien

Graphics Vision Visualization (GV2)
Trinity College, Dublin, Ireland
E-mail: parissebastien@free.fr

Scott Heckbert

Institute for Sustainable Solutions
Portland State University
Portland, OR, USA. 97201
E-mail: scott.heckbert@gmail.com

Ian Bishop

Department of Geomatics
University of Melbourne
Victoria, Australia 3010
E-mail: idbishop@unimelb.edu.au

Bas Straatman

Energy and Environmental Systems Group, ISEEE, University of Calgary
2500 University Dr NW, Calgary, AB

T2N 1N4, CANADA
Phone: (403) 210-7621
Fax: (403) 210-3894
E-mail: b.straatman@mun.ca

Danielle J. Marceau

Department of Geomatics Engineering
University of Calgary, Calgary, Alberta, Canada
2500 University Dr NW, Calgary, AB
T2N 1N4, CANADA
Tel: 403-220-5314
E-mail: dmarceau@ucalgary.ca

Roger White

Department of Geography
Memorial University of Newfoundland
St. John's, NL A1B 3X9
Canada
E-mail: roger@mun.ca

Harutyun Shahumyan

Urban Institute Ireland
University College Dublin, Richview Campus Dublin 4 Ireland
E-mail: harutyun.shahumyan@ucd.ie

Inge Uljee

VITO - Flemish Institute for Technological Research Boeretang 200
2400 Mol
Belgium
E-mail: inge.uljee@vito.be

Andrew Crooks

Department of Computational Social Science
Krasnow Institute for Advanced Study
379 Research 1 Building,
MS 6B2, George Mason University,
Fairfax, VA 22030, USA
Tel: (001)(703) 993-4640
Fax: (001)(703) 993-9290
E-mail: acrooks2@gmu.edu

Andrew Hudson-Smith

Centre for Advanced Spatial Analysis
University College London
1st Floor - 90 Tottenham Court Road
London, W1T 4TJ, UK
Tel: (44) (0)20 3108 3877
Fax (44) (0)20 3108 3258
E-mail: asmith@geog.ucl.ac.uk

Ateen Patel

Centre for Advanced Spatial Analysis

University College London
1st Floor - 90 Tottenham Court Road
London, W1T 4TJ, UK
Tel: (44) (0)20 3108 3877
Fax (44) (0)20 3108 3258
E-mail: ateen.patel@ucl.ac.uk

Erez Hatna

Centre for Advanced Spatial Analysis (CASA)
University College London (UCL)
90 Tottenham Court Road
London W1T 4TJ, UK
Email erezh51@gmail.com

Itzhak Benenson

Department of Geography and Human Environment
Tel Aviv University, Israel
Ramat Aviv, 69978, Tel Aviv
+97236409896
E-mail: bennya@post.tau.ac.il

Dr. Antoni Moore

School of Surveying
University of Otago
PO Box 56, Dunedin
New Zealand
Tel: 0064 3 4797589
Fax: 0064 3 4797586
E-mail: tony.moore@otago.ac.nz

CHAPTER 1

Challenges and Perspectives in Geosimulation

Danielle J. Marceau[1,*] and Itzhak Benenson[2]

[1]Department of Geomatics Engineering, University of Calgary, Calgary, Alberta, Canada and [2]Department of Geography, Tel-Aviv University, Tel-Aviv, Israel

Abstract: Geosimulation is a rapidly growing field of scientific investigation that involves the use of object-based spatially-explicit models in order to understand, through simulation, the dynamics of complex, adaptive human-driven geographic systems. Grown on cellular automata and multi-agent systems, it is progressively evolving towards an integrated framework in which physical objects of the environment and animated agents, especially humans, are investigated in interaction with each other. Simulating humans and other animated agents acting and interacting within an evolving environment raises conceptual and methodological challenges. The first one is the adequate representation of geographic space through meaningful spatially located objects. The implementation of objects' flexible geometry, neighborhood relations, rules of objects movement and relocation, all in 3D virtual space and at different scales are among the recent advances that we discuss in the paper. The second challenge is validation of the Geosimulation models. This paper presents several recent qualitative and quantitative techniques, including the pattern-oriented approach that aims at reproducing patterns that can be observed in reality. It is followed by an overview of the content of the book in respect to the concepts and methods that address these two challenges. We conclude that collaboration among researchers for sharing datasets and comparing models is a critical step towards the credibility and operability of Geosimulation models.

INTRODUCTION

The term *Geosimulation* has been popularized through the publication in 2004 of the book entitled: *Geosimulation: Automata-based modeling of urban phenomena* written by Itzhak Benenson and Paul Torrens [1]. These authors define Geosimulation as object-based spatially-explicit modeling of dynamic systems. To develop and investigate these models, Geosimulation employs the advanced methods of Complex Systems theory, Geographic Information Science, Object-Oriented Programming and Geovisualization.

Compared to traditional models of chemical or ecological systems that describe the temporal dynamics of a spatial system in terms of state variables representing the aggregate or spatially distributed properties of system's components [2], Geosimulation deals with discrete entities representing interacting animate and non-animate real-world agents and objects, such as pedestrians, drivers, land-users and developers, who act and interact within the space consisting of land parcels, buildings, roads, agriculture lots and open spaces. Taking into account instantaneous and delayed interactions between agents and objects of the same and different types, feedbacks, hierarchy of relationships, and heterogeneity of the objects' properties and behaviors, Geosimulation aims at understanding and reproducing the dynamics of the real-world complex, adaptive spatial systems, which dynamics is characterized by bifurcations, emergence, resilience, path dependence, and criticality. In Geosimulation models, space and time are defined explicitly; entities are situated in an environment and evolve through time. Following a bottom-up approach, Geosimulation considers the dynamics of a geographic system as a phenomenon emerging from the actions and interactions between multiple entities that compose it. Appreciating the complexity of the real-world and, usually, human-driven geographic systems, Geosimulation explores possible future paths a system can take through the testing of alternative scenarios.

Based on the superiority of non-animated objects or animated agents, Benenson and Torrens [1] consider two major classes of Geosimulation models, namely Cellular Automata (CA) and Multi-Agent Systems (MAS).

In their standard formalism, CA models represent geographic space by the regular 2D lattice of cells, each characterized by a finite set of states. The state of each cell is updated recurrently, according to transition rules that

***Address correspondence to Danielle J. Marceau:** Department of Geomatics Engineering, University of Calgary, Calgary, Alberta, Canada, 2500 University Dr NW, Calgary, AB, T2N 1N4, Canada; Tel: 403-220-5314; Email: dmarceau@ucalgary.ca

account for the cell's previous state and for the states of the neighboring cells [3]. When applied in a geographic context, standard CA consider raster representation of space and describe the dynamics of the cells' properties. While the raster-based partition dominates, the popularity of the irregular partitions is increasingly growing [4-8].

Numerous applications confirm that despite their conceptual simplicity, CA can realistically reproduce patterns of the land-use dynamics of cities and regions [9]. Neighborhood-dependent transition rules capture and sum up essential information governing local socio-economic-ecological interactions between the population group and modes of development. In addition, geographic CA go beyond the local interactions and include mid and even long-range effects as the influence of roads and other networks, as well as the central cities [10-11].

Over the last two decades, CA have been increasingly used to simulate a wide range of spatio-temporal phenomena from fire propagation [12-13] to vehicle and pedestrian flows [14-15]. The dominant field of their application in geography, however, is land-use/land-cover dynamics at the urban and regional levels [16-24]. The potential of CA models for ecological impact assessment and land-use and social planning is widely recognized and makes them a basis of spatial decision-support systems [10].

MAS aim at capturing collective behavior of human decision-makers within the environment that supports their activities. Agents "behave" and, compared to immobile spatial units of the CA models, can move within the environment in order to fulfill their objectives. Agents can communicate, negotiate, and learn; they can remember their past, make plans for the future and adapt their behavior to the system's changes [25-27]. MAS are largely used to simulate the interactions of social actors in various contexts, including urban dynamics [28-30], vehicle and pedestrian traffic [31-33], spatial planning [34], economic markets [35], land-use changes [36-37], natural resource management [38, 25], and human-wildlife interactions [39-41].

When applied in ecology, MAS are often called Individual-Based Models (IBM). IBMs are widely employed to study the spatio-temporal dynamics of plant and animal population and communities [42], including animal movement, schooling, species dispersal, and bioenergetics [43-44]. MAS/IBM are appropriate tools for studying the intricate aspects of biological and socio-economic systems consisting of heterogeneous individuals that interact in space at local and intermediate scales. This heterogeneity and interactions play a critical role in the dynamics of social and ecological systems [45, 42]. [46] and [47-48] have lately suggested a standard framework to facilitate the description and construction of MAS/IBM models and comparison of their results. Furthermore, the standard ODD (Overview, Design concepts and Details) protocol for representing simulation models in which human agents are involved, as recently proposed by [47-48], enables a formal description of each aspect of the designing and implementation of the model, thus facilitating model review, comparison, and replication. We expect this protocol to be largely accepted. The paper of [49] that compares the respective merits and flaws of three agent-based social simulation models using ODD, is one of the first attempts in this direction.

[1] and [50] have further proposed to consider CA and MAS as specific implementations of the general Geographic Automata Systems (GAS) framework. GAS automata directly represent real-world entities, characterized by states, location, and automation rules that depend on their own state and the state of the automata they are related to. The geographic automata can represent immobile entities such as land parcels and buildings or mobile agents such as animals and humans. Geo-referencing – direct, according to coordinates, and indirect, by pointing to the other objects - is employed for establishing neighborhood and other spatial relationships in GAS. GAS transition rules include the rules of movements and account not only for the dynamics of the agents' properties, but also for the time variation of relationships between them. Combining CA and MAS, GAS offers the possibility of coupling human and natural systems to study the evolving collective of interacting human agents within the varying environment [34,46,51].

Despite the fact that Geosimulation is still in its infancy, the constantly growing scientific literature on the subject confirms that this avenue of research has a major influence in many disciplines. The objective of this introduction is to critically review two challenges of today: a) the need for an adequate representation of the geographic objects and geographic space, and b) the need for appropriate techniques to evaluate Geosimulation models and relate the model results to the complex and dynamic geographic reality. The book's contributions that address these issues are then introduced.

ADEQUATE REPRESENTATION OF THE GEOGRAPHIC SPACE

Geosimulation, by definition, demands direct and adequate representation of the geographic entities. First, this is necessary for assuring the realism of the simulations, *i.e.* for capturing the attributes of the agents and objects that comprise the geographic system, and providing a description of their behavior and interactions. Second, an adequate representation of the geographic space ensures that the model is designed and implemented at the proper spatial and temporal scale and can be adequately compared to the patterns that are observed in reality.

From Lattice of Cells to Virtual Environments of Geographic Entities

Scientists who early recognized the potential of high-resolution CA modeling in geography also acknowledged the necessity to relax all aspects of the CA formalism. [52] and [53] were among the first to claim that geographic CA should allow for neighborhoods which extent and shape vary in space, transitions rules that depend on location, and non-local dependencies and constraints. This conceptual framework is being gradually implemented and recent CA models widely employ extended neighborhoods and account for distant infrastructure elements when formulating the transition rules [54, 11, 18, 17, 55-56]. Rigid lattice partitions and uniform neighborhoods are often substituted by irregular tessellations that reflect the inherent structure of the geographic space, such as as Voronoi diagrams [57-58, 30], Delaunay triangles [5], and spatial graphs [7].

Definition of the neighborhood relationships via topological contiguity between the units of irregular spatial partition is only the first step in representing the real-world relationships between the geographic entities [10]. The next step is in extending the CA formalism towards flexible neighborhoods and transition rules that account for the location of streets, 2D- and 3D-visibility, and, more generally, human spatial cognition [59, 32, 50]. Spatially-explicit CA that employ neighborhoods that are based on irregular but still rigid location of buildings, cadastral parcels, or forest stands [29, 60-61] are still insufficient for this purpose. The CA models of [5] and [7-8] provides a next step; they are based on a land partition in which parcels can undergo geometric transformation, defined by the influence and evolution of their neighbors. Note that evolving spatial units demand neighborhood relations that are defined contextually and not only on the basis of distance or topology [50, 46], as it is implemented in [7-8].

Two full-featured implementations of the GAS framework were recently proposed. [62] consider vector agents, *i.e.* the geographic automata that can change their geometry. In addition to the ability to move, a vector agent is able to modify its shape, and its transition rules include spatial growth and contraction. [63] further consider neighborhood relationships that vary during simulation time, in respect to the land-use changes in the model. In addition their model allows for relationships between spatially non-adjacent cells. Compared to the rigid objects and relationships of the basic CA models, the GAS systems that allow for the dynamically varying objects and relationships indicate a critical step toward an adequate representation and, thus, a better understanding of the complex human-environmental systems [64].

Advances in 3D Geovisualization considerably improve the realism of the GAS simulations. As an example, [5] introduced a 3D view of a city into a CA model while [65] used the height of the building as a leading factor of the transition rules. Implicit representation of the 3D objects in the geosimulation models are now substituted by an explicit one, when the model is developed within a 3D virtual geographic environment, where agents navigate and react to changes [66]. Comparing to the 2D representation, 3D virtual environments considerably improve the "sense of realism" which is particularly useful when simulating the interactions between human agents and their environment [67]. This tendency is strongly supported by the recent development of the simulation software, which quickly advance towards realistic and spatially explicit 3D Geosimulation environments [68-69].

From Scale Dependency to Multi-Scale Representation

While numerous studies have emphasized the critical role of scale in spatial analysis and modeling [70-73], relatively few attempts have been made to systematically investigate the importance of spatial and temporal scale in Geosimulation. However, spatio-temporal resolution is at the core of decisions made during model design, implementation, calibration and result analysis.

Modeling land-use dynamics with raster-based CA illustrates the effects of the cell size, neighborhood configuration, and time resolution on the ways the researcher establishes model transition rules and, thus, on the spatial patterns generated by the model [74-90]. When based on the neighborhood relationships that explicitly reflect the contextual relations between the real-world objects [7-8], the geosimulation model resolves the problem of possible qualitative changes of transition rules in respect to the change of the model spatio-temporal scale [81]. Non-uniform view of the relationships might demand accounting for spatio-temporal variations of transition rules and, as [82] have demonstrated, this variation can be estimated using the historical records of land-use changes.

To implement the idea that the relationship between distant phenomena can be considered at lower spatial resolution than that between the closer ones, [83] proposed the concept of multi-resolution CA. The same geographic area is represented in these CA by several hierarchically organized partitions of the increasing cell size. The influence of the area that is adjacent to a given cell is considered, as in the basic CA, at the highest possible resolution, while the influence of the distant phenomena is represented by the cells taken from the lower-resolution partitions. The hierarchical spatial structure logarithmically reduces the number of interactions that should be considered for the full representation of the phenomena, including local and distant influence. The idea of a hierarchically structured CA was further generalized by [84] who consider a multi-layer CA in which the neighbors of a cell are acquired from the same and the other layers, the latter of arbitrarily coarser or finer resolutions. These multi-layer representations relate between the micro-interactions of model entities and the macro-level structures, where the global patterns are expected to emerge [85].

The influence of spatial scale on the agents' behavioral rules in MAS has been little explored so far. An exception is provided by [86] who investigated the scale dependence of an agent-based model of the land-cover changes by varying the resolution of the landscape on which agents make decisions. Demonstrating considerable impact of scale on model performance, design, implementation and application, the authors recommend running models at a variety of spatial scales and exploring the range of possible model outcomes. Scale in MAS influences not only the transition rules, but, also, scenarios development and the ability to communicate with the planners and decision-makers. [87] discuss multi-scale scenarios as a mean to engage stakeholders in understanding the driving forces and possible responses of a complex system.

To understand the potential of MAS for studying multi-scale social systems, [88] and [89] propose to distinguish between agents that act at different hierarchical levels and then include the knowledge on spatial hierarchy into agents' behavioral rules. Emergence of the collective responses in such a case can be illustrated by [34] model of participatory planning, where each agent constructs an individual image of the spatial environment and, simultaneously, participates in collective decision making on land-use allocation. An even further-going attempt is made by [90] who advocates that emergence and adaptation, characteristic of the complex systems, demand a dynamic reassessment of system's hierarchy and scale. Geosimulation yet lacks meaningful examples of such a dynamic definition of scale.

An adequate representation of space and spatial relationships is a precondition of the correspondence between the Geosimulation model and reality.

BRIDGING THE GAP BETWEEN THE MODEL AND REALITY

Adequate calibration and validation is critical for ensuring the credibility and the usefulness of the whole geosimulation exercise [91]. Calibration aims at establishing a good fit between the real-world dataset and the model outcomes and is considered successful if estimates of the model parameters can be obtained and the dataset dynamics can be imitated. As such, successful calibration guarantees goodness of fit, but does not warrant that the model rules reflect the real world dependencies. Model validation, *i.e.*, the proof that the model's structure, rules and the rules' parameters reflect the real-world phenomena is commonly considered as a bold challenge [92].

A commonly accepted view is that a model is valid if it is consistent with the intended application according to three dimensions: conceptually, operationally, and numerically [91, 93]. The numerical fit is most easily formalized. The difference between the geosimulation model and the real-world dynamics is often based on aggregate measures of similarity between two maps, such as the Kappa index [94], and [95-98] generalized and deepened this approach by proposing several advanced techniques to aggregate measuring the uncertainty associated to the land-use changes.

Conceptual validation is essentially more controversial. Extremists argue that the whole issue is inapplicable to complex human systems and a model can never represent the whole 'truth', *i.e.*, provide complete and sufficiently accurate representation of the real world in all its complexity [99]. The maximum demand can be the model's goodness of fit within a limited range of parameters, and, thus, the researcher's goal becomes to establish the limits of model's credibility [100]. A step towards an evaluation of this kind is conducting an analysis of a model's sensitivity to the parameters and initial conditions. An extended example is provided by [101] who conducted a systematic evaluation of the CLUE land-use model at different spatial scales. The authors conclude that the scale at which they evaluate the model strongly influences its predictive power: the latter improves exponentially with a decrease in spatial resolution. However, conducting a systematic sensitivity analysis for a multi-parametric system is a computationally very intensive task, and [102] propose to apply computer agents for this service. [99] further advocates that taking into account theoretical, empirical, parametrical, and temporal uncertainties, the researcher should develop the methods for evaluating the quality of models, rather than attempting to validate them. Developing this idea, [103] also suggests using the term *evaluation* instead of validation, to encompass the entire process of calibration and verification of the complex system models. The popular SLEUTH model can be considered as an extreme example of this approach: possible land-use transitions in this CA model are intentionally limited to five simple rules, which parameters are chosen to ensure the best fit to the real data. In case the model outcome diverges from the data, the parameters are updated [104-106].

Appreciating the problems of validation for complex and open spatial systems, the majority of researchers do not attempt to provide universal models that aim at describing the wide spectrum of geographic phenomena; they rather prefer to specify the domains in the parameters' space where the model performs better or worse. [107] apply this view for validating agent-based models of land-use change and define "invariant regions" where the dynamics of the land-use type is almost certain, and "variant regions" where the land-use dynamics depends on a particular series of events and is thus path dependent. The authors use this approach to determine the predictive accuracy of a model at simulating land-use patterns that are consistent with real-world processes. However, the mainstream in geosimulation avoids the issue of the validation that demands developing models based on several initial datasets and, then, validating them with the datasets that were not employed in the model development. Due to the high cost of the database construction, typical geosimulation model is limited to the single set of the land-use maps. Our view is that sharing of the land-use datasets and research cooperation in investigating models' abilities and limitations are critical for making geosimulation operational. The paper of [108] is the first step in this direction.

Pattern-oriented modeling (POM) can be considered as a reasonable compromise between the validation and evaluation views. A pattern is defined as a clearly identifiable structure in the data extracted from nature that goes beyond random variation and indicates the existence of an underlying process [109]. POM considers the model as successful if it is able to repeat many distinct patterns observed in nature [110]. The POM framework is defined as a computer protocol and is applied in two steps. First, a set of primary patterns are extracted from the field observations; second the patterns resulting from the simulations are compared to them. Both protocols have been applied by [111] to evaluate an agent-based model of whale-watching excursions in the St. Lawrence estuary. Further in this direction, [112] propose to apply POM to compare models with the Turing-like "interrogation test" [113] In order to understand the model's predictive ability, they suggest using the outcomes of other models as a surrogate of the real-world patterns. The calibration and validation techniques can then be tested versus artificially generated patterns, given the model rules are not known to the investigator. Different from the real-world datasets, the surrogate ones are potentially unlimited and cover the entire spectrum of parameters.

Another promising approach to models' qualitative evaluation is based on expert estimates. [114] evaluated an agent-based model designed to simulate the land-use impact of municipal planning policies and stakeholders' decision, by presenting the simulation outcomes to a group of experts and studying their judgment on the credibility of the model. The experts' feedbacks provided valuable view on the limitations of the model and its usefulness for understanding actors' interactions during a land-use re-designation process. Further in that direction, the *companion modeling* approach demands regular interactions between stakeholders and modelers to ensure that the design and implementation of the geosimulation model closely match the experts' opinion on the real-world system. Several examples of companion modeling can be found in integrated natural resource management [115].

To conclude, establishing correspondence between the models and reality is the major challenge Geosimulation is facing today. Further developments and, especially, case-studies, are critical for making geosimulation models operational. We need a variety of models properly scaled in space and in time, which objects and agents adequately representing real-world infrastructure entities and human decision-makers. Adequate representation of the rules of human decision-making is even a bolder challenge. We expect more studies that would critically assess the knowledge that can be gained through the geosimulation exercise versus observations of the real-world dynamics, thus specifying the perspectives and limits of Geosimulation as a tool for investigating geographical reality.

OVERVIEW OF THE BOOK CONTENT

The book presents novel concepts and methods in the field of Geosimulation. These include new approaches to representation of the geographic space through the use of a variable grid CA (Chapter 2), vector agent-based simulation (Chapter 3), 3D multi-agent system (Chapter 4), and virtual geographic environments (Chapter 5). The calibration techniques for agent-based models are presented in Chapter 6. Chapters 7 and 8 demonstrate the potential of agent-based modeling to increase knowledge and enrich theories of urban residential dynamics and economic markets, respectively.

In the Chapter 2, Roger White, Harut Shahumyan and Inge Ulje combine processes operating at different spatial scales in a land-use CA through simultaneous representation of the land-use map at several resolutions. Conventional CA models are based on a single spatial partition, and cell's neighborhood consists of several rings of adjacent cells. Processes that operate at a larger scale are either ignored or taken into account through the linkage with another model (*e.g.* transportation, economic, or population growth). The authors propose to aggregate the information about the distant influence based on the larger cells, which size increases with the distant from the target cell. This variable grid approach is tested in simulating land-use patterns and regional activity levels in Dublin, Ireland. The authors compared their approach to the traditional approach of two linked, micro- and macro-scale, models. The results indicate that the model that is based on variable grid provides better prediction of land-use patterns and regional activities than the pair of linked models. In addition, the structure of the multi-layer model is simpler and allows a better representation of the dynamics of the phenomenon being investigated.

As we discussed above, CA grid of identical cells facilitates computations but does not represent meaningful entities of the real world. To overcome this limitation, Antoni Moore (Chapter 3) proposes to extend Geographic Automata System to the system Geographical Vector Agents (GVA). A GVA is able to modify its own geometry while interacting with the other agents. The author implements GVA for investigation of the classic Von Thunen’s theory. He implements three types of agents, some establishing and modifying the land parcels and some residing in these parcels and defining their land-uses. Based on likelihood estimates of prices, transportation cost and other objects' and agents' properties and interactions, the GVA model adequately recreates von Thunen's rings of agriculture land-uses around the urban core.

In Chapter 4, Andrew Crooks, Andrew Hudson-Smith and Ateen Patel present achievements of the 3D agent-based urban modeling. They underline that the recent enhancements of virtual globe technologies and increased availability of fine-scale data have created an unprecedented boost to 3D urban representation and modeling. Animating agents within a 3D virtual city enables direct accounting for an individual's sense of place and, thus, facilitates a wide range of experiments aimed at investigating complex behaviors that mimic humans' movements, interactions, and adaptation to varying urban conditions. The authors present different techniques of developing 3D agent-based models, review commercial open source game engines, and call upon coupling ABM and a 3D visualization software by data transfer. They provide examples of such linkage, while emphasizing that the adequate and non-trivial information on agents' behavior is crucial for conveying 3D geosimulation model beyond just sophisticated visual representation.

Pursuing a similar objective, Mehdi Mekni, Bernard Moulin and Sébastien Paris present, in Chapter 5, the Multi-Agent Geosimulation framework that allows simulating large number of interacting heterogeneous agents within a Virtual Geographic Environment (VGE). The authors present a novel approach for the automatic generation of a geometrically accurate VGE that enhances spatial semantic information and thus supports agents’ knowledge about their environment and situated reasoning. The Conceptual Graphs formalism is proposed for precise representation of the semantic information to be used by the agents. The proposed framework is further applied to solve a path planning problem within heterogeneous urban environment.

As presented above, the calibrating of the ABM is a multi-faceted and challenging issue. In Chapter 6, Scott Heckbert and Ian Bishop present three methods of calibration and illustrate them with three case studies. The first study estimates hunters' preference functions for the ABM model of hunting in the forest area. The second describes participatory techniques that aim at disclosing context-dependent information provided by the experts in high-resolution urban planning. The authors propose integration of the modern technologies, such as GPS and 3D visualizations of landscapes, for enhancing the perception and response of participants and thus assisting their decision making. The last case study considers experimental economic games, in which participants mimic real-world trading behavior of economic agents. An application regards agricultural agents who choose the quantity of fertilizer to buy or sell given their calculated profitability. The novel experimental interface developed by the authors and used by the participants of an economic game in laboratory is directly integrated with the ABM, ensuring consistency in terms of equations and parameters in the models.

Urban theory states that the local heterogeneity of the income residential distribution is low. Erez Hatna and Itzhak Benenson (Chapter 7) present several opposite examples and design an agent-based model to investigate the emergence and persistence of the heterogeneous residential patterns as those observed in nine Israeli cities. Their MAS model investigates the hypothesis that urban heterogeneity is a consequence of the householders' tolerance to the closely located neighbors of the lower economic status. The model city is open and the families of different income and of different tolerance to the poorer neighborhoods immigrate into the city, find a dwelling to settle, reside and emigrate. The analysis of the model reveals that a low fraction of tolerant families is sufficient for the emergence and persistence of heterogeneous residential patterns, thus illustrating the use of MAS for acquiring additional knowledge about the data-rich urban system and enhancing urban theory.

In a similar attempt to test hypotheses and contribute to existing economic theory Baas Straatman and Danielle Marceau present an agent-based model designed to investigate whether an economic market acts as an open system that continuously incorporates new adaptive characteristics and in this way increases its functional complexity. Two novel aspects are introduced in the study. First this is a *constructive* agent-based model that generates new components and relationships between them. Second, a measure of evolutionary activity is applied to evaluate the impact of adding new technologies in the simulated economic market. The model simulates an artificial market in which producer and consumer agents exchange products and perform activities. New technologies, developed as opportunities of invention and innovation appear during the interactions among the agents. The simulations reveal an increase in the number of agents and diversity of products and technologies, which in turn generates a higher level of activities in the system, confirming the hypothesis that an economic market is an open-ended system.

REFERENCES

[1] I. Benenson and P. Torrens, "Geosimulation: Automata based modeling of urban phenomena", Wiley, 2004.

[2] H. Haken, "Synergetics: An Introduction", Berlin, Springer, 1983.

[3] S. Wolfram, "Cellular automata as models of complexity", Nature, vol. 311, pp. 419-424, 1984.

[4] R. Hegselmann and A. Flache, "Understanding complex social dynamics: a plea for cellular automata based modeling", Journal of Artificial Societies and Social Simulation vol. 1, no. 3, 1998. [Online]

[5] F. Semboloni, "The growth of an urban cluster into a dynamic self-modifying spatial pattern", Environment and Planning B, vol. 27, pp. 549-564, 2000.

[6] D. O'Sullivan, "Exploring spatial process dynamics using irregular cellular automaton models", Geographical Analysis, vol. 33, pp. 1-18, 2001.

[7] N. Moreno, F. Wang, and D.J. Marceau, "Implementation of a dynamic neighbourhood in a land-use vector-based geographic cellular automata model", Computers, Environment and Urban Systems, vol. 33, pp. 44-54, 2009.

[8] N. Moreno, F. Wang, and D.J. Marceau, "A geographic object-based approach in cellular automata modeling", Photogrammetric Engineering and Remote Sensing, vol. 76, pp. 183-191, 2010.

[9] M. Batty, "Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals", MIT Press, 2005.

[10] R. White, and G. Engelen, "High-resolution integrated modelling of the spatial dynamics of urban and regional systems", Computers, Environment and Urban Systems, vol. 24, pp. 383-400, 2000.

[11] G. Engelen, R. White, and I. Uljee, "The MURBANDY and MOLAND models for Dublin", European Commission Joint Research Center, Ispra, Italy, Dublin, ERA-Maptec, 2002.

[12] S.G. Berjak and J. W. Hearne, "An improved cellular automaton model for simulating fire in a spatially heterogeneous Savanna system", Ecological Modelling, vol. 148, pp. 133-151, 2002.

[13] S. Yassemi, S. Dragicevic, and M. Schmidt, "Design and implementation of an integrated GIS-Based cellular automata model to characterize forest fire behaviour", Ecological Modelling, vol. 210, pp. 71-84, 2008.

[14] J. Wahle, L. Neubert, J. Esser, and M. Schreckenberg, "A cellular automaton traffic flow model for online simulation of traffic", Parallel Computing, vol. 27, pp. 719–735, 2001.

[15] T. Sun and J. Wang, "A traffic cellular automata model based on road network grids and its spatial and temporal resolution's influences on simulation", Simulation Modelling Practice and Theory, vol. 15, pp. 864-878, 2007.

[16] M. Batty, M., Y. Xie, and Z. Sun, "Modeling urban dynamics through GIS-based cellular automata", Computers, Environment and Urban Systems, vol. 23, pp. 205-233, 1999.

[17] X. Li and A. G.-O. Yeh, "Neural-network-based cellular automata for simulating multiple land use changes using GIS", International Journal of Geographical Information Science, vol. 16, pp. 323-343, 2002.

[18] F. Wu, "Calibration of stochastic cellular automata: the application to rural-urban land conversion", International Journal of Geographical Information Science, vol. 16, pp. 795-818, 2002.

[19] C.M. Almeida, J. M. Gleriani, E.F. Castejon and B.S. Soares-Filho, "Using neural networks and cellular automata for modelling intra-urban land-use dynamics", International Journal of Geographical Information Science, vol. 22, no. 9, pp. 943-963, 2008.

[20] K.H. Lau and B. H. Kam, "A cellular automata model for urban land-use simulation", Environment and Planning B, vol. 32, pp. 247-263, 2005.

[21] A. Ménard and D.J. Marceau, "Simulating the impact of forest management scenarios in an agricultural landscape of Southern Quebec, Canada, using a geographic cellular automata", Landscape and Urban Planning, vol. 79, pp. 253-265, 2007.

[22] C. He, N. Okada, Q. Zhang, P. Shi, and J. Li, "Modelling dynamic urban expansion processes incorporating a potential model with cellular automata", Landscape and Urban Planning, vol. 86, pp. 79-91, 2008.

[23] I. Santé, A. M. Garcia, D. Miranda, and R. Crecente, "Cellular automata models for the simulation of real-world urban processes: a review and analysis", Landscape and Urban Planning, vol. 96, pp. 108-122, 2010.

[24] Z. Shen, M. Kawakami, and I. Kawamura, "Geosimulation model using geographic automata for simulating land-use patterns in urban partitions", Environment and Planning B, vol. 36, pp. 802-823, 2009.

[25] F. Bousquet and C. Le Page, "Multi-agent simulations and ecosystem management: a review", Ecological Modelling, vol. 176, pp. 313-332, 2004.

[26] D.J. Marceau, "What can be learned from multi-agent systems"? in Monitoring, Simulation and Management of Visitor Landscapes, R. Gimblett, ed., University of Arizona Press, 2008, pp. 411-424.

[27] M. Bithell, J. Brasington, and K. Richards, "Discrete-element, individual-based and agent-based models: Tools for interdisciplinary enquiry in Geography"? Geoforum, vol.39, pp. 625-642, 2008.

[28] I. Benenson, "Multi-agent simulations of residential dynamics in the city", Computers, Environment and Urban Systems, vol. 22, no. 1, pp.25-42, 1998.

[29] I. Benenson, I. Omer, and E. Hatna, "Entity-based modeling of urban residential dynamics: the case of Yaffo, Tel-Aviv", Environment and Planning B, vol. 29, pp. 491-512, 2002.

[30] J. Jackson, B. Forest, and R. Sengupta, "Agent-based simulation of urban residential dynamics and land rent change in a gentrifying area of Boston", Transactions in GIS, vol. 12, pp. 475-491, 2008.

[31] D. Helbing, P. Molnar, I.J. Farkas, and K. Bolay, "Self-organizing pedestrian movement", Environment and Planning B, vol. 28, pp. 361-383, 2001.

[32] M. Batty, J. Desyllas and E. Duxbury, Safety in numbers? "Modelling crowds and designing control for the Notting Hill Carnival", Urban Studies, vol. 40, no. 8, pp.1573–1590, 2003.

[33] I. Benenson, K. Martens, and S. Birfir, "PARKAGENT: an agent-based model for parking in the city", Computers, Environment and Urban Systems, vol. 32, pp. 431–439, 2008.

[34] A. Ligtenberg A. K. Bregt, and R. Van Lammeren, "Multi-actor-based land use modelling: spatial planning using agents", Landscape and Urban Planning, vol. 56, pp. 21-33, 2001.

[35] L. Tesfatsion, "Agent-based computational economics: modeling economies as complex adaptive systems", Information Sciences, vol. 149, pp. 263-269, 2003.

[36] K. Lim, P. J. Deadman, E. Moran, E. Bondizio, and S. McCracken, "Agent-based simulations of household decision making and land use change near Altamira, Brazil", in Integrating geographic information systems and agent-based

modeling techniques for simulating social and ecological processes. H. R. Gimblett, ed., Oxford University Press, 2002, pp. 277-310.

[37] D.C. Parker, S. M. Manson, M.A. Janssen, M.J. Hoffmann, and P.J. Deadman, "Multi-agent systems for the simulation of land-use and land-cover change: a review", Annals of the Association of American Geographers, vol. 93, pp. 314-337, 2003.

[38] H.R. Gimblett, C. A. Roberts, T.C. Daniel *et al.*, "An intelligent agent-based model for simulating and evaluating river trip scenarios along the Colorado River in Grand Canyon National Park", in Integrating Geographic Information Systems and Agent-Based Modeling Techniques, H. R. Gimblett, ed., Oxford University Press, 2002, pp. 245-275.

[39] M. Anwar, C. Jeanneret, L. Parrott, and D.J. Marceau, "Conceptualization and implementation of a multi-agent model to simulate whale-watching activities in the St. Lawrence estuary in Quebec, Canada", Environmental Modeling and Software, vol. 22, pp. 1775-1787, 2007.

[40] P.D. Grosman, J. A. G. Jaeger, P.M. Biron, C. Dussault, and J.P. Ouellet, "Reducing moose-vehicle collisions through salt pool removal and displacement: an agent-based modeling approach", Ecology and Society, vol. 14, pp. 2-17, Online, 2009.

[41] M. Musiani, M. Anwar, G. McDermid, M. Hebblewhite, and D.J. Marceau, "How humans shape wolf behavior in Banff and Kootenay national parks, Canada", Ecological Modelling, vol. 221, pp. 2374-2387, 2010.

[42] V. Grimm and S.F. Railsback, "Individual-based Modeling and Ecology", Princeton University Press, 2005.

[43] B. Breckling, U. Middelhoff, and H. Reuter, "Individual-based models as tools for ecological theory and application: Understanding the emergence of organizational properties in ecological systems", Ecological Modelling, vol. 194, pp. 102-113, 2006.

[44] D. A. Bennett and W. Tang, "Modelling adaptive, spatially aware, and mobile agents: Elk migration in Yellowstone", International Journal of Geographical Information Science, vol. 20, no. 9, pp. 1039-1066, 2006.

[45] O.P. Judson, "The rise of the individual-based model in ecology", TREE, vol. 9, pp. 9-14, 1994.

[46] I. Benenson, S. Aronovich, and S. Noam, "Let's talk objects: generic methodology for urban high-resolution simulation", Computers, Environment and Urban Systems, vol. 29, pp. 425–453, 2005.

[47] V. Grimm, U. Berger, F. Bastiansen *et al.*, "A standard protocol for describing individual-based and agent-based models", Ecological Modelling vol. 198, pp. 115-126, 2006.

[48] V. Grimm, D.L. DeAngelis, J.G. Polhill, J. Giske, and S.F. Railsback, "The ODD protocol for describing individual-based and agent-based models: A first update", Ecological Modelling, vol. 221, pp. 2760-2768, 2010.

[49] J.G. Polhill, D. C. Parker, D.G. Brown, and V. Grimm, "Using the ODD protocol for describing three agent-based social simulation models of land-use change", Journal of Artificial Societies and Social Simulation, vol. 11, no. 2, 2008. [Online].

[50] P. Torrens and I. Benenson, "Geographic automata systems", International Journal of Geographical Information Science, vol. 19, pp. 385-412, 2005.

[51] M. Monticino, M. Acevedo, B. Callicot, T. Cogdill, and C. Lindquist, "Coupled human and natural systems: a multi-agent approach", Environmental Modelling and Software, vol. 22, pp. 656-663, 2006.

[52] M. Takeyama and H. Couclelis, "Map dynamics: integrating cellular automata and GIS through Geo-Algebr", International Journal of Geographical Information Science, vol. 11, pp. 73-91, 1997.

[53] M. Batty, "Urban evolution on the desktop: simulation with the use of extended cellular automata", Environment and Planning A, vol. 30, no. 11, pp. 1943-1967, 1998.

[54] K.C. Clarke, S. Hoppen, and L.J. Gaydos, "A self-modifying cellular automata model of historical urbanization in the San Francisco Bay area", Environment and Planning B, vol. 24, pp. 247-261, 1997.

[55] B.C. Pijanowski, D. G. Brown, B.A. Shellito, and G.A. Manik, "Using neural networks and GIS to forecast land use changes: a land transformation model", Computers, Environment and Urban Systems, vol. 26, pp. 553-575, 2002.

[56] S. Fang, G. Z. Gertner, Z. Sun, and A.A. Anderson, "The impact of interactions in spatial simulation of the dynamics of urban sprawl", Landscape and Urban Planning, vol. 73, pp. 294-306, 2005.

[57] W. Shi and M. Y. C. Pang, "Development of Voronoi-based cellular automata. An integrated dynamic model for geographical information systems", International Journal of Geographical Information Science, vol. 14, pp. 455-474, 2000.

[58] A. Flache and R. Hegselmann, "Do irregular grids make a difference? Relaxing the spatial regularity assumption in cellular models of social dynamics", Journal of Artificial Societies and Social Simulation, vol. 4, 23 pp., 2001.

[59] J. Portugali, "Self-Organization and the City", Berlin, Springer, 2000.

[60] D. Stevens, S. Dragicevic, and K. Rothley, "iCity: A GIS-CA modelling tool for urban planning and decision making", Environmental Modelling and Software, vol. 22, pp. 761-773, 2007.

[61] A.-H. Mathey, E. Krcmar, S. Dragicevic, and I. Vertinsky, "An object-oriented cellular automata model for forest planning problems", Ecological Modelling, vol. 212, pp. 359-371, 2008.

[62] Y. Hammam, A. Moore, and P. Whigham, “The dynamic geometry of geographical vector agents”, Computers, Environment and Urban Systems, vol. 31, pp. 502-519, 2007.

[63] S. Sarkar, K. A. Crews-Meyer, K.R. Young, C.D. Kelley, and A. Moffett, “A dynamic graph automata approach to modeling landscape change in the Andes and the Amazon”, Environment and Planning B, in press, 2010.

[64] D. A. Bennett and D. McGinnis, “Coupled and complex: Human-environment interaction in the Greater Yellowstone Ecosystem, USA”, Geoforum, vol. 39, pp. 833-845, 2008.

[65] L. Benguigui, D. Czamanski, and R. Roth, “Modeling cities in 3D: a cellular automaton approach”, Environment and Planning B, vol. 35, pp. 413-430, 2008.

[66] B. Moulin, W. Chaker, J. Perron, P. Pelletier, H. Hogan, and E. Gbei, “MAGS project: multi-agent geosimulation and crowd simulation”, in Proceedings of the COSIT' 2003 Conference, Ittengen, Switzerland, September 2003.

[67] D. Loiterton and I. Bishop, “Simulation, calibration and validation of recreational agents in an urban park environment”, in Monitoring, Simulation, and Management of Visitor Landscapes. Gimblett, H. R. and H. Skov-Petersen, eds, The University of Arizona Press, 2008, pp. 107-122.

[68] VISSIM, http://www.trafficgroup.com/3d/, visited 12/03/2020, 2010.

[69] Crowd Dynamics, http://www.crowddynamics.com/, visited 12/03/2020, 2010.

[70] S. Openshaw, “The modifiable areal unit problem”, GeoAbstracts, University of East Anglia, Norwich, 1984.

[71] A.S. Fotheringham and D. W. S. Wong, “The modifiable areal unit problem in multivariate statistical analysis”, Environment and Planning A, vol. 23, pp. 1025-1044, 1991.

[72] D.J. Marceau, “The scale issue in the social and natural sciences”, Canadian Journal of Remote Sensing, vol. 25, no. 4, pp. 347-356, 1999.

[73] K. Kok and A. Veldkamp, “Evaluating the impact of spatial scales on land use pattern analysis in Central America”, Agriculture, Ecosystems and Environment, vol. 85, pp. 205-221, 2001.

[74] G. D. Jenerette and J. Wu, “Analysis and simulation of land-use change in the central Arizona Phoenix region, USA”, Landscape Ecology, vol. 16, pp. 611-626, 2001.

[75] X. Liu and C. Andersson, “Assessing the impact of temporal dynamics on land-use change modeling”, Computers, Environment and Urban Systems, vol. 28, pp. 107-124, 2004.

[76] C.A. Jantz, and S. Goetz, “Analysis of scale dependencies in an urban land-use change model”, International Journal of Geographical Information Science, vol. 19, pp. 217-241, 2005.

[77] A. Ménard and D.J. Marceau, “Exploration of spatial scale sensitivity in geographic cellular automata”, Environment and Planning B., vol. 32, pp. 693-714, 2005.

[78] V. Kocabas, V. and S. Dragicevic, “Assessing cellular automata model behaviour using a sensitivity analysis approach”, Computers, Environment and Urban Systems, vol. 30, pp. 921-953, 2006.

[79] N. Samat, “Characterizing the scale sensitivity of the cellular automata simulated urban growth: a case study of the Seberang Perai Region, Penang Tate, Malaysia”, Computers, Environment and Urban Systems, vol. 30, pp. 905-920, 2006.

[80] Y. Pan, A. Roth, Z. Yu, and R. Doluschit, “The impact of variation in scale on the behavior of a cellular automata used for land-use change modeling”, Computers, Environment and Urban Systems, vol. 34, pp. 400-408, 2010.

[81] I. Benenson, “Warning! The scale of land-use CA is changing”, Computers, Environment and Urban Systems, vol. 31, pp. 107-113, 2007.

[82] J.-G. Hasbani, N. Wijesekara, and D.J. Marceau, “An interactive method to dynamically create transition rules in a land-use cellular automata model”, in: Cellular Automata, Intech Publisher, Accepted for publication, 2010.

[83] J. Van Vliet, R. White, and S. Dragicevic, “Modeling urban growth using a variable grid cellular automata”, Computers, Environment and Urban Systems, vol. 33, pp. 35-43, 2009.

[84] A.R. Kiester and K. Sahr, “Planar and spherical hierarchical, multi-resolution cellular automata”, Computers, Environment and Urban Systems, vol. 32, pp. 204-213, 2008.

[85] C. Ratzé, F. Gillet, J.P. Muller, and K. Stoffel, “Simulation modelling of ecological hierarchies in constructive dynamical systems”, Ecological Complexity, vol. 4, pp. 13-25, 2007.

[86] T.P. Evans and H. Kelley, “Multi-scale analysis of a household level agent-based model of landcover change”, Journal of Environmental Management, vol. 72, pp. 57-72, 2004.

[87] R. Biggs, C. Raudsepp-Hearne, C. Atkinson-Palombo, *et al.*, “Linking futures across scales: a dialog on multiscale scenarios”, Ecology and Society, vol. 12, 17 pp., 2007.

[88] A. Giret and A. Botti, “On the definition of meta-models for analysis of large-scale MAS”, in Second German Conference on Multi-agent system technologies, Erfurt, Germany, September 2004, Springer.

[89] E. Adam and R. Mandiau, “Roles and hierarchy in multi-agent organizations”, in 4th International Central and Eastern European Conference on Multi-Agent systems, Budapest, Hungary, September 2005, Springer.

[90] S.M. Manson, “Does scale exist? An epistemological scale continuum for complex human-environment systems”, Geoforum, vol. 39, pp. 776-788, 2008.
[91] E.J. Rykiel, “Testing ecological models: the meaning of validation”, Ecological Modelling, vol. 90, pp. 229-244, 1996.
[92] S.M. Manson, “Validation and verification of multi-agent systems”, in Complexity and Ecosystem Management. M. A. Janssen ed., Edward Elgar, 2000, pp. 63-74.
[93] M.A. Louie and K. M. Carley, “Balancing the criticisms: validating multi-agent models of social systems”, Simulation Modelling Practice and Theory, vol. 16, pp. 242-256, 2008.
[94] H. Visser, and T. de Nijs, “The Map Comparison Kit”, Environmental Modelling and Software, vol. 21, pp. 346-358, 2006.
[95] R.G. Pontius, “Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions”, Photogrammetric Engineering and Remote Sensing, vol. 68, pp. 1041-1049, 2002.
[96] R.G. Pontius, D. Huffaker, and K. Denman, “Useful techniques of validation for spatially explicit land-change models”, Ecological Modelling, vol. 179, pp. 445-461, 2004.
[97] R.G. Pontius and J. Malanson, “Comparison of the structure and accuracy of two land change models”, International Journal of Geographical Information Science, vol. 19, pp. 243-265, 2005.
[98] R.G. Pontius, A. H. Versluis, and N.R. Malizia, “Visualizing certainty of extrapolations from models of land change”, Landscape Ecology, vol. 21, pp. 1151-1166, 2006.
[99] N. Oreskes, “Evaluation (not validation) of quantitative models”, Environmental Health Perspectives Supplements 106(S6), pp. 1453-1460, 1998.
[100] C.S. Holling, “Adaptive Environmental Assessment and Management”, Wiley and Sons, 1978.
[101] K. Kok, A. Farrow, A. Veldkamp, and P.H. Verburg, “A method and application of multi-scale validation in spatial land use models”, Agriculture, Ecosystems and Environment, vol. 85, pp. 223-238, 2001.
[102] Y. Li, A. J. Brimicombe, and C. Li, “Agent-based services for the validation and calibration of multi-agent models”, Computers, Environment and Urban Systems, vol. 32, pp. 464-473, 2008.
[103] S.M. Manson, “Challenges in evaluating models of geographic complexity”, Environment and Planning B, vol. 34, pp. 245-260, 2007.
[104] K.C. Clarke and L. Gaydos, “Loose coupling a cellular automaton model and GIS: long-term growth prediction for San Francisco and Washington/Baltimore”, International Journal of Geographical Information Science, vol. 12, pp. 699-714, 1998.
[105] C.A. Jantz, S.J. Goetz, and M. K. Shelley, “Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore-Washington metropolitan area”, Environment and Planning B, vol. 31, pp. 251-271, 2003.
[106] K.C. Clarke, “Mapping and modelling land-use change: An application of the SLEUTH model”, in Landscape Analysis and Visualisation: Spatial Models for Natural Resource Management and Planning. C. Pettit, Cartwright, W., Bishop *et al.*, eds, Springer, 2008, pp. 353-366.
[107] D. G. Brown, S. Page, R. Riolo, M. Zellner, and W. Rand, “Path dependence and the validation of agent-based spatial models of land use”, International Journal of Geographical Information Science, vol. 19, pp. 153-174, 2005.
[108] R.G. Pontius, W. Boersma, J-C. Castella *et al.*, “Comparing the input, output, and validation maps for several models of land change”, Annals of Regional Science, vol. 42, pp. 11-37, 2008.
[109] T. Wiegand, F. Jelisch, I. Hanski, and V. Grimm, “Using pattern-oriented modeling for revealing hidden information: a key for reconciling ecological theory and application”, OIKOS 100, pp. 209-222, 2003.
[110] V. Grimm, K. Frank, F. Jeltsch, R. Branbl, J. Uchmanski, and C. Wissel, “Pattern-oriented modelling in population ecology”, The Science of the Total Environment, vol. 183, pp. 151-166, 1996.
[111] C. Chion, P. Lamontagne, S. Turgeon *et al.*, “Analysis of emerging patterns to select and validate an agent-based model of whale-watching excursions”, Ecological Modelling, Accepted for publication, 2010.
[112] T. Svoray and I. Benenson, “Scale and validity of environmental microsimulation”, Ecological Complexity, vol. 6, pp. 77–79, 2009.
[113] J. Copeland, “Artificial intelligence: A philosophical introduction", Wiley-Blackwell, 1993.
[114] M. Kieser and D.J. Marceau, “Simulating a land development planning process through agent-based modelling”, Intech Publisher, in press, 2010.
[115] F. Bousquet, G. Trébuil and B. Hardy, “Companion modeling and multi-agent systems for integrated natural resource management in Asia”, International Rice Research Institute (IRRI), 2005.

CHAPTER 2

Activity Based Variable Grid Cellular Automata for Urban and Regional Modelling

Roger White[1,*] and Harutyun Shahumyan[2]

[1]Memorial University of Newfoundland and VITO - Flemish Institute for Technological Research and [2]Urban Institute Ireland, Inge Uljee, VITO - Flemish Institute for Technological Research

Abstract: Traditional cellular automata (CA) based models of land use change represent only local processes. Processes operating over longer distances are captured in traditional spatial interaction based models; these are then be linked to the CA so that the linked models cover a range of scales. An alternative approach presented here includes processes operating at all scales within the CA model itself. This is done by increasing the size of the cell neighbourhood to include the entire modelled area, so that long range effects are included in the cellular transition functions. In order for this approach to be computationally reasonable, a variable grid is used, so that as the distance from the centre of a cell neighbourhood increases, the size of cells in the neighbourhood also increases, with cells of successive Moore rings, n, of a cell neighbourhood scaling as 3^n from the basic raster cell size. Since all Moore rings but the smallest consist of composite cells, in general the state of neighbourhood cells must be characterized by a vector of activity levels, where each cell of the raster level is assigned a quantity of the activity corresponding to its land use (*e.g.* each residential cell is assigned a population). Tests of this approach show that it gives clearly better predictions of both land use patterns and regional activity levels (*e.g.* regional populations) when compared with conventional approaches. It is also "riskier" in the Popperian sense and thus more powerful as an explanatory tool.

INTRODUCTION

Cellular automata (CA) are currently the most powerful and efficient technique available for modelling the dynamics of land use. They come in great variety. Some are raster based while others make use of cadastral polygons. Some model only two land uses—urban and rural—while others use as many as several dozen, and still others use quantitative measures of cell state, like population, rather than categorical definitions of land use. And some are linked dynamically to other models in order to augment the CA capabilities with the power and results of other approaches. In short, the field of CA modelling is developing rapidly, and some models are becoming sufficiently powerful and realistic to begin to be used as practical tools by planners. This chapter describes one approach—activity based variable grid cellular automata—that shows great promise in that it simplifies existing models while simultaneously increasing the range of phenomena that are modelled and improving the quality of the results.

MODELLING LAND USE, ACTIVITIES, AND SPATIAL INTERACTION: THE BACKSTORY

In conventional CA the state of each cell, representing its land use, depends on the state of cells in a surrounding neighbourhood, with the nature of the dependence expressed in cell state transition rules. The transition rules in turn represent attraction and repulsion effects between different land uses, and thus capture some spatial interaction or distance decay effects. In addition, in more realistic formulations, cell space heterogeneity is introduced in order to represent various factors that may have an impact on land use patterns but are not captured by the neighbourhood effect. For example, physical factors like topography and soil conditions can affect land use decisions, as can legal constraints on land use like zoning regulations, and accessibility considerations—some sites are more accessible than others to the transport infrastructure and thus to the region at large. These cell space heterogeneities are then introduced into the transition rules so that together with the neighbourhood effect they determine the land use dynamics [1].

The concept, and often the implementation, is simple but powerful; however, there are a few loose ends that must be tied up. Specifically, in a pure CA the number of cells of each state or land use depends entirely on the internal

***Address correspondence to Roger White:** Department of Geography, Memorial University of Newfoundland, St. John's, NL A1B 3X9, Canada; E-mail: roger@mun.ca

Danielle J. Marceau and Itzhak Benenson (Eds)

dynamics generated by the neighbourhood rules as applied to the initial configuration [2]. But there is no reason to expect that either the total area of each land use or the proportions among these land uses that result from the CA dynamics will be appropriate, even if the spatial patterns that are generated are realistic. Consequently, most usefully realistic CA are forced or constrained in some way in order to ensure approximately correct amounts of each land use; the forcing may, for example, be a means for representing urban growth in the region being modelled.

A number of different forcing techniques have been used, and the optimal approach depends on the nature of the model. In the case of models which operate with only a few active land uses, for example urban and rural, the forcing may be implemented through parameters that affect the probability of conversion from rural to urban, as in the Sleuth model [3]. In models with a larger number of land uses, such as those discussed in this chapter, the forcing may be as simple as a file of land use demands, with the CA constrained to achieve the specified number of cells of each type [4], or as complex as separate, linked, models that generate the land use demands [5], [6]. For example, a demographic model may be used to generate demands for land for housing, an economic model used to generate demands for land for use by various sectors of the economy, and an agricultural model used to drive requirements for agricultural land. If the forcing models are linked dynamically in both directions with the CA land use model, then land use changes will also affect the behaviour of the forcing models. For example, increasing urbanization will remove specific parcels of land from what is available to the agricultural sector, and thus alter the output of the agricultural model. This is a major step into integrated modelling of heterogeneous phenomena.

In addition to activity models, it has also frequently been necessary to link a regional dynamic model to the land use model in order to capture spatial interaction effects that operate over longer distances. A conventional CA through its transition rules defined on the cell neighbourhood gives a good representation of the spatial interaction or distance decay effects that operate at scales up to that of the neighbourhood radius. But there is no representation of these effects at longer distances. When modelling small regions, this is not a problem. But for larger areas, such as metropolitan regions or entire countries, longer distance effects become too important to ignore. The obvious solution is to adopt a traditional regionalized spatial interaction based model—for example a model like those used in transportation demand modelling.

In this approach the modelled area is divided into regions such as census tracts, counties, or provinces, and the interactions among the regions are modelled [7]. The interactions are typically in terms of migration of activities as the regions compete with each other to attract activities from each other based on their current levels of activity and their distance from each other. While the structure of such a model is quite similar to that of a transportation model, the time scale is different, with each time step normally representing a year, as opposed to a typical time step of minutes or hours for a transportation model. And the interactions represent not the movements of people going about their daily activities, but the relocation of the activities themselves—the places of residence, work, and play.

Such transport models are widely used and are known to perform well [8]. The activity migration models also seem to perform reasonably well, although some applications can be problematic. Where regions are small and homogeneous, or otherwise coherent, performance tends to be good. For example, a spatial interaction based activity migration model of The Netherlands used the 40 Corop regions into which the country is divided. These are relatively small (c. 1000 km^2) urban centred regions. Since activity is concentrated in the centre of each region, inter-regional distances are relatively well defined and meaningful, and the model results were good [6]. In contrast, in the five-county greater Dublin area of Ireland, the four peripheral counties are relatively large (up to 2335 km^2) and largely rural, with several urban centres in each one. In this case the concept of distance between counties is less well defined, with the location of county centroids between which distances are measured being somewhat arbitrary: whether the centroid is located with respect to the area of the county or with respect to the population distribution, in either case it is likely to be nowhere near any of the major population centres, and probably also far removed from any of the major transport routes. More meaningful locations can be found, but they are not objectively defined, being based on the judgment of the modeller.

The spatial interaction model generates regional activity levels for each time step based on long distance spatial interaction effects. These regional activity levels are then translated into demands for land for the various corresponding land uses—*i.e.* cell demands—which are used to constrain the CA land use model. On the basis of the local attraction and repulsion interaction effects captured in the neighbourhood-based transition rules of the CA, cells change state until the various land use demands are satisfied. In this manner both long and short distance interaction effects are included in the land use model.

This linked model approach, in which a regional dynamics model of activities drives a CA land use model and in turn receives feedback from the CA, has proved to be relatively successful in applications to the greater Dublin area—in particular the Moland model [7] (Fig. **1**)—as well as to The Netherlands [6], [9], Puerto Rico, and elsewhere. Nevertheless the approach is relatively cumbersome in several respects. First, the number of parameters is relatively large, since in addition to the parameters in both component models which must be calibrated, there are a number of others required to link the models so that they function as a single integrated model. Second, spatial processes are represented at only two scales: that of the regional model and that of the CA neighbourhood. Scales between these two are ignored. Third, activities are not explicitly represented at the cellular level. This amounts to an implicit assumption that every cell of a given land use within a region has the same amount of activity, since nothing is known about the variation of densities within a region. Fourth, each land use corresponds to a single activity. Thus, for example, residential cells are assigned population but no other activity, and agriculture cells never contain population. In reality, of course, multiple land uses are common.

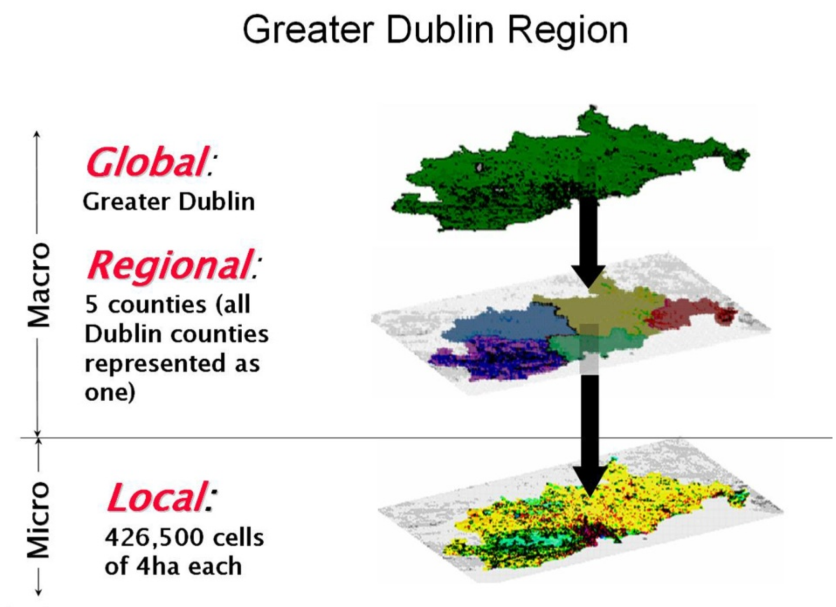


Figure 1: The three spatial scales of the Moland model of the greater Dublin area.

These observations suggest that an alternative approach, one in which interactions at all scales, activities, and land use are all handled in a single, unitary model is desirable. An activity based, variable grid (multi-scale) CA approach has many advantages, and minimizes the problems mentioned in the previous paragraph.

THE VARIABLE GRID CA APPROACH

The basic strategy in this approach is to eliminate the regionalized activity model by including its functions within the CA. In particular, activities as well as land uses are assigned to individual cells, and spatial interactions at all scales are handled by defining the neighbourhood of each cell to be the entire modelled region. Then the transition rules defined over the neighbourhood can include a representation of distance decay effects over all distances.

The Variable Grid and Cell States

The idea is simple, but the problem is how to keep CA execution times reasonable when the cell neighbourhood now contains not just a few cells but hundreds of thousands or even millions. The solution is to aggregate cells in the periphery of the neighbourhood into progressively larger super-cells as the distance from the central cell increases. In this way the number of cells in a neighbourhood is reduced to a few dozen at most [10, 11]. Since the super-cells are treated as unitary, their location must be defined by a centroid. In this sense they are quite analogous to the regions of the conventional spatial interaction models. But since they are small nearby and grow larger at progressively greater distances, the loss of detail is least where the detail is most important, *i.e.* close to the cell for which the neighbourhood is defined, and greater farther away where detail is less relevant (Fig. **2**).

Specifically, if the resolution of the basic CA raster, typically 100 m to 500 m in a land use application, is designated as level L_0, and super-cells are aggregates of 3 X 3 lower level cells, then each super-cell of level L contains 32 cells of the level L-1 grid, or (3^{2L}) cells of the fundamental L_0 grid. The evaluation of a cell's neighbourhood is carried out using the eight L_0 cells immediately surrounding the cell, and then using the eight cell rings of successively larger super cells beyond that. Since in general, unlike a conventional

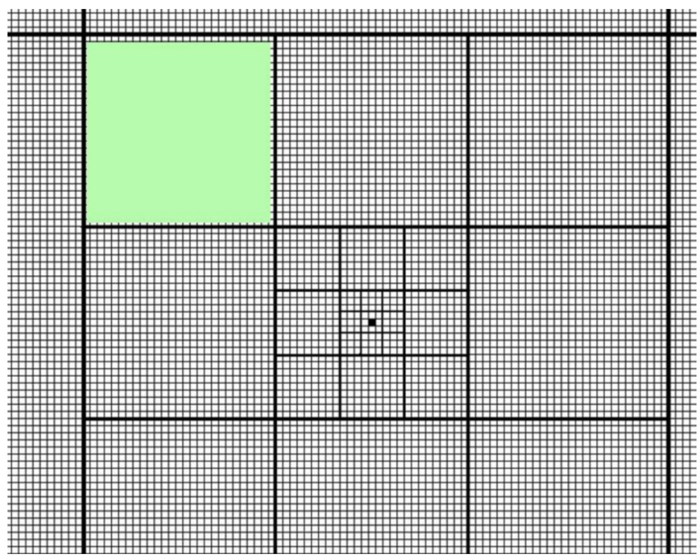


Figure 2: The variable grid template, with one of the super-cells shown in green.

CA, each super-cell includes a number of different land uses and activities, its state is represented by a vector of activity totals—the totals of activities on all basic L_0 cells in the super-cell. Normally land uses that do not correspond to a modelled activity such as parks are imputed a nominal activity level of one unit per basic cell, so that the total activity level is simply the cell count for that land use in the super-cell; but other values could be assigned, representing quality of the land use—for example the varying quality of the parkland cells for recreational purposes.

The neighbourhood of each L_0 cell is defined in terms of the cell template shown in Fig. **2**, with the template centred in turn on each basic level cell. Thus the cell state activity vectors for each super-cell in the template must be established for each basic cell. Nevertheless, the updating of the activity vectors from one iteration to the next is relatively efficient since it is not necessary to do a complete summation of the values of all L_0 cells in the neighbourhood. To calculate the new values, the variable grid template is displaced one cell at a time, and with each displacement, to get the activity totals for the state vectors of the neighbourhood of the new cell it is only necessary to subtract the totals for the trailing L_0 edge of the super-cells and add the totals for the new leading edge, which in turn becomes the trailing edge of the super-cells of the neighbourhood of the next cell.

The Neighbourhood Effect

The neighbourhood effect represents the quality or desirability of the cell neighbourhood for a particular activity or land use. Since the CA transition rules must choose a state for the cell from among all possible states, whether represented by land use or activity, for each cell a neighbourhood effect must be calculated for each possible state the cell could acquire. Once the activity vectors have been calculated for cells in the neighbourhood, the neighbourhood effect can be calculated as a weighted sum of activity values, with the weights depending on both the particular activity or land use, and the distance of the cell from the L_0 cell i for which the effect is being calculated:

where $N_{ki} = \sum_{k} \sum_{j} w_{lkji} A_{kj}$

N_{ki} = the neighbourhood effect for activity k on cell i

w_{lkji} = the weight applied to activity l in level L≥0 cell j when calculating the neighbourhood effect for activity k

A_{kj} = the level of activity k in level L≥0 cell j

The weights applied to cells j in the neighbourhood depend systematically on the distance of the cells in the neighbourhood from cell i (Fig. **3**); the zero-distance weight (w_{kk0i}) for the effect of an activity on itself represents inertia, or the extra costs involved in changing location. The pattern of the distance

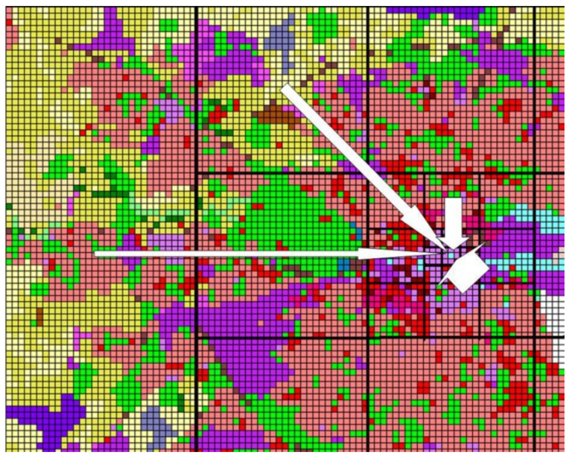


Figure 3: Schematic representation of variation by distance of weights used in the calculation of the neighbourhood effect.

dependence of the weights represents distance decay effects of various kinds. For example the weights may decline as an inverse power function of distance; in this case they replicate a traditional gravity model formulation of distance decay. But they may be negative over close distances and rise to positive values at longer distances; this configuration would represent a close-quarters repulsion effect of the activity being weighted on the activity for which the neighbourhood effect is being calculated. In fact, in all applications to date, the typical pattern of calibrated weights for most pairs of land use that are attractive to each other falls into two parts: an inner segment in which weights are very high but fall rapidly with distance, and an outer segment in which the weights are much lower but decrease much more gradually with distance, and furthermore decrease with something like an inverse power or negative exponential relationship (Fig. **4**; see also Table **1**). This strongly suggests that decisions involving land use or activity location include two distinct neighbourhood concepts: (1) the actual local neighbourhood, with a radius on the order of 0.5 km to 1.5 km, which is perceived in one way, and (2) the region, including all area beyond the local neighbourhood, in which traditional distance decay relationships characterise neighbourhood evaluation.

Table 1: Calibrated parameters for the period 1990-2000 for the variable grid model.

GENERAL PARAMETERS	
Parameter	**Calibrated Value**
random	0.4
band	1
epsilon	0.82
lambda	1.33
phi	0.9985

DISTANCE DECAY PARAMETERS (RESIDENTIAL ON RESIDENTIAL)		
log3 distance	**weight**	
0	600	
0.315	280	
1	30	
1.315	15	
2	4	
2.315	2	
3	1.3	
6	1.1	
8	0	
ACCESSIBILITY PARAMETERS		
Network Element	**Importance**	**Dist. Dec. (a_{jk})**
Motorways	0	0
Mot. Junctions	0.4	0.4
Dual Highway	0.4	3
National Highway	0.4	2
Regional Road	0.3	2
Local Road	0.3	2
Rail	0	0
Rail Station	0.4	4
Light Rail	0.4	2

Diseconomies of Agglomeration

One important influence on the location of activity and land uses, and especially on variations on the

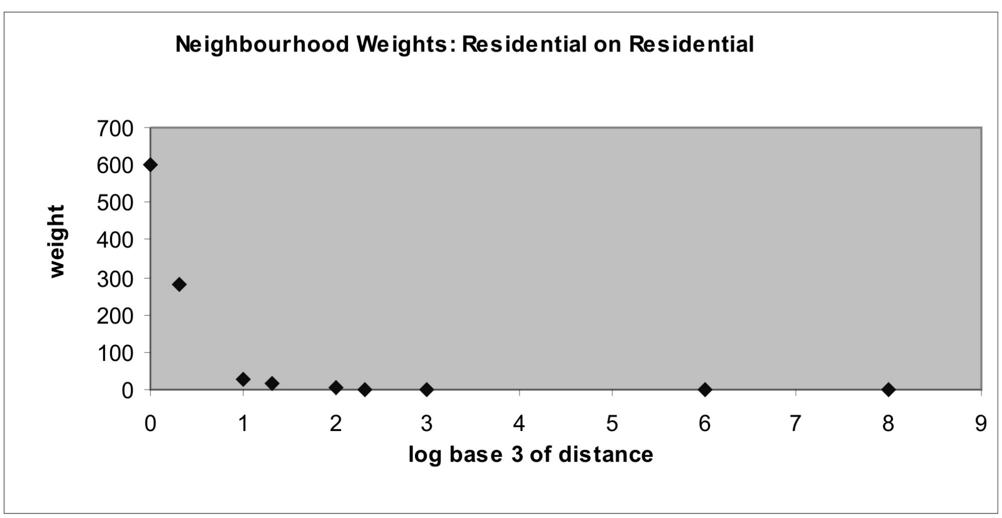


Figure 4: Neighbourhood evaluation weights as a function of super-cell distance for the effect of residential activity on residential activity (self-attraction).

density of activities at various locations, are the diseconomies that arise from concentrations of activity. Activities generally attract each other, and therefore tend to generate clusters of increasing size and density [12]. But these concentrations in turn generate negative externalities, largely in the form of higher land prices and increased congestion. As these diseconomies grow with increasing concentration of activity, they eventually tend to discourage the location of additional activity in the area. These diseconomies can be modelled in the variable grid framework as an additional, repulsive, neighbourhood effect, operating at longer distances and tending to counter the largely positive neighbourhood effects that lead to growing clusters of activity.

Diseconomies are calculated as follows:

where $E_{iK} = \dfrac{2}{1+e^{\lambda_K\left(\frac{N_{i,pop}}{N_{crit}}-1\right)}}$ $\qquad N_{crit} = \varepsilon\langle N_{init}\rangle$

E_{iK}= relative level of diseconomies of agglomeration for sector K on cell i

$N_{i,pop}$ = the population potential, that is, the neighbourhood effect calculated using only population (*i.e.* all w_{lkji} = 0 except w_{kkji} where k=population), and excluding inertia)

N_{crit}= critical level of population potential at which diseconomies appear

$\langle N_{init}\rangle$ =mean value of initial population potentials for all cells

λ_K = relative importance of economies or diseconomies for activity K

ε = parameter expressing critical level relative to initial mean population potential

The parameter λ_K accommodates the fact that some activities like financial services show very little sensitivity to diseconomies of agglomeration, whereas others, for example manufacturing, are highly sensitive. The parameter ε establishes the neutral point for the diseconomies effect. Both parameters are relatively powerful in models of urban regions such the Dublin application discussed below.

Heterogeneous Cell Space and Agents

Unlike the case in abstract CA, the cell space in geographical applications cannot be considered homogeneous. In general each cell has unique characteristics that may affect its desirability as a site for various activities. For example, a steep hillside may be somewhat suitable for housing, but unsuitable for manufacturing. Cells may therefore be characterized by slope values. All such site-specific characteristics of the cell space that are thought to influence cell state transitions are combined into a suitability measure, so that this aspect of the non-homogeneity of the cell space can be summarized by a series of suitability maps (Fig. **5a**), one for each possible land use or activity state of a cell; there would be, for example, one suitability map for housing and another for manufacturing.

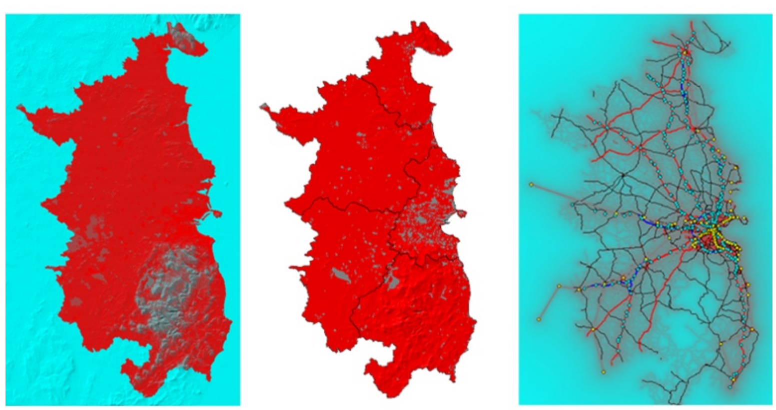


Figure 5: Cell space non-homogeneities. a. (left) suitability, representing here primarily slope; b (centre) legal restrictions, here primarily land classified as parks; c. (right) accessibility to various elements of the transport infrastructure.

Other cell space non-homogeneities can also be characterized in this way. Two that are included in the models discussed in this chapter are zoning and accessibility to the transport infrastructure. The zoning maps are straightforward, showing which cells are available (or unavailable) for particular activities at various time periods (Fig. **5b**). The accessibility measures are calculated as weighted distances from various elements of the transport network (*e.g.* motorway interchange, local road, rail station):

$T_{ji} = (1+D_{ik} / a_{jk})^{-1}$

where

T_{ji} = the accessibility of cell i to transport element k as evaluated by activity j

D_{ik} = the Euclidean distance from the cell i to the nearest cell through which the network k passes, and

a_{jk} = a coefficient representing the importance of accessibility to network element k for activity j.

A set of coefficients representing the relative importance to a particular activity of accessibility to the various network elements is used to derive a set of overall accessibility measures, one for each activity. These accessibility measures are calculated for each cell on the map (Fig. **5c**).

Each activity or land use can be thought of as representing a population of agents who are not modelled directly, but whose locational decisions, or the results of those decisions, are. From this perspective it is realistic to assume that agents corresponding to a particular land use or activity, for example individuals looking for a place to reside, or retailers looking for a shop site, are not all identical. People looking for a place to live have highly variable tastes and abilities to pay, and retail companies have quite different requirements in terms of site characteristics and location. Without modelling the agents directly, this heterogeneity can simply but adequately be represented by a stochastic perturbation on the calculations that determine a cell's desirability for the various activities:

where $r = 1 + (-\ln[rand])$ *rand* is a uniform random variate.

The perturbation has a highly skewed pdf, so that most perturbations are very small and a few are very large. This stochastic element permits new clusters to be generated, some of which are "adopted" by the system and grow significantly; this process allows the model to maintain stable cluster-size frequency relationships. The stochastic perturbation also permits the model to explore new types of location, some of which will prove to be very successful, that would not be discovered in a purely deterministic framework.

Transition Potentials

The various factors that affect the desirability of a cell as a location for a particular activity or land use are combined into a single summary measure, the transition potential. A transition potential is calculated for each activity, and its value is determined for each cell on the map. Each cell is thus characterized by a vector of transition potentials. Transition potentials are calculated as follows:

where $VT_{ki} = rE_{ki}S_{ki}Z_{ki}T_{ki}N_{ki}$

VT_{ki} = the transition potential for activity k on cell i

r = the random perturbation

E_{ki} = diseconomies of agglomeration

S_{ki} = the intrinsic suitability of the cell for the activity

Z_{ki} = the land use zoning status of the cell for the activity

T_{ki} = the accessibility of the cell to the transportation network

N_{ki} = the pure neighbourhood effect from the influence curves and activity levels

The Transition Rule

The transition rule is then to change each cell to the *land use* for which it has the highest potential, subject to the constraint that all *activities* are allocated: each time a cell receives a land use it also receives a quantity of the corresponding activity. Potentials of each cell are ranked, and then all cells are ranked according to their highest potential. Cell transitions (which are frequently to the existing land use, so in fact there is no change of state) proceed in order from the cell with the highest potential. At each transition a quantity of the corresponding activity is also allocated to the cell. In the case of cells which do not change state, the existing activity is retained. When all of a particular activity has been allocated, the corresponding potentials are ignored, so subsequently some cells may be assigned land uses other than those for which the potential is highest.

Activity is assigned as a function of potential relative to the mean potential of the area in which the cell is located:

where $A_{ki} = \langle A_k \rangle \left(\frac{V_{ki}}{\langle V_k \rangle} \right)^{\gamma_k}$

A_{ki} = the amount of activity k assigned to cell i

V_{ki} = potential—*i.e.* the neighbourhood effect excluding inertia

$<A_k>$ = mean level of activity k in cells containing the activity in the region = mean level of activity k in cells containing the activity in the region

$<V_k>$ = mean potential of the region

γ_k = a parameter to scale the magnitude of the effect

In other words, the better the location, as measured by the pure neighbourhood effect, the more activity that is assigned to a cell. Densities thus vary by location and are generally higher in large clusters.

This version of the variable grid model is something of a hybrid in that it is both land use and activity based: cell transitions are defined in terms of land use, but the potentials by means of which the transitions are determined are calculated from activity levels; and as well, the number of cells of each land use is activity constrained. Consequently while the total amount of each activity will always be correct in terms of the totals given by the data used for calibration or specified in a scenario, by the end of a simulation the resulting cell totals for each land use may be quite different from actual or scenario specified totals.

During initialization, each activity is distributed over all cells of the corresponding land use. If the simulation is to be run over a calibration period, for which data is available for both initial and final years, activity densities per cell can be determined for initial and final years, and from these, the rate of change of density, δ_k, calculated. Then at each iteration t the activity level on each cell which does not change land use is adjusted as

$$^{t}A_{ki} = \delta_k \, ^{t-1}A_{ki}$$

This tends to minimize the final errors in cell numbers for the various land uses. However, due to the allocation of activity to new cells, with the amount allocated being a function of cell potential, the calculated densities will drift away from actual values, so the δ_k must be adjusted in order to maintain approximately correct cell totals. Note that, unlike the case with the Moland model, there is no constraint on the total number of cells of each land use, so the total for the entire area, as well as the counts for the individual regions, can drift from actual values. An adjustment factor, φ_k, is therefore introduced and applied to the δ_k in order to keep cell counts close to actual values. This parameter is adjusted during the calibration process.

APPLICATIONS

The activity based variable grid model has, *a priori*, several advantages over the linked model approach. As mentioned previously, it represents processes flexibly across a range of spatial scales, rather than at only two.

Furthermore, it is simpler, since the CA includes the functionality contained in the linked macro-scale model of the other approach, and consequently there are fewer parameters to be calibrated, since all parameters of the macro-model and the parameters involved in linking the models are eliminated. Finally, it offers the potential for high resolution modelling of activity location, since it attributes activity to cells rather than relatively large regions; and since results are calculated at the level of cells, they can be aggregated to any desired set of units, such as census tracts, municipalities, or counties. But does the model actually perform better? Or is the greater simplicity bought at the cost of larger errors in the output? In general, how does the approach perform? We examine several applications to throw some light on these questions.

The variable grid approach has been tested in several contexts. In one, an application to greater Vancouver, Canada [13], land use only was modelled; the activities modelling aspect was turned off by setting both activity levels on all cells, $A_{ki,}$ and the rate of change of densities, $\delta_{k,}$ to unity. In the others—one to Portugal at 500 m resolution, and two to the Dublin area, at 200 m resolution—the full model was used. In the case of the Dublin applications, applications of the linked CA - macro-model are also available, thus permitting a comparison of the two frameworks. All three applications actively model the dynamics of five land uses—ports, industry, commerce, public and private services, and residential. However, only one activity, population, is actively modelled; this is a temporary limitation while the model undergoes further development and testing. We will focus the discussion on the more recent of the Dublin applications.

The Dublin Application of the Activity Based Variable Grid Model

The modelled area consists of five counties in the greater Dublin area: the old Dublin county and counties Louth, Meath, Kildare, and Wicklow. This area is subdivided for statistical purposes into 629 Electoral Districts (EDs). These are small, relatively homogeneous areas with populations typically in the range of 2000-3000. Electoral district data is useful for initializing the model, since population must be distributed to individual cells and the large number of EDs results in a smoother and more accurate distribution. In addition, model output can be aggregated by ED and compared with census data to test the ability of the model to make reasonable small-area population forecasts. Activity data (population in the case of the variable grid model; population and employment data in the case of the Moland model) is from the census of Ireland for 1986, 1991, 2002, and 2006, interpolated to the years 1990, 2000, and 2006 in order to conform to the dates of the available land use maps. Land use data are provided by the Land Management Unit of the Joint Research Centre of the European Commission. They are based on Corine data and augmented with local ground data in order to provide additional urban land use categories, ones corresponding to demographic and economic categories.

County Level Results

Results are aggregated to the county level for both calibration and testing purposes. Trying to calibrate to the 629 individual population and cell counts of the EDs would be hopelessly complex; summary measures like root mean square error (RMSE) are used, but lose spatial aspects of the calibration problem. Calibration to minimize errors in predicted county populations and cell counts is relatively straightforward and permits at least a rough control on spatial aspects of the calibration. It is not infrequently the case that a calibration that minimises a measure of global error like RMSE hides large spatial biases in the errors. It is preferable to aim for errors that are of similar size across all regions even at the expense of a global error somewhat higher than the minimum that can be achieved. The parameters to be calibrated are those discussed in the previous section, and include the weighting parameters that define the distance decay relationship in the neighbourhood effect and the parameters for calculating accessibility to the transport system. Values calibrated over the period 1990-2000 are shown in Table **1**. However, due to the stochastic term the variable grid model gives varying ED and county population estimates for a given parameter set, because populations are tied to residential cells and land use transitions are affected by the perturbation. Consequently, to get reliable output, the variable grid model must be run a number of times with a given set of parameter values and the results averaged over the runs.

The same data has also been used to calibrate and test the linked model approach. The Moland model, in which the CA is linked to a regional activity model defined on the five counties, was provided by the Land Research Unit of the Joint Research Centre, and the applications to the Dublin area were carried out by the Urban Institute Ireland, University College Dublin [7]. These applications use the same data as the variable grid applications (with the exception of ED data), and so the results are directly comparable. In the case of the Moland model, the county population predictions are essentially independent of the land use dynamics, being calculated in the linked deterministic regional model.

Over the 1990-2000 calibration period, the performance of the variable grid model was clearly better, with a root mean square error (RMSE), expressed as a percent of total population in the final year, of 0.98%, while the RMSE of the Moland results for the same period was 2.31%. Errors in population predictions by the variable grid model were smaller for all five counties. These results are shown in Fig. **6**. Constant share predictions are shown also for comparison purposes. The constant share is a frequently used "null" prediction; it assumes that each county grows so as to maintain its share of the population so that the regional distribution remains unchanged. The constant share

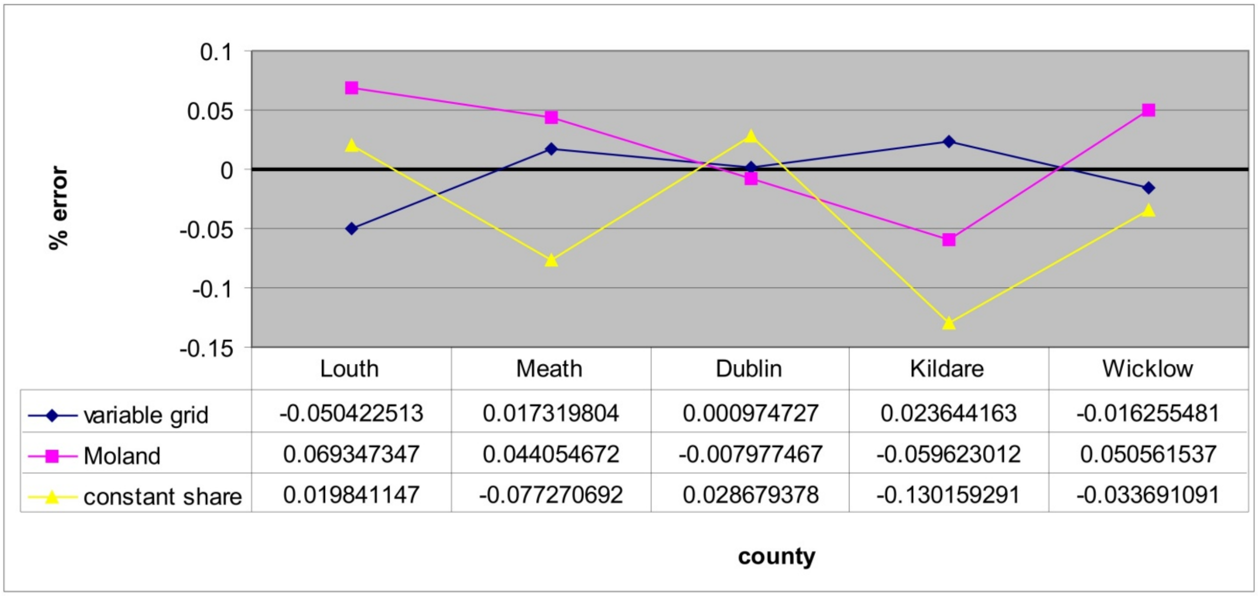


Figure 6: Percent errors, 2000, in county population predictions by three approaches: variable grid (blue), Moland (pink) and constant share (yellow).

RMSE is 5.47%. The results for the variable grid model represent the mean of 80 runs of the model using the parameter values shown in Table **1**. Since the calibration process of the variable grid model must attempt simultaneously to optimise both population and cell counts for each region, errors in county residential cell counts, as well as the total for the modelled area, are also shown (Table **2**).

Table 2: Errors in cell numbers averaged over 80 runs of the calibrated variable grid model.

County	Cell Error
Louth	-5.88%
Meath	-5.45%
Dublin	0.80%
Kildare	0.69%
Wicklow	-1.77%
RMSE	29.00

Beyond the calibration period, it is normal to validate a model by running it over a second period for which data is available to see how well the model performs over a period for which it was not optimised. In such an exercise the model is initialized with data for the start year of the validation period. Such results were not available for the Moland model, so a comparison of validation runs is not possible. However, both models were calibrated for the 2000-2006 period; the results are shown in Fig. **7** and Table **3**. The relative results of the two models are reversed for this period: Moland performs better than for the earlier period, and better than the variable grid model, which performs not as well as for the earlier period. Both, however, give much better results than the constant share approach.

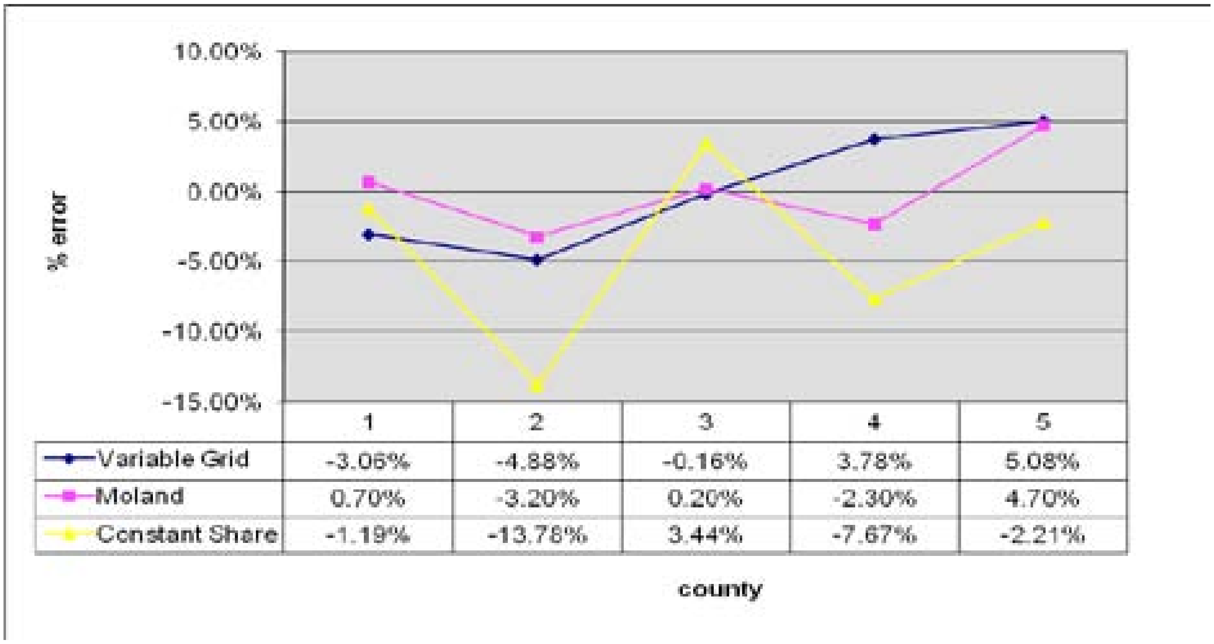


Figure 7: Percent errors, 2006, in county population predictions by three approaches: variable grid (blue), Moland (pink) and constant share (yellow).

Table 3: Overall root mean square errors (RMSE) expressed as a percent of total population; the variable grid result is the mean of 17 runs of the model.

Model	**RMSE (2000)**	**RMSE (2006)**
Variable Grid	0.98%	1.20%
Moland	2.31%	1.67%
Constant Share	5.47%	6.15%

ED Level Results

In the variable grid simulations, population allocated to cells is also aggregated to the EDs. At this level errors are greater than at the county level, partly because these errors were not used as a calibration criterion, but mostly because small areas are more variable and harder to predict. In particular, if an ED has only one residential cell (and that is the case with a number of EDs), to which all the population is attributed, if that cell disappears during the course of the simulation, then all population is removed as well, and the error in predicted population is immediately -100%. Similarly, if residential cells are added to EDs with very few such cells, the resulting allocation of population is likely to create errors well above 100%. Most of the large errors are in very rural EDs. This is essentially a data rather than a modelling problem.

For the 80 runs of the calibrated model to 2000, the RMSE for ED populations is 46.8%. An error map for one of these runs is shown in Fig. **8**. Red and orange show under-predictions and green over-predictions. More work is needed in order to understand how reliable and useful variable grid population predictions at this scale can be. Currently results are being further analysed in order to understand how errors are related to ED characteristics in terms of degree of urbanization and distance from urbanized EDs. Indications from current results are that EDs in which more than 80% of the cells are characterized by urban land uses, including more than 50% residential, have mean errors (RMSE) in predicted population that are about one-half those for EDs which have at least 80% urban but less than 50% residential cover, and approximately one-fourth those for rural EDs. There is little difference in size of error between rural EDs that are peri-urban, *i.e.* adjacent to urban EDs, and those which are truly rural.

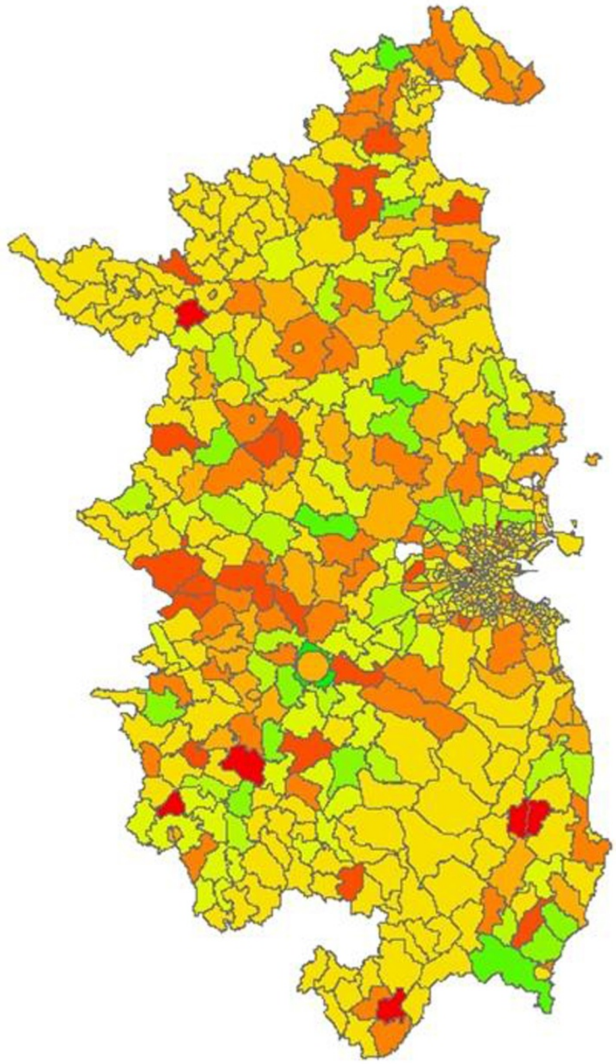


Figure 8: Errors in Electoral District (ED) population predictions for the year 2000. Scale is from dark red (-100% error) to dark green (≥100% error).

Land Use

Normally in CA modelling of geographical systems the emphasis is on the spatial patterns of land use that are generated by the model. Are they realistic? What are their fractal dimensions? How good are they according to various statistical tests? In this chapter we have ignored these issues; the emphasis has been on using land use modelling to predict activity totals by region. But using the activity based variable grid approach it is possible to generate an unrealistic or erroneous land use pattern that nevertheless gives good estimates of population by regions. So it is also important as part of the model validation process to ensure that the land use patterns generated are realistic. In this respect the calibration reported in this chapter performs well. The general pattern of urban development for the period 1990-2000 generated by the model (details vary from run to run because of the stochastic perturbation) is very similar to the observed pattern. The simulated land use map for the year 2000 for part of the modelled area from one of these runs is shown in Fig. **9** with the actual land use for the same year shown in Fig. **10** for comparison purposes.

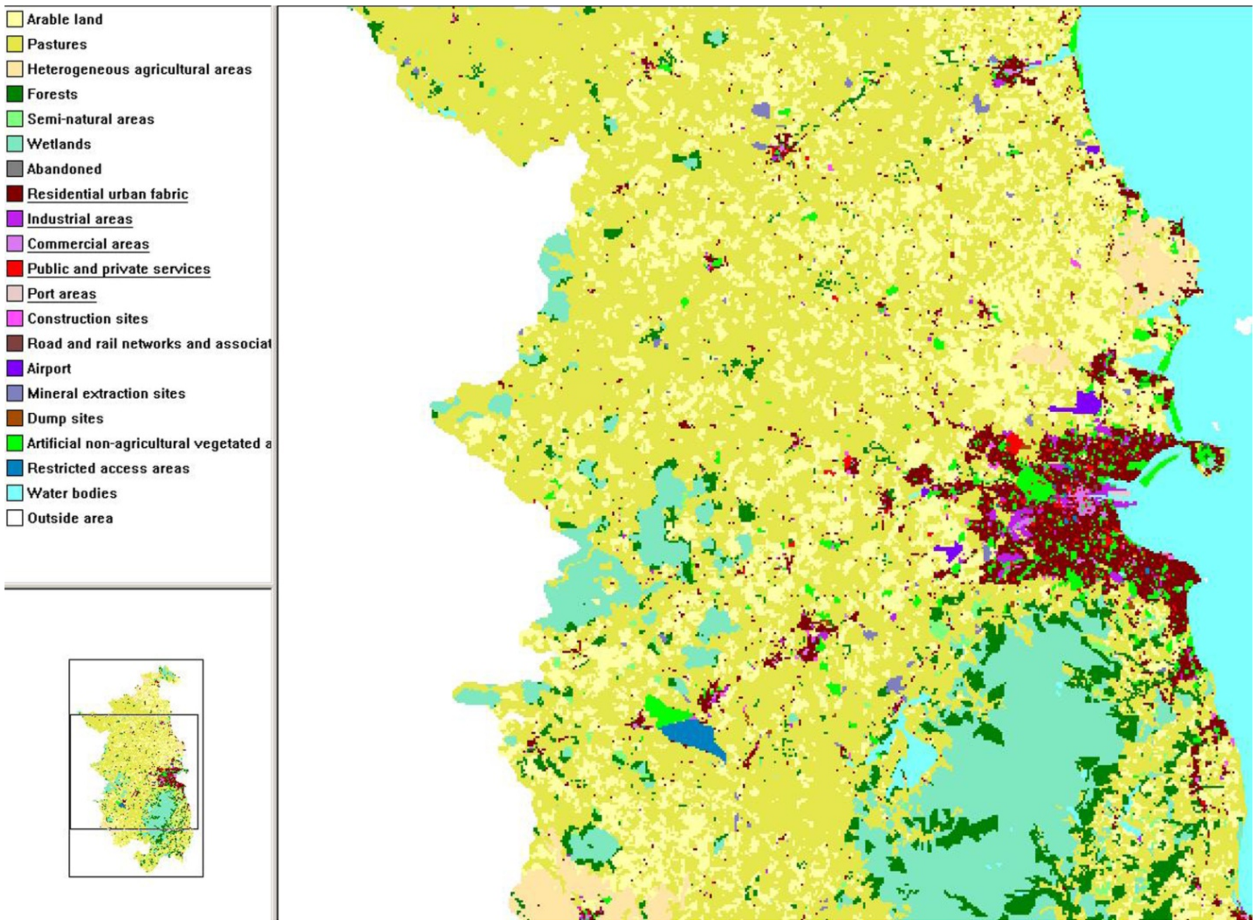


Figure 9: Part of a land use map for the year 2000 generated by the variable grid model.

CONCLUSIONS

The activity based variable grid approach offers two major advantages over a linked macro-model – CA approach. First, it is simpler: it is one model rather than two, with many fewer parameters. Second, it models more phenomena of interest at the micro scale: activities as well as land use. These advantages are of major practical importance if CA modelling is to move into the world of everyday application in the form of desktop tools for spatial planning and policy development. Currently planners must rely on history, experience, and intuition to understand the future consequences of current policies, since available tools like GIS are extremely good at descriptive tasks but do not deal with long term dynamics, and in particular do not capture the cumulative effects of repeated interactions among the actors or components of the system. Dynamic, integrated models, on the other hand, are directed specifically at understanding these long term consequences, and thus have the potential for providing a valuable complement to the existing suite of planning and policy tools. Because of the level of detail of both land use and activity location available in the output of a activity based variable grid CA, this approach can be thought of as something close to a dynamic, predictive GIS, and thus a natural extension of that tool.

From a scientific point of view, too, the approach is ambitious. In a conventional constrained CA, total cell counts for each land use are determined exogenously, so only the location is the result of the CA dynamics. In spatial interaction based activity location models there is typically no representation of land use and so no cell numbers to determine. In the activity based variable grid model, however, both cell numbers and activity levels per region are determined simultaneously with a single set of parameters, including one for adjusting density trends. Of course cell numbers per region are simply a summary expression of the land use pattern, and it is this that is actually determined by the parameter set. Since to be successful the activity based variable grid model must produce reasonable results for both activity distribution and the land use pattern with many fewer parameters than the linked models it replaces, it is much more risky in the Popperian sense; and since on balance it seems to give equally good, if not better predictions, it must be judged more powerful.

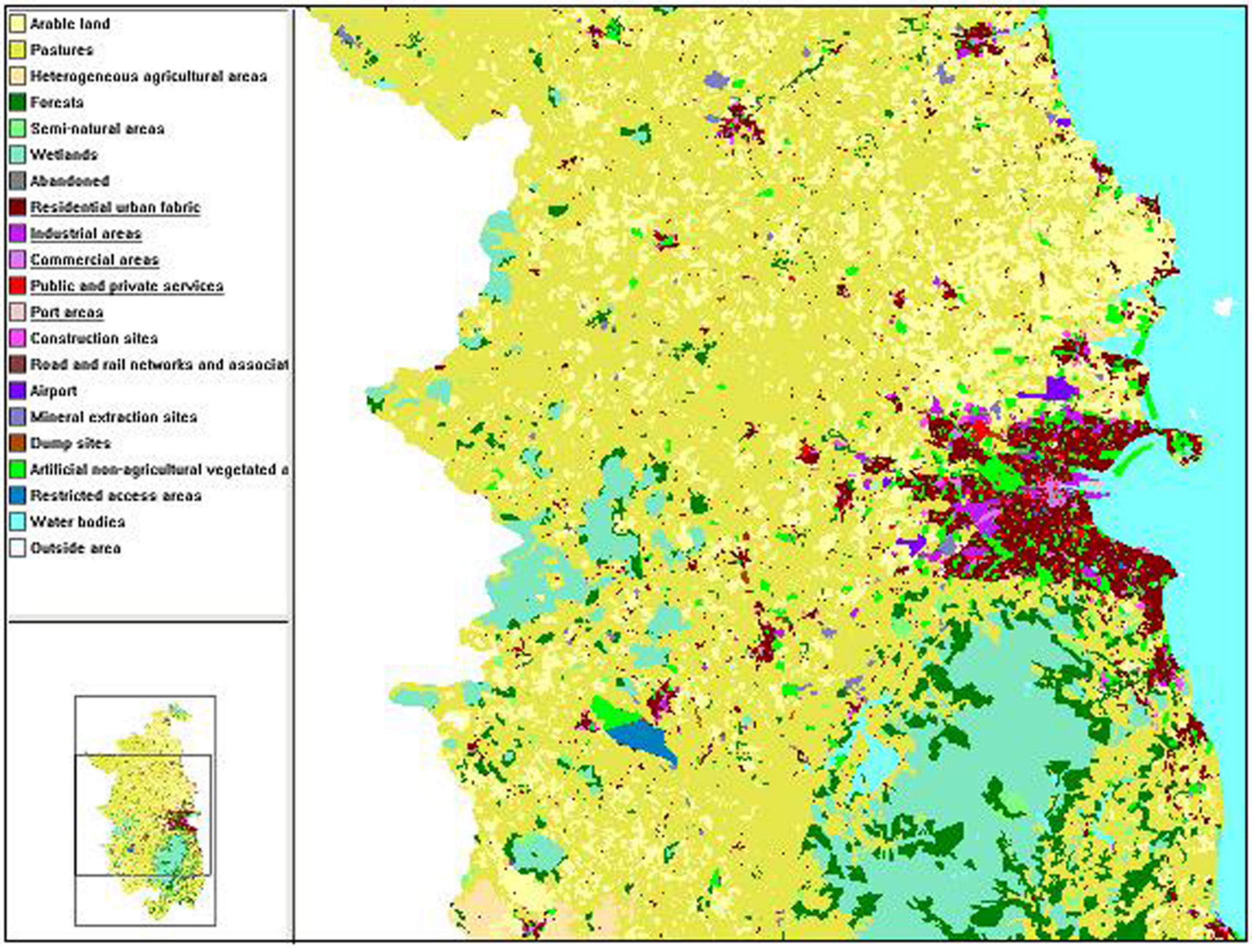


Figure 10: Actual land use for the year 2000 for the area shown in Fig. **9**.

ACKNOWLEDGMENTS

This work was supported by the Belgian research agency Belspo under project SD/TA/01A (MultiMode), the Science Foundation Ireland through the Urban Environment Project of the Urban Institute Ireland at University College Dublin, and the Social Sciences and Humanities Research Council of Canada (SSHRC).

REFERENCES

[1] R. White, G. Engelen G, and I. Uljee, 'The use of constrained cellular automata for high-resolution modelling of urban land use dynamics', Environment and Planning B, vol. 24, pp. 323-343, 1997.

[2] S. Wolfram, 'A New Kind of Science', Wolfram Media, 2002.

[3] K. Clarke and L. Gaydos, 'Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore', International Journal of Geographic Information Science, vol. 12, pp. 699-714, 1998.

[4] R. White and G. Engelen, 'Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban form', Environment and Planning A, vol. 25, pp. 1175-1199, 1993.

[5] R. White and G. Engelen, 'Cellular automata as the basis of integrated dynamic modelling', Environment and Planning B, vol. 24, pp. 235-246, 1997.

[6] G. Engelen, R. White, and T. De Nijs, 'Environment explorer: Spatial support system for the integrated assessment of socio-economic and environmental policies in the Netherlands', Integrated Assessment, vol. 4, pp. 97-105, 2003.

[7] G. Engelen, C. Lavalle, J. Barredo, M. van der Meulen, and R. White, 'The MOLAND modelling framework for urban and regional land use dynamics', in Modelling land-use change: progress and applications, E. Koomen, J. Stillwell, A. Bakem, and H.J. Scholten, eds, Springer, pp. 297-320, 2007.

[8] D. Hensher and J. Button, 'Handbook of transport modelling', Emerald Group, 2000.

[9] T. de Nijs, R. de Niet, and L. Crommentuijn, 'Constructing Land-Use Maps of the Netherlands in 2030', Journal of Environmental Management, vol. 72, pp. 35-42, 2004.

[10] C. Andersson, S. Rasmussen, and R. White, 'Urban settlement transitions', Environment and Planning B, vol. 29, pp. 841-865, 2002.

[11] R. White, 'Modelling multi-scale processes in a cellular automata framework', in Complex Artificial Environments, J. Portugali, ed., Springer-Verlag, pp. 165-178, 2005.

[12] P. Verburg, J. Ritsema, T. de Nijs, M. Dijst, and P. Schot, 'Determinants of land-use change patterns in the Netherlands', Environment and Planning B, vol. 31, pp. 125-150, 2004.

[13] J. van Vliet, S. Dragecivic, and R. White, 'Modelling urban growth using a variable grid cellular automaton', Computers, Environment and Urban Systems, vol. 33, pp. 35-43, 2009.

CHAPTER 3

Geographical Vector Agent Based Simulation for Agricultural Land Use Modelling

Antoni Moore

School of Surveying, University of Otago, New Zealand

Abstract: Spatial structures of anthropogenic origin tend to exhibit complex patterns made up of heterogeneous and irregular objects. While the urban structure is a typical example and has received a lot of attention, the agricultural domain can be included in this group of structures, also inheriting its attendant properties. Complex domains of this kind form a challenge to geo-simulation techniques, notably cellular automata (CA). Like Geographic Automata Systems (GAS), Geographical Vector Agents (GVA) have been introduced as a vector geometry alternative to CA, which bases its operations on geographically unrealistic regular cells. Indeed, the geometric element of GAS has already had a GVA-based structure imposed on it, with irregular and dynamic GVAs in particular being explored. Given the irregular nature of both domain and technology, it is this type of GVA that is applied to the modelling of agricultural land use and is reported on in this chapter. A simple theoretical scenario was presented and evaluated, an attempt to parameterise GVA to model Von Thünen's theory of agricultural land use. GVA was found to handle well the irregular geometry and neighbourhood demands as well as the rules and states relating to this hypothetical agricultural scenario.

INTRODUCTION

One of the key issues in spatial modelling is the abstraction of a given phenomenon. Many real-world phenomena manifest themselves as objects with regular / irregular boundaries (*e.g.* land use patterns or man-made objects such as buildings) or even an imprecise (vague) geometric boundary (*e.g.* forest, mountains). Spatial model and simulation techniques are mostly based on the property of spatial phenomena characteristics (Goodchild, 1992) [1]. Benenson and Torrens (2004) [2] have pointed out, "if modelled phenomena are an abstraction of real-world phenomena, why should modelled objects differ from their counterparts in the real world?" (p.4).

With this in mind, Geographical Vector Agents (GVA) have been introduced as a generic spatial modelling framework populated by objects as agents (Hammam *et al*, 2007) [3]. Agents are objects that computationally possess a goal or set of goals (Luck *et al*, 2003) [4], having the ability to make decisions in the process towards achieving the goal(s). The GVA fits within the geometric element of the Geographic Automata System (GAS) framework, developed by Torrens and Benenson (2005) [5] to provide a dynamic vector-based modelling structure as an alternative to the prevailing and less spatially realistic Cellular Automata (CA).

The use of CA for dynamic spatial modelling is well established with complex spatial phenomena, especially city growth and differentiation (Batty, 2000) [6], (Batty, 2001) [7]; (Torrens and O'Sullivan, 2000) [8]. This is down to the simplicity of cellular space to model the real world and the ease of implementation in terms of computational complexity. The cellular structure has also formed the underlying spatial data structure for agent systems from the early spatial agents (*e.g.* Sanders *et al*, 1997) [9] to more recent Multi Agent Systems (*e.g.* Batty *et al*, 2003) [10].

However, the division of space into regular square cells is a limiting assumption in a spatial simulation domain (Benenson and Torrens, 2004) [2]. There have been attempts to modify the limited geometry of CA, including linear cells (Wahle *et al*, 2001) [11] and Voronoi polygons (Shi and Pang, 2000) [12]. Efforts to move beyond the rigid topology-based neighbourhood concepts of CA have used Delaunay triangle links (Semboloni, 2000) [13] and planar graphs (O'Sullivan, 2000) [14], (O'Sullivan, 2001) [15].

A drive to escape the constraints of the fixed cell led to the development of the Object-Based Environment for Urban Simulation (OBEUS), the original GAS (Torrens and Benenson, 2005) [5]. GA are distinguished from CA by

***Address correspondence to Dr. Antoni Moore:** School of Surveying, University of Otago, PO Box 56, Dunedin, New Zealand; Tel: 0064 3 4797589; Fax: 0064 3 4797586; E-mail: tony.moore@otago.ac.nz

Danielle J. Marceau and Itzhak Benenson (Eds)

having the means to explicitly store its own geometry as well as rules for the movement of that geometry and rules for the alteration of neighbourhood effects. Two types of agent were distinguished – the fixed agent (*e.g.* land parcel) and the non-fixed agent (*e.g.* social "actors") of no physical manifestation. Fig. **1** represents this continuum of spatial simulation system geometry from square cells (CA) to autonomous vector objects, including GVAs.

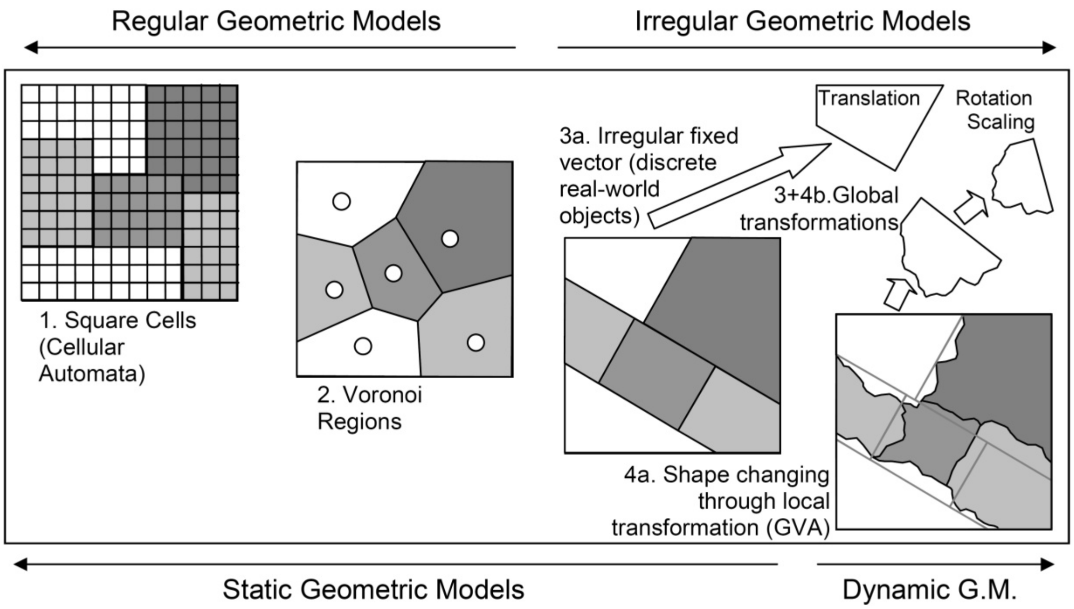


Figure 1: Various methods of discrete spatial modelling, arranged on a continuum from regular to irregular geometry, and static to dynamic models. The form of vector agent discussed in this paper is type 4a.

The GVA is an object that can represent dynamic and non-dynamic; regular and irregular vector boundaries and specifically develops the GAS "georeferencing convention" element (originally defined by Torrens and Benenson, 2005) [5] to represent a wide range of geospatial phenomena. Like GAS, GVA is grounded in letting the "geography" lead the simulation, with geometry considered a subset of that when regarded in its widest sense. However, objects with dynamic geometry are prevalent in the real world (*e.g.* natural habitats and urban areas) and are subject to irregular growth. These concepts have not been addressed by spatial automata and agents. In Torrens and Benenson's paper (2005) [5], the existence of objects that change their geometric form over time is briefly mentioned: "…there are instances in which georeferencing is dynamic for the geographic automata that represent infrastructure objects, for example when land parcel objects are sub-divided during simulation" (p.392). Moving towards this state of being are the aforementioned non-fixed objects, which point to a vector object but have no geometry in themselves ("indirect georeferencing").

Recognising that geometry may have an important role in these systems, a dynamic and irregular geometric framework was added to GAS. Algorithms that controlled object size and boundary complexity were implemented and tested visually with selected real-world instances (Hammam *et al*, 2007) [3]. Thus, like geographic automata, the vector agent can be a direct abstraction of a real-world entity, but CA cells have the disadvantage of needing explicit grouping to form an extended object (Torrens and Benenson, 2005) [5]. Moreno *et al* (2009) [16] introduced VecGCA, a vector-based cellular automata, like GVA representing space through irregular vector objects, but representing real-world entities. In particular, they investigated the effect of irregular and dynamic neighbourhood on land-use simulation runs.

This paper reports on work that implements GVA in a theoretical spatial model context, a step towards the calibration of vector agents with real world processes. In the next section, the generic GVA model is outlined in terms of its implementation within the agent modelling shell Repast Simphony and how the dynamic geometry is driven (the latter summarized from Hammam *et al*, 2007) [3] with additional exposition of newer GVA developments: neighbourhoods, states and their associated rules (in line with the GAS framework). The theoretical spatial model used is von Thünen's model of agricultural land use (Hall, 1966) [17] and its assumptions and parameterisation (spatial and non-spatial) in a GVA context will be introduced in section 3. Initial experimental results from von Thünen simulations will then be presented, followed by analysis, discussion and conclusions.

THE GEOGRAPHICAL VECTOR AGENT

What it is

The Geographical Vector Agent is a spatial agent that is physically and explicitly defined by a Euclidean geometry, able to change its own geometric characteristics while interacting with other agents in its neighbourhood using a set of rules (Hammam *et al*, 2007) [3].

Its specific properties include:

- an ability to represent multifarious discrete geographic phenomena through an irregular (or regular) vector data structure
- a dynamic nature that manifests itself either through boundary manipulation or by being able to move as a whole (translation)
- a basis on real-world objects or on a non-deterministic shape boundary
- an ability to drive its own geometric nature, including location in space and boundary configuration
- being part of a dynamic neighbourhood structure (enabled by Delaunay triangulation of GVA centroids)
- storing and manipulating its own (a)spatial attributes (states), the change of which is subject to transition rules, neighbourhoods and time
- the potential to model many application domains such as agricultural land use or urban differentiation, facilitated by the flexibility of the GVA geometric controls, neighbourhood, time, states and transition rules.

GEOMETRIC CHANGE THROUGH MIDPOINT, EDGE AND VERTEX DISPLACEMENT

The algorithms that implement geometric change in this initial implementation of the GVA were chosen to adequately capture real-world objects with a range of areal and boundary complexities and characteristics. One such algorithm is random midpoint displacement, adapted from irregular fractal geometry generation as it can adequately represent real-world entities that are not smoothly formed, as is the case with most natural objects (Laurini and Thompson, 1992) [18], the source of the vast majority of geographic geometric forms (including landscape features).

Midpoint displacement in the way described above is a form of Brownian Motion (Kenkel and Walker, 1996) [19]. Considering a single line segment (part of a polygon boundary) as an initiator, recursive subdivision by midpoint displacement occurs, with the displacement normally made perpendicular to the line (though in this case it is subject to an angle α_{md}). The following is a generalised algorithm for simple displacement (adapted from Laurini and Thompson, 1992) [18]:

$$P_{new\_md} = \tfrac{1}{2}(P_1 + P_2) + \mu\sigma_0 2^{-lh} \quad \textbf{(1)}$$

where $\mathbf{P_1}$ and $\mathbf{P_2}$ are the start and end points of the line segment being subdivided, represented in vector form. The second group of terms (governing the amount of perpendicular displacement) includes $\boldsymbol{\mu}$, which is a random number from a Gaussian distribution; $\boldsymbol{\sigma_0}$ is the standard deviation of that Gaussian curve (equal to 1; mean = 0), $\mathbf{l}$ is the level of recursivity (in this case incremented by an amount << 1), and $\mathbf{h}$ is the Hurst exponent specifying the roughness of an object (Voss, 1988) [20]. The fractal dimension D is equal to $(2 - h)$. With this relation the fractal dimension D of regular Brownian motion ($h = 0.5$) is 1.5. When $h < 0.5$ a shape is considered rough and when $h > 0.5$ the shape is assumed to be smooth.

The other two algorithms that operate alongside midpoint displacement as the basis for the stochastic vector agent generation method are vertex and edge displacement. The latter in particular promotes the generation of more regular looking shapes.

Vertex displacement displaces a single existing point on the boundary by a random amount on a bearing α_{vd} (*i.e.* without fractal modelling parameters):

$$P_{new\_vd} = P + \mu\sigma_0 \quad \textbf{(2)}$$

Edge displacement applies this equation to two consecutive points on the boundary, with magnitude and direction (*i.e.* α_{ed}) of displacement equal for both points.

SHAPE CONSTRUCTION AND EVOLUTION IN VECTOR AGENTS

The following account illustrates how the three algorithms are used in tandem to evolve the vector shapes (Fig. **2**). The first stage is a point allocation in space, whether random or more directed (for this simulation the allocation is driven by a random value taken from a normal distribution). Then a second point is placed at a random angle and distance from the first. Midpoint displacement is used to place the third point and close the polygon (Fig. **2a - d**). Consequently, the geometry starts evolving using one of midpoint displacement (Fig. **2e, f**), edge displacement (Fig. **2g, h**) or vertex displacement (Fig. **2i, j**), according to predefined probabilities of use.

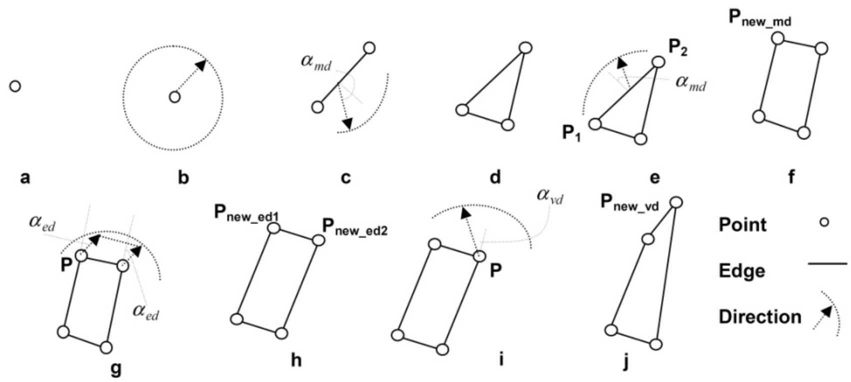


Figure 2: How a vector agent shape is born and evolves in the spatial simulation domain: **(a)** initialising by random point, **(b)** allocating second point by random displacement, **(c, d)** applying the random new point displacement and accomplishing closed polygon, **(e, f)** choosing any edge randomly and applying the new point displacement, **(g, h)** edge displacement, **(i, j)** vertex displacement.

An investigation of various probabilistic combinations of the algorithms (Hammam *et al*, 2007) [3] is summarized pictorially in Fig. **3** as a guide for recommended use. The schematic takes into account the observed ability of midpoint displacement to adjust the shape to the desired level of complexity (fractal dimension), taking into account the possible additive effects of using other algorithms. Control in the temporal axis as well would entail the use of vertex and edge displacement to model slow and fast shape growth rates respectively.

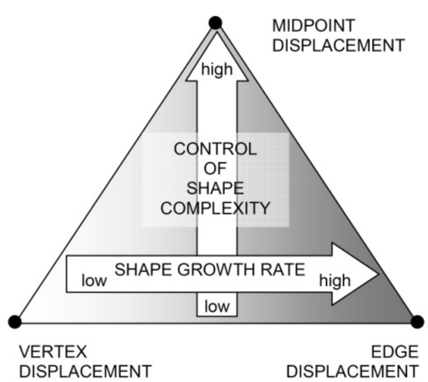


Figure 3: Using the three algorithms to control the complexity and growth of a shape.

GENERIC MODEL IMPLEMENTATION IN REPAST

An agent that is physically represented by the explicit and dynamic geometry demonstrated in the previous section is quite clearly a spatial agent, with the spatial conventions that implies. For example, it is defined within a coordinate system based upon a Euclidean geometry and subject to some projective transformation of the whole space. It is also of course a temporal agent, subject to change at (mainly) regular time increments.

Repast-S

A Geographic Vector Agent application for the von Thünen scenario was developed in the Repast (Recursive Porous Agent Simulation Toolkit) framework, in Java (Howe *et al*, 2006) [21]. It was decided to use Repast for this implementation, as since the original GVA model development (developed in Java; Hammam *et al* 2007) [3] this toolkit had added support for geographic vector data that had not necessarily pre-existed in a dataset (*e.g.* a shapefile of parcels) and for enabling geometric change within the toolkit environment. In other words, it was now possible to evolve non-deterministic GVA objects in the Repast framework, reaping the ancillary benefits that a mature agent toolkit provides.

Repast Simphony (Repast-S) was released in 2006 (Repast, 2009) [22], introducing “contexts” and “projections” for spatial agent model builders. The context is the core data structure, a “proto-space” that acts as a container for agents that are semantically similar, or localized. This context can be expressed geographically (enabling the implementation of GVA) or otherwise (*e.g.* networks, grids) through one or more projections. For instance, spatial modelling over Euclidean space is enabled by a Geography projection. Moreover, agents can simultaneously exist in a non-geographically (abstract) represented network projection; this is enabled by having agents and projections independent of each other.

Another attribute of note for Repast-S is that context-sensitive behaviour can be modelled for the proto-agents within the context. They use “watchers” to determine if a trigger for a certain behaviour applies (Howe *et al*, 2006) [21]. For example, this mechanism is used in the application to initiate GVA, triggered by the correct environmental circumstances (*i.e.* if the state of another agent changes).

Therefore the agents implemented here are goal-agents, according to Luck *et al*’s (2003) [4] taxonomy of agents. Such an agent can perceive information from passive or static objects and other agents in the environment, make decisions about actions according to a given task or goal, then perform those actions.

GVA in Repast

For the implementation described in this paper, there is one context, VecContext, to which all agents belong. (See the UML diagram in Fig. **4**.) The agents exist in a hierarchy, with SimpleAgent as the top level agent. There are two children of SimpleAgent, GeometricAgent, which groups the types of agent that have a geometry (the GVA itself – VecAgent – and MakerAgent). The other child agent type is IndividualAgent, which does not have a geometry in itself; rather it links to other agents that do have a geometry. Describing the children of GeometricAgent, VecAgent is a dynamic geometry agent that behaves physically in the way described in section 2 and MakerAgent creates the VecAgents in the first place, putting them into the context. The latter has a simple geometry that does not change in this implementation.

This is a description of the abstract level classes – they are designed to be generic and built on with specific scenarios. The application-defined agent classes holding von Thünen-specific members and methods (CityAgent, FarmAgent and EconomicAgent, which extend MakerAgent, VecAgent and IndividualAgent respectively) will be described in the next section.

Four projections exist within the context, one geographic to hold the geometry of the GVA and three network representations:

a) a vector network holding linear links to neighbouring vector objects (*i.e.* the VecAgents, MakerAgent) according to a Delaunay triangular network for a topological representation to model agent neighbourhood

b) a network that includes the links from IndividualAgents (or instantiations of any of its possible children classes) to the VecAgent or MakerAgent geometry currently associated with it (AssociationNetwork).

c) a network specific to the Von Thünen implementation that maintains economic linkages from each Individual (Economic)Agent to the Maker (City)Agent (EconNetwork).

Fig. **5** shows the relationship of the context, the four projections and the agents that inhabit them. Supporting the action of the VecAgent, the Repast Geometry and GeometryFactory classes (from the JTS – Java Topology Suite - library) have been extended (as VAGeometry and VAGeometryFactory) to include methods that are specific to evolving a GVA (*i.e.* midpoint, edge, vertex displacement) and its location definition in geographic space. The geometric objects themselves are also extended forms of the Point, LineString and Polygon classes (VAPoint, VALineString, VAPolygon; only the latter of which will be implemented as part of a GVA in this research) in the JTS library.

To form the basis for neighbourhood operations, this implementation of GVA uses a Delaunay triangular network (as in Semboloni, 2000) [13], constructed on the centroids of the objects (in the VAGeometry class). The network is then accessed by the VANeighbour class as needed in the course of a simulation run. At this stage in the GVA's development, the Delaunay neighbourhood (where an object is a neighbour of another object if there is a Delaunay triangle link connecting their respective centroids) is only accessed by methods that detect overlap topology so it can be avoided. In this way the evolution of the object is restricted so that neighbouring objects are not superimposed. These can be implemented on immediate and extended (*i.e.* two or more links) Delaunay neighbourhoods to reduce computing resources being used to guarantee this planar enforcement. The extended Delaunay neighbourhood is achieved via a recursive function which calculates the neighbours of neighbours (and so on) of the focus object.

Returning to the GAS definition (Benenson and Torrens, 2004) [2], there is separation of geometry, their neighbourhood and transition rules (the latter, along with states is dealt with in the application specific outline in section 4). Geometry rules are explicitly dealt with, though neighbourhood rules are not at this stage.

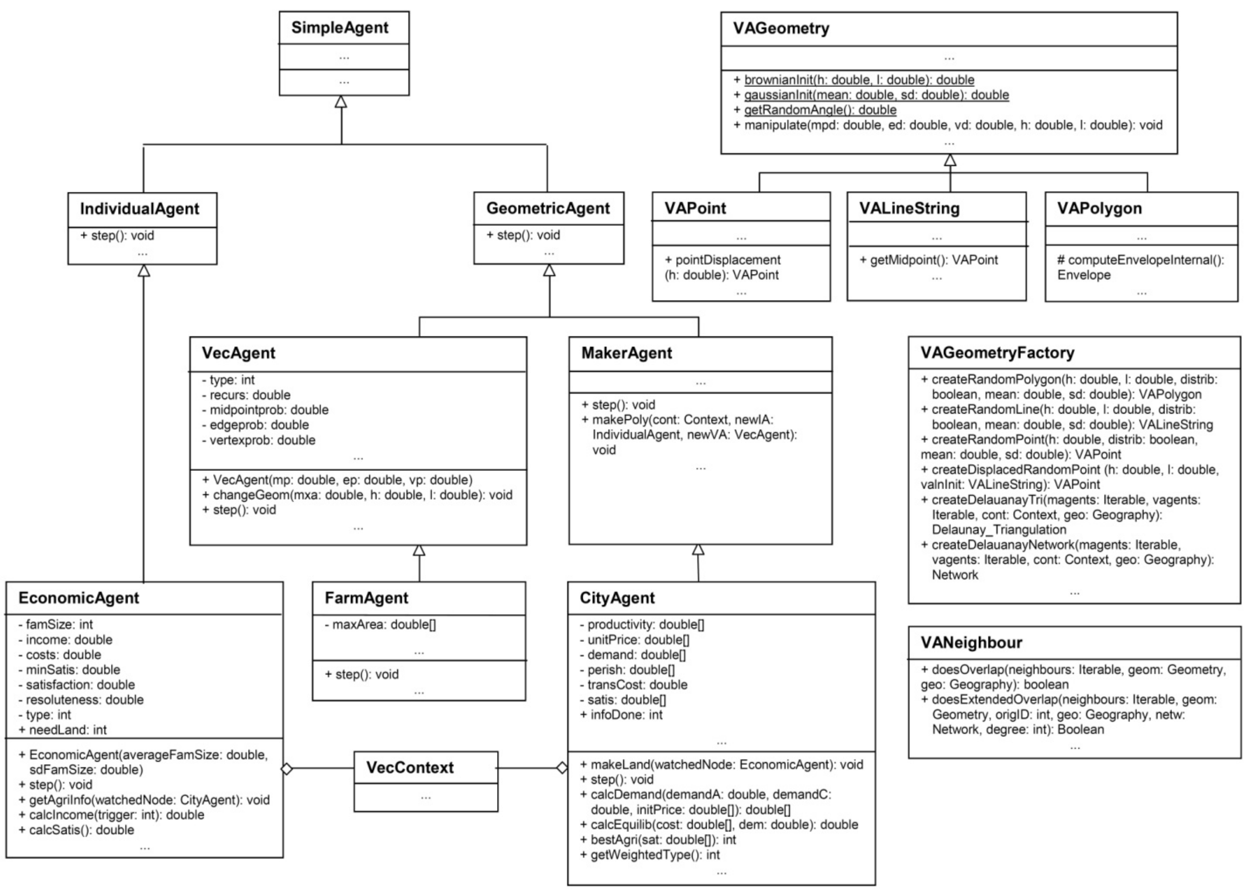


Figure 4: Simplified UML diagram of the GVA Von Thunen simulation.

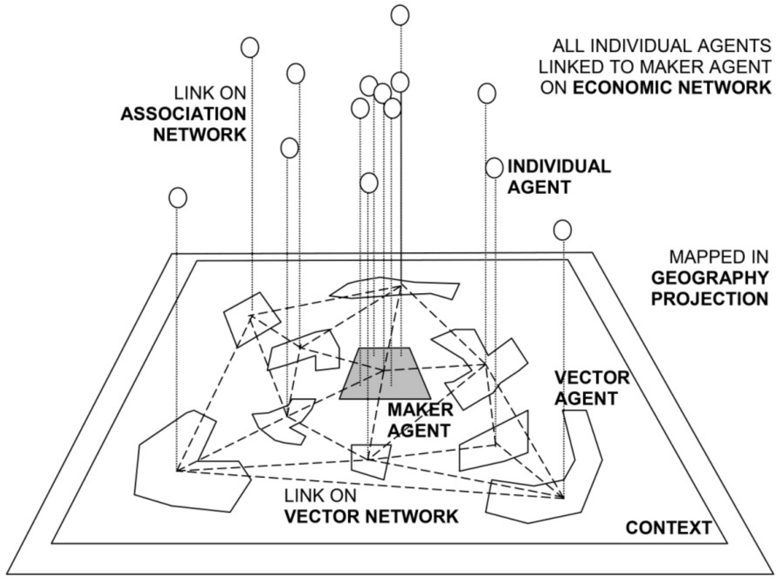


Figure 5: Schematic of the relationship of the context, the four projections and the agents that inhabit them.

A GEOGRAPHICAL VECTOR AGENT MODEL FOR VON THÜNEN'S AGRICULTURAL LAND USE THEORY

Von Thünen's theory

Johann Heinrich von Thünen developed his basic agricultural land use model of the relationships between market, production and distance in 1826 (Hall, 1966) [17]. The relative costs of transporting different agricultural products to a central market determined the agricultural land use around that market. The most productive activities (*e.g.* intensive agriculture, forestry) would compete for the closest land, pushing out the less productive activities (*e.g.* grain farming, livestock farming).

In terms of spatial geometry, this should lead to the development of "... fairly sharply differentiated concentric rings or belts [which] will form around the town, each with its own staple product" (Hall, 1966, p.8) [17], though Lösch (1954) [23] has said that the formation of concentric zones is not inevitable with Von Thünen's assumptions. The concentric rings are ordered in distance from the city market according to their relative perishability and transportation cost. Therefore the order will be intensive agriculture (eg perishable fruits, vegetables and dairy), forest resources, grain farming then livestock farming (i.e livestock have the ability to transport themselves, thus minimizing transport cost). Fig. **6** illustrates the theoretical spatial arrangement of concentric agricultural land use for the isolated case.

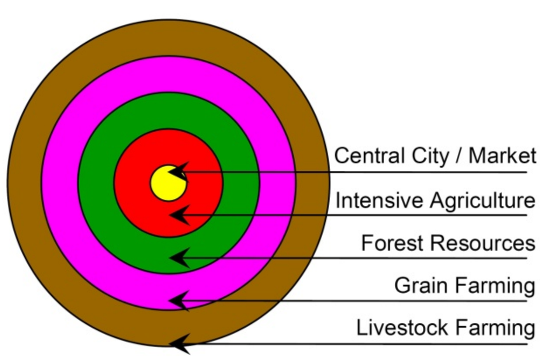


Figure 6: The concentric ring pattern of von Thünen's agricultural land use model.

The model rests on a set of assumptions:

- there is one isolated market in an isolated state having no (trade) interactions with the outside
- the land surrounding the market is entirely flat and its fertility is uniform (*i.e.* it is isotropic)
- there are no transport infrastructures such as roads and rivers (though von Thünen did investigate the modification of transport cost and ensuing land use patterns due to the introduction of a navigable river to the model), the farmers transporting their produce to market (the type of product will also affect transportation cost)

The production costs, market price and transportation cost can be related in the following land rent equation (presented here in modified form) developed by Dunn (1954, cited in Henshall, 1967) [24]. This is an attempt to convert von Thünen's largely descriptive theory into a normative model:

$$E = Y/(p - c - rd) \quad \textbf{(3)}$$

Where E = rent per unit of land, Y = yield per unit of land, p = the market price per unit of yield, c = the average production costs per unit of yield, r, the transport costs per unit weight per unit distance from the market and d represents distance from the market.

Since the late 1950s / early 1960s there have been numerous other schemes to quantify von Thünen's theory as a process. Garrison and Marble (1957) [25] presented a partial equilibrium von Thünen model that can be solved by simultaneous equations. The first computational model was published by Stevens (1968) [26], calculating optimal locations while generating equilibrium location rents. Day and Tinney (1969) [27] developed a dynamic and discrete model, which Okabe and Kume (1983) [28] extended to accommodate continuous modelling (incorporating a demand function) and output land use patterns for two types of good. They found that land use is generally unstable, sometimes changing cyclically in 2-year intervals. Along with Kanagaya's (2003) [29] Generalised Thünen model, the continuous modelling approach recognises that in reality, many land uses may co-exist in the same space.

Recently, there have been a few projects to capture the process behind von Thünen's theory in an agent-based modelling environment. Macmillan and Huang (2008) [30] developed an agent-based model of an agricultural society on a grid-based landscape, which can produce classic von Thünen patterns.

Most directly relevant to this research is Sasaki and Box's (2003) [31] work, using von Thünen to verify the use of agent environments for land use simulation. In this research, cellular automata and an agent-based model were the approaches used (in the Swarm environment, an alternative agent modelling framework). Turton (2003) [32] investigated the use of intelligent multiagent models for von Thünen on a cellular spatial structure, using an earlier version of Repast. More recently, Millington *et al* (2008) [33] implemented a Mediterranean style agricultural model in the NetLogo environment

For our implementation, the agent structure and parameters used by Sasaki and Box (2003) [31] will be referred to extensively, only using a vector geometry approach in a Repast environment.

The GVA von Thünen structure

Our von Thünen simulation is driven by the action of three sets of agents, the FarmAgent (an extension of the VecAgent), the CityAgent (which extends the MakerAgent class) and the EconomicAgent (extending the IndividualAgent class). Refer back to the Fig. **4** UML diagram.

The CityAgent represents the market in a von Thünen economy but also has a geometric footprint, is responsible for the creation of FarmAgents and "houses" all city-dwelling EconomicAgents. Specifically, it:

- initialises and stores values for productivity of, unit price for, demand for, perishability of, and unit transport cost per distance unit per, each agricultural good type; and the overall satisfaction of EconomicAgents farming that type.

- Gathers information on satisfaction and costs accrued from each EconomicAgent and synthesises it
- Calculates demand and equilibrium price for each good type and also the overall satisfaction of the EconomicAgents farming that type, at every increment
- "broadcasts" this information so that every EconomicAgent can potentially make use of it
- Creates a new FarmAgent and initiates its dynamic geometry or allocates an empty FarmAgent for any EconomicAgent needing it
- Forges and severs network links of EconomicAgents and FarmAgents as necessary (via the AssociationNetwork)
- Is implicitly responsible for selling the various product types that EconomicAgents produce

The FarmAgent is the pure geometric embodiment of the GVA. As such it has no direct equivalent in von Thünen theory, but it is anticipated that groups of individual FarmAgents of a specific agricultural type will form a concentric ring pattern around the market or approximation thereof. Specifically, the FarmAgent:

- Stores its own single agricultural type (out of Intensive Agriculture, Forestry, Grain Farming and Livestock Farming) and the maximum area that a farm of that type can have
- Stores the probabilities governing the method of boundary manipulation as well as roughness and recursivity parameters
- Manipulates the boundary at each increment, provided that the FarmAgent area is less than the maximum allowed area and is not empty
- Also manages Delaunay neighbourhood information to ensure that there are no polygonal overlaps in the process of manipulation

The EconomicAgent represents the human decision maker in the system. As with the FarmAgent, there is no direct von Thünen equivalent (and it has no geometric manifestation) but such agency of individuals is something that agent-based modelling facilitates. Specifically, the EconomicAgent:

- Represents a family that (via the AssociationNetwork) is either linked to the CityAgent or is linked to a sole family occupancy FarmAgent and, via the EconNetwork, is permanently linked to the CityAgent
- Initialises and stores its own properties of income, costs, satisfaction and resoluteness
- Adopts and stores the agricultural type of the FarmAgent it is linked to (otherwise it is coded as a city dweller)
- is responsible for collecting current agricultural market information from the CityAgent
- at every increment, will calculate its own income and satisfaction
- will make a decision on what to do (city dweller or farmer), based on its own state and what the CityAgent is telling it (*i.e.* it calculates its rent based on equation 3 and uses this information in the decision)

The GVA von Thünen model simulation run

The aim of the simulation is to demonstrate the use and effect of irregular and dynamic GVA in a simple theoretical scenario. The hypothesis is that using these start-up parameters the end result will be a concentric ring pattern, rings in order of distance from market, being intensive agriculture (I), forest resources (F), grain farming (G) and livestock farming (L).

The start up parameters for the model are specified in Tables 1 and 2. As mentioned before, unless the parameters are geometry specific, they have been adapted from Sasaki and Box's 2003 [31] CA model, with relative values for parameters largely kept the same:

(Demand: $F \approx G > I \approx L$; Initial Unit Price: $I > L > G > F$; Perishability: $I > F \approx G > L$; Productivity: $F \approx I > G > L$; Maximum Area: $L > G > F > I$).

Given the importance of geometry in a GVA implementation, the relevant parameters have been separated into their own table, including spatial parameters directly from von Thünen (unit transportation cost) and Sasaki and Box (*e.g.* number of economic agents, which has an implicit but profound spatial effect). The maximum area for each of the agricultural land uses have also been adapted from Sasaki and Box, though importantly the area relates to a single shape, not a group of cells, as would be the case with CA.

Table 1: Model parameters and their values: these are the parameters relating to the *economic* aspects of the model

Variables	Value	Description
averageFamilySize	5	Average Family Size
stDevFamilySize	1	Standard Deviation of Family Size
incomeCityDwellers	300	Income for City Dwellers
initialDemandI	300	Initial Demand for Intensive Agriculture
initialDemandF	800	Initial Demand for Forestry Resources
initialDemandG	800	Initial Demand for Grain Farming
initialDemandL	300	Initial Demand for Livestock
initialUnitPriceI	800	Initial Unit Price of Intensive Agriculture
initialUnitPriceF	300	Initial Unit Price of Forestry Resources
initialUnitPriceG	500	Initial Unit Price of Grain Farming
initialUnitPriceL	700	Initial Unit Price of Livestock
perishabilityI	0.7	Perishability of Intensive Agriculture
perishabilityF	0.5	Perishability of Forestry Resources
perishabilityG	0.5	Perishability of Grain Farming
perishabilityL	0.3	Perishability of Livestock
productivityI	1.2	Productivity of Intensive Agriculture
productivityF	1.5	Productivity of Forestry Resources
productivityG	0.9	Productivity of Grain Farming
productivityL	0.5	Productivity of Livestock
resoluteFam	0.1	Resoluteness of Families

Table 2: Model parameters and their values: these are the parameters that have the most influence on *geometric* configuration of the model results, either directly or indirectly

Variables	Value	Description
unitTransCost	50	Unit Transportation Cost
numeconagents	100	Number of Economic Agents
maxAreaI	2	Maximum Area Allowed for Intensive Agriculture
maxAreaF	8	Maximum Area Allowed for Forestry Resources
maxAreaG	14	Maximum Area Allowed for Grain Farming
maxAreaL	20	Maximum Area Allowed for Livestock
propEmptyFarms	50	Percentage of Empty Farms (probability of reoccupation)
maxxcoord	100	Maximum X Extent
maxycoord	100	Maximum Y Extent
probmidpointdisplacement	0.4	Probability of Midpoint Displacement
probedgedisplacement	0.5	Probability of Edge Displacement

probvertexdisplacement	0.1	Probability of Vertex Displacement
roughnessH	0.8	Roughness parameter (h)
recursivityInterval	0.001	Recursivity Interval
initDispersal	2	Initial Polygon Dispersal
neighDegree	2	Degree of Neighbourhood

The displacement probabilities, roughness and recursivity interval directly affect the invocation of equations for midpoint, vertex and edge displacement (equations 1 and 2) as well as parameterise those equations. The neighbourhood degree specifies the amount of Delaunay links from a FarmAgent that constitutes its neighbourhood – for example, if the degree is 2, then any FarmAgent or CityAgent that can be reached in two links or less is a neighbour (in this simulation it is used to limit the check for neighbour overlap as the FarmAgent expands).

Another geometric parameter of interest is the probability that an empty farm will be occupied by a new farmer (expressed as a percentage). This was introduced as a realistic addition facilitating dynamic model runs with more changes in farm occupancy as a result. Finally, the polygon dispersal parameterises a normal distribution with its standard deviation, governing the spread of FarmAgent initialisation, as measured with distance from the CityAgent. Fig. **7** illustrates the workings of the simulation. Initially, the context VecContext is created and populated with the Geography and three Network projections. A single CityAgent and n=100 EconomicAgents are initialised. Since the CityAgent has a geometric footprint it is added to the Geography perspective. Since the EconomicAgents are resident in the CityAgent but have no geometric presence, they are each linked with the CityAgent through an AssociationNetwork link yet not added to the Geography perspective. Once EconomicAgents start farming, they will be linked to their own FarmAgent on the Association Network. They also maintain a permanent link with the CityAgent via an economic network (EconNetwork) for all economic transactions.

The two agent types initially present in the VecContext now run their transition rules at each iteration. The CityAgent gathers information from each EconomicAgent, calculates synthesised market information (demand and price for each agricultural type) from the gathered resource, then broadcasts the information, incrementing the infoDone variable as it does so.

The EconomicAgents will not have fired until this point, as they need the market information in order to decide what to do in the next round. A Repast watcher annotation on infoDone tells the EconomicAgents when to act (*i.e.* the moment infoDone changes in value). The income for each EconomicAgent is then calculated. For city dwellers, the income is a constant rate of 300. For farmers, the income is calculated from the following algorithm, which is adapted from equation 3.

if EconomicAgent is linked on the AssociationNetwork to a FarmAgent **then**

revenue = price [agricultural_type] * productivity [agricultural_type] * (1 + FarmAgent.area)

costs = proportionLabourForce [agricultural_type] +

(FarmAgent.distanceFromCity * unitTransCost) * perishability [agricultural_type]

income = revenue – costs

else if EconomicAgent is linked on the AssociationNetwork to the CityAgent **then**

income = incomeCityDwellers

end if

The EconomicAgents will now compare their adjusted satisfaction (which is equal to income) with their minimum allowed satisfaction, choosing to remain where they are (city or farm) if the threshold is met. If the agent is a farmer and the threshold is not met, then they will vacate their land and return to the city. This will mean removing their link with their FarmAgent on the Association Network and forging a new link with the CityAgent on the same network. If the EconomicAgent is linked to the CityAgent and wants to be a farmer then the needLand variable is incremented. This is another watcher variable that triggers the CityAgent into action on the Economic Agent's behalf.

The above account is initially what happens incrementally. However, as seen above, the CityAgent may need to acquire land for an EconomicAgent wishing to farm. They can do this by first checking for any empty land (FarmAgent) left by previous unsuccessful farmers. If there is no empty land or the CityAgent was unable to acquire empty land (governed by propEmptyFarms), then the CityAgent will initiate a new FarmAgent (with the probability parameters for future midpoint, edge and vertex displacement). It will then try to place the farm in empty space at a distance randomly extracted from a normal distribution (parameters defined by the mean of the distances of existing FarmAgent centroids from the centroid of the CityAgent and the predefined polygon dispersal standard deviation value). If it tries to site the new farm on an already existing farm, it will try again until it succeeds. Then the geometric procedure shown pictorially in Fig. **2a-d** will be followed, leaving a triangular farm shape to be geometrically manipulated in future iterations.

The FarmAgent is also linked with its farmer EconomicAgent (and its tie with the CityAgent severed on the AssociationNetwork); they are both also assigned their agricultural type code (randomly, but biased towards the more successful agricultural types). The new FarmAgent is also linked to its geometric (and topological) neighbours through the VANetwork. In future iterations, the FarmAgent will attempt to manipulate its own boundary, with the predetermined probabilities of the method of manipulation. The provisos to this are:

- empty land does not change its own boundaries
- FarmAgents with a prospective area greater than the maximum area for its agricultural type will not change
- FarmAgents will not overlap with its neighbours (defined by Delaunay neighbourhood to a degree of 2); it will repeatedly try to change its boundary until it succeeds:

```
if FarmAgent.area <= maximumArea [agricultural_type] then
FarmAgent.geometry.manipulate {prob_midpoint_displacement, prob_edge_displacement, prob_vertex_displacement,
    roughness, recursivity_interval}
    for each adjacent FarmAgent on the AssociationNetwork
    // Neighbourhood degree = 1
    if FarmAgent.geometry overlaps with adjacentFarmAgent.geometry then
        reject new FarmAgent geometry
        END
    else
        // Neighbourhood degree = 2, actually handled as a recursive function
        for each adjacent FarmAgent on the AssociationNetwork
            if FarmAgent.geometry overlaps adjacentFarmAgent.geometry then
                reject new FarmAgent geometry
                END
            else
                // Neighbourhood degree = 3, and so on (actually handled as a
                // recursive function)
            end If
        loop
    end if
    loop
    Accept change in geometry
```

else

Do not change geometry

end if

END

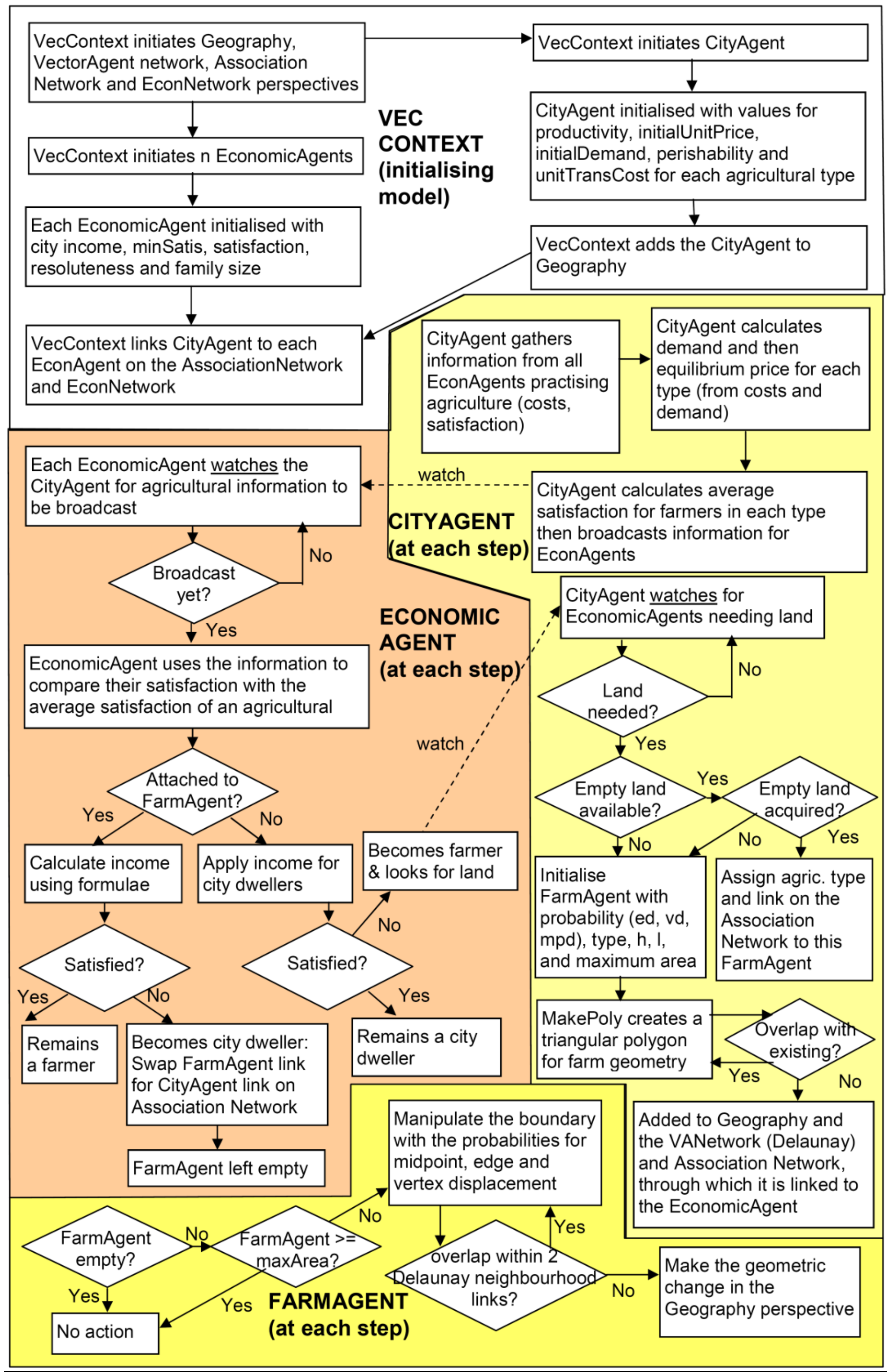


Figure 7: The steps in the GVA Von Thünen simulation.

RESULTS

Like Sasaki and Box [31], the results shown in Fig. **8** were arrived at by experimenting with different combinations of parameters. Emphasis was not on the values of the parameters themselves but in ensuring that they remained the same relative to each other (given the theoretical nature of this simulation). Model runs varied between being too dynamic, where farms changed type at every increment and too static, where farm types locked in at an early stage. An optimal level between the two was sought (Sasaki and Box characterises this as a balance between chaos and order, employing a parameter to indicate a model run's place on that continuum), with the emphasis of parameter change on the geometric drivers, to try and achieve the concentric ring land use pattern shown in Fig. **6**.

We can see that the pattern evolved in Fig. **8** has similarities to the idea of specific land uses specializing at certain distances from the city, while not forming a full concentric ring pattern in itself. In this typical run there is an initial farmland rush for all agricultural types. The FarmAgents themselves exhibit rapid growth until about 100 increments, when the advancement of the farm development front starts to slow down (the number of EconomicAgents in the city are exhausted; also the farm maximum areas for each agricultural type are reached), stabilising completely by 200 increments.

Two key phenomena occur to bring this pattern about – normally between 50 to 100 increments, the outermost intensive agriculture farms are purged, leaving a higher density of such farms closer to the city. Second, the emphasis on larger area for livestock farmers (and to a lesser extent grain farmers) and the room to expand serves to stabilise these types of farms towards the edge of the farm distribution. In comparison, forestry resource agents appear to lead a stable existence throughout.

The boundary form of the polygons themselves appear somewhat convoluted, as FarmAgents try to grow, even in a confined space. As a result, the geometric arrangement appears virtually tessellated at the global scale beyond 500 increments. Also at this stage in the simulation, an increasing litter of small (mostly triangular) empty farms becomes apparent, mostly visible at the edges of the distribution.

Out of the simulation runs with the parameters recorded in Tables 1 and 2, this is the pattern evolution type that is most common. Other runs lead to weaker patterns (when compared to the von Thünen ideal) but are less common.

The graphs in Fig. **9** back this description up. Fig. **9a** shows the average distance of FarmAgents of each land use type from the central CityAgent. After an initial fluctuation the basic order becomes established a little after 200 increments and is stable thereafter. The order is as expected of a von Thünen spatial representation though there is a reversal of the expected average distance of forestry and grain farming. The average (distance*area) graph in Fig. **9b** establishes the correct order. This graph also follows the same fluctuation-stability pattern. For the distance graph, the occurrence of empty land is at an increasing magnitude from the CityAgent. This reflects the increasing tessellation of areas closer to the CityAgent and decreasing opportunities for empty land to occur in these areas.

Fig. **9c** shows the averaged percentage of Delaunay neighbours that are of the same type and is meant to be an indication of homogeneity of areal clusters, another indicator that a contiguous pattern is being reached. It is calculated through the following formula:

$$d = (1 / n) \sum_{i=1} (m_{ti} / m_i)*100 \quad \textbf{(4)}$$

where:

n = number of FarmAgents

m_{ti} = number of Delaunay neighbours for FarmAgent i of same agricultural type t

m_i = total number of Delaunay neighbours for FarmAgent i

This graph shows a steady increase in this proportion over the simulation run, supporting the argument of contiguity. However, given the results in Figures 9a and 9b establishing stability soon after 200 increments, the subsequent increase in contiguous neighbours should be put down to the increasing density of empty farms accumulating in the peripheral areas of the simulation space.

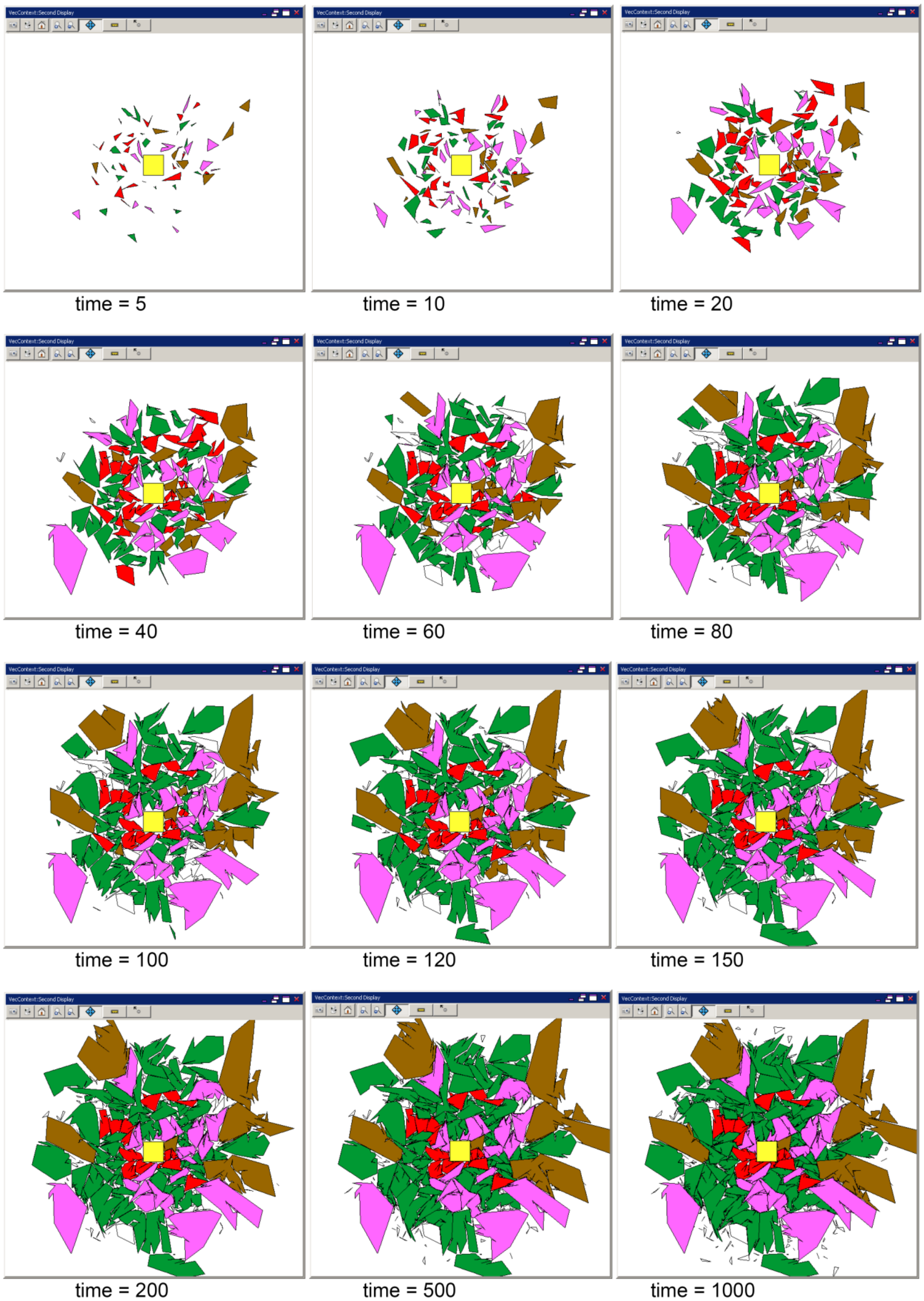


Figure 8: The von Thünen GVA output for the first 1000 steps (see Figure 6 for legend; white = empty farms)

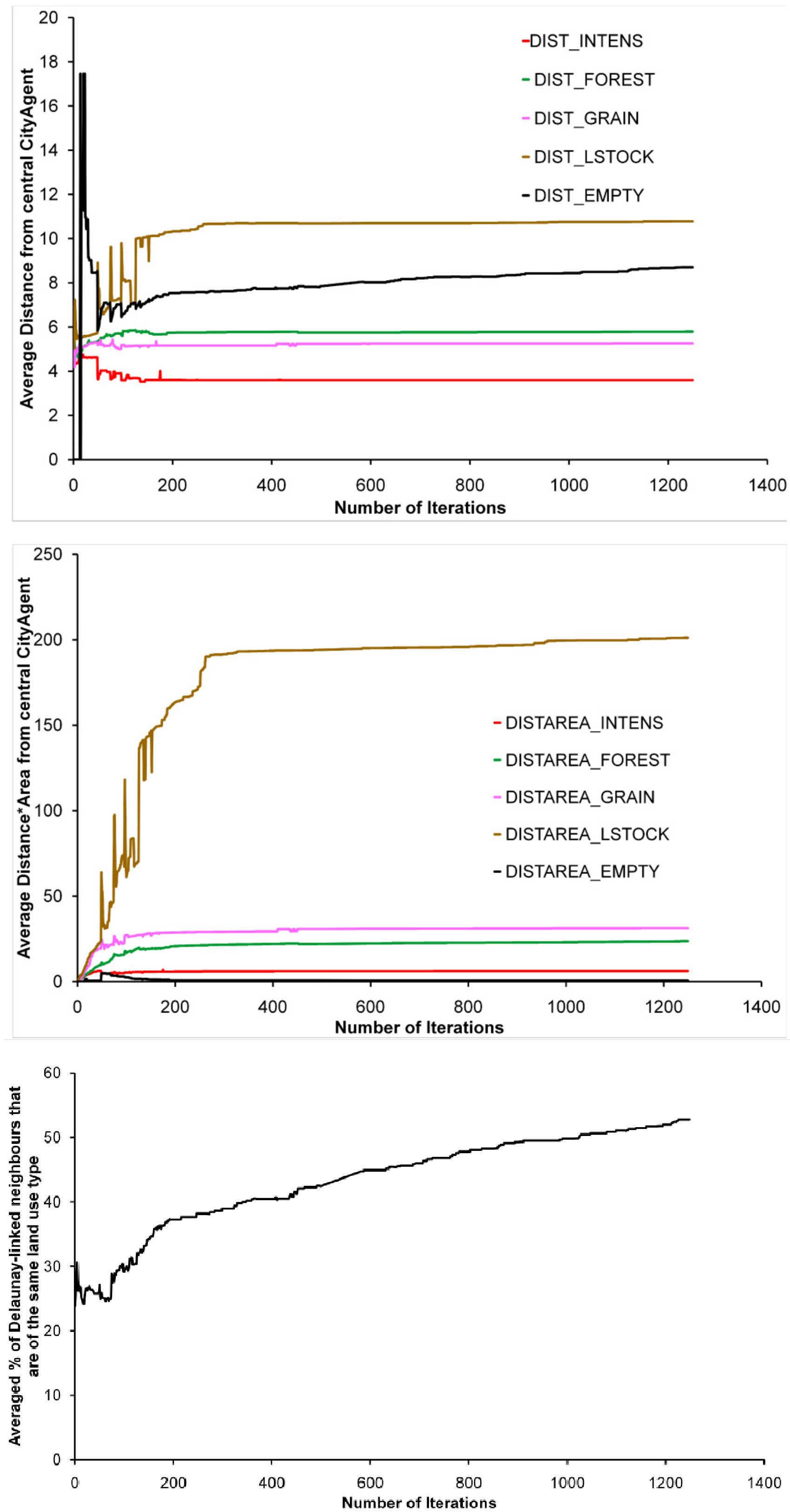


Figure 9: Graphs of simulation performance over time (in increments). **(a)** Average distance of each land use type from the central CityAgent **(b)** Average distance*area of each land use type from the CityAgent **(c)** Averaged percentage of Delaunay-linked neighbours that are of the same land use type (overleaf).

DISCUSSION AND CONCLUSION

Geographical Vector Agents (GVAs) have been introduced as a geometry-led form of spatial modelling (Hammam *et al*, 2007) [3], a vector-based version of cellular automata that has irregular objects with dynamic boundaries instead of static cells as units of space. This chapter documents the next stage in the development of GVA, a move from a simple implementation using fractal geometry as the mechanism of boundary manipulation (*i.e.* simple geometric rules to create realistic real-world shapes) to adapting the thinking of that initial stage for deployment in the Repast agent modelling environment. Repast offers a structure that can accommodate dynamic vector geometry, placing it in geographic space in parallel with network projections that enable representation of topological, land attachment and economic relationships. This augmented version of a GVA environment has introduced two types of spatial agent to interact with GVAs, a MakerAgent that can initialize VecAgents as well as perform other management tasks and an IndividualAgent, which adopts a role similar to that of a free-roaming agent in a conventional multi-agent system simulation. These three types of agent in their mapped and networked environment have been extended to model a basic theory of agricultural land use (von Thünen), using domain specific parameters (adapted from a von Thünen CA simulation developed by Sasaki and Box, 2003) [31] but mostly through altering more generic geometric controls. The model results give visual evidence that the desired concentric ring pattern of land use is forming in part and in the correct order from the city, though forestry resources and grain farming are reversed in order (Wu has stated that, "…in land use simulation, it is rare that a simulation can match exactly the real-world pattern" 1999, p. 202) [34]. This points to the general usefulness of the model but it should be remembered that these model results are not intended to contribute to the agricultural land use domain per se as a predictive tool, but to demonstrate the simulation of a simple approximation of a real-world phenomenon and in so doing test the applicability of the model's architecture.

The results have been verified through quantitative methods such as graphing the average distance of each agricultural land use type and measuring the changing averaged proportion of contiguous neighbours, but there are various other issues and areas of improvement to be highlighted. One of the reasons for the successful pattern is due to a specific vector property, indivisible area. The larger areas of the livestock and grain land uses make them have a more stable presence. Although CA would have this area subdivided (as with Millington *et al*, 2008) [33], making for a less integrated and less realistic simulation it remains the more efficient format for spatial modelling – due to the numerous polygons generated, particularly if there is a high proportion of empty farms (and the ones that are not empty have their boundaries continually increase in complexity), the GVA simulation is very computer hungry and quickly slows right down. Strategies that could be employed to mitigate this include removing empty farms that have been lying dormant for a while and stopping boundary change for a VecAgent after a given number of failed tries (if, for example the agent is trapped). Perhaps when this happens the agent can snap to the neighbouring polygons, making for a true tessellation. Snapping is just one of the geometric processes that can potentially be applied in a GVA context in addition to fractals – depending on the function of the model being implemented, any transformation or static / dynamic spatial pattern could be realised.

There is a role played by the prolonged non-availability of some empty land, which feeds back into the model through additional changes in land use type. This may lead to another unpredictable effect – a farm with an area larger than the maximum area for its type (*e.g.* a former livestock farm of large area has turned to intensive agriculture). Other minor modifications that could be made to this specific model include a tighter coupling of the geometric and economic elements, say through making geometry or rate of growth / area linked to economic well-being. As an example, this approach may be linked with a transformation not implemented here, say that an object can have its area reduced due to a downturn in economic fortune (as opposed to leaving the object empty, as is currently the case). Practically, this would be simple to implement, involving the removal of a check for area decrease from the current programming code.

Finally, the role of neighbourhood is underplayed, merely being used as a localized overlap check. One way in which neighbourhood could be used is to promote stability of contiguous groups of farms of the same type through collaboration, or instability through competition. A neighbourhood search could also reveal empty farms nearby that could be occupied by a farmer wanting to expand. Millington *et al* (2008) [33] has their farmer agents owning multiple farms, with penalisation applied if the owned farms are too fragmented. The authors also implement multiple cities / markets which is another possible initiative.

What has been demonstrated in this paper is (to borrow from the GAS definition) the successful interplay of geometry, geometric rules, states, transition rules and to a lesser extent, neighbourhood (but not neighbourhood rules). Beyond tweaking this theoretical GVA model example, the main future challenge is to implement GVA in Repast for a real world scenario (*e.g.* real agricultural land use, urban growth and differentiation). An added factor implicit in this is having to handle real geographic data. Also in keeping with true conditions, the development of models where more than one type of spatial behaviour is being played out in the model space (*i.e.* different agents having different rules) is a priority.

ACKNOWLEDGEMENTS

The author would like to thank Yasser Hammam, Peter Whigham and Claire Freeman for their contribution to the development of GVAs.

REFERENCES

[1] M. F. Goodchild, 'Geographical information science', International Journal of Geographical Information Systems, vol. 6, no.1, pp. 31-45, 1992.
[2] I. Benenson and P. Torrens, 'Geosimulation: Automata-Based Modelling of Urban Phenomena', Wiley, 2004.
[3] Y. Hammam, A. Moore, and P. Whigham, 'The dynamic geometry of Geographical Vector Agents', Computers, Environment and Urban Systems, vol. 31, no.5, pp. 502-519, 2007.
[4] M. Luck, M. D'Inverno, and S. Munroe, 'Autonomy: Variable and generative', in Agent Autonomy, H. Hexmoor, C. Castelfranchi, and R. Falcone, eds., Kluwer Academic Publishers, pp. 9-22, 2003.
[5] P. Torrens and I. Benenson, 'Geographic Automata Systems', International Journal of Geographic Information Science, vol. 10, no.4, pp. 385-412, 2005.
[6] M. Batty, 'Geocomputation using cellular automata', in GeoComputation, S. Openshaw and R.J. Abrahart, eds, Taylor and Francis, pp. 95-126, 2000.
[7] M. Batty, 'Cellular dynamics: modelling urban growth as a spatial epidemic', in GeoComputational Modelling: Techniques and Applications, Advances in Spatial Science, M.M. Fischer and Y. Leung, eds, Springer, pp.109-141, 2001.
[8] P. Torrens and D. O'Sullivan, 'Cities, cells, and complexity: Developing a research agenda for urban GeoComputation', in Proceedings of the 5th International Conference on GeoComputation, University of Greenwich, London, [Electronic Resource - CDROM], 2000.
[9] L. Sanders, D. Pumain, H. Mathian, F. Guerin-Pace, and S. Bura, 'SIMPOP: A multi-agent system for the study of urbanism', Environment and Planning B, vol.24, pp.287–305, 1997.
[10] M. Batty, J. Desyllas, and E. Duxbury, 'The discrete dynamics of small-scale spatial events: agent-based models of mobility in carnivals and street parades', International Journal of Geographical Information Science, vol. 17, no.7, pp. 673-697, 2003.
[11] J. Wahle, L. Neubert, J. Esser, and M. Schreckenberg, 'A cellular automaton traffic flow model for online simulation of traffic', Parallel Computing, vol.27, pp.719–735, 2001.
[12] W. Shi and M.Y.C. Pang, 'Development of Voronoi-based cellular automata: An integrated dynamic model for geographical information systems', International Journal of Geographical Information Science, vol. 14, no.5, pp. 455-474, 2000.
[13] F. Semboloni, 'The growth of an urban cluster into a dynamic self-modifying spatial pattern', Environment and Planning B, vol. 27, no.4, pp. 549-564, 2000.
[14] D. O'Sullivan, 'Graph-based cellular automata models of urban spatial processes", Unpublished PhD Thesis, Bartlett School of Architecture and Planning, University College London, University of London, London, UK, 2000.
[15] D. O'Sullivan, 'Graph-cellular automata: A generalised discrete urban and regional model', Environment and Planning B, vol. 28, pp.687-707, 2001.
[16] N. Moreno, F. Wang, and D. Marceau, 'Implementation of a dynamic neighbourhood in a land-use vector-based cellular automata model', Computers Environment and Urban Systems, vol.33, no.1, pp.44-54, 2009.
[17] P. Hall, 'Isolated State: an English edition of der Isolierte Staat', Pergamon Press, 1966.
[18] R. Laurini and D. Thompson, 'Fundamentals of spatial information systems', Academic Press, 1992.
[19] N.C. Kenkel and D.J. Walker, 'Fractals in the biological sciences', COENOSES, vol. 11, pp.77–100, 1996.
[20] R. F. Voss, 'Fractals in nature: From characterization to simulation", in The Science of Fractal Images, H.O. Peitgen and D. Saupe, eds, Springer-Verlag, pp.21-70, 1988.

[21] T. Howe, N. Collier, M. North, M. Parker, and J. Vos, 'Containing agents: Contexts, projections and agents', in Proceedings of the Agent 2006 Conference on Social Agents: Results and Prospects, Argonne National Laboratory, 2006.

[22] Repast. Recursive Porous Agent Simulation Toolkit, 2010. [Online] Available: repast.sourceforge.net/index.html. [Accessed Nov. 30th, 2010].

[23] A. Lösch, 'The economics of location' (trans. W.W. Wogblom), New Haven, 1954.

[24] J. Henshall, 'Models of agricultural activity', in Socio-economic models in geography, R.J. Chorley and P. Haggett, eds. Methuen, pp.425-458, 1967.

[25] W.L. Garrison and D.F. Marble, 'The spatial structure of agricultural activities', Annals of the Association of American Geographers, vol.47, pp.137 - 144, 1957.

[26] B. Stevens, 'Location theory and Programming models: the von Thünen case', Papers of the Regional Science Association, vol. 21, pp. 19-34, 1968.

[27] R.H. Day and E.H. Tinney, 'A dynamic von Thünen model', Geographical Analysis, vol. 1, pp. 137-151, 1969.

[28] A. Okabe, and Y. Kume, 'A Dynamic von Thünen model with a demand function', Journal of Urban Economics, vol. 14, no. 3, pp. 355-369, 1983.

[29] K. Kanagaya, 'Generalised Thünen and Thünen-Ricardo models for Asian land use', in Modelling geographical systems - Statistical and computational applications, B. Boots, A. Okabe, R. Thomas, eds., Kluwer, pp. 181-199, 2003.

[30] W. Macmillan and H.Q. Huang, 'An agent-based simulation model of a primitive agricultural society', Geoforum, vol. 39, pp. 643-658, 2008.

[31] Y. Sasaki, and P. Box, 'Agent-based verification of von Thünen's location theory', Journal of Artificial Societies and Social Simulation, vol. 6, no. 2, paper 9, 2003. [Online] Available: jasss.soc.surrey.ac.uk/6/2/9.html. [Accessed Nov. 30th, 2010].

[32] I. Turton, 'Modelling land use development using multi-agent systems', in Proceedings of the 7^{th} International Conference on GeoComputation, University of Southampton, UK, [Electronic Resource - CDROM], 2003.

[33] J. Millington, R. Romero-Calcerrada, J. Wainwright, and G. Perry, 'An agent-based model of mediterranean agricultural land-use/cover change for examining wildfire risk', Journal of Artificial Societies and Social Simulation, vol. 11, no. 4, paper 4, 2008. [Online] Available: jasss.soc.surrey.ac.uk/11/4/4.html. [Accessed Nov. 30th, 2010].

[34] F. Wu, 'GIS-based simulation as an exploratory analysis for space-time process', Journal of Geographical Systems, vol. 1, pp. 199-218, 1999.

CHAPTER 4

Advances and Techniques for Building 3D Agent-Based Models for Urban Systems

A.T. Crooks[1,*], A. Hudson-Smith[2] and A Patel[2]

[1]Department of Computational Social Science, Krasnow Institute for Advanced Study, George Mason University, United States of America and [2]Centre for Advanced Spatial Analysis (CASA), University College London, United Kingdom

Abstract: There is a growing interest in relating agent-based models to real-world locations by combining them with geographical information systems (GIS) which can be seen with the increase of geosimulation models in recent years. This coincides with the proliferation of digital data both in the two and three dimensions allowing one to construct detailed and extensive feature-rich and highly visual 3D city models. This chapter explores some of these developments in relation to our own initial work on building 3D geospatial agent-based models of urban systems and the technologies that allow for such models to be created. These range from coupling agent-based models with 3D visualization, to building 3D agent-based models in 3D animation and rendering packages, and to using 3D virtual worlds for the creation of agent-based models.

INTRODUCTION

Agent-based modeling (ABM) is increasingly being used as a tool for the spatial simulation of a wide variety of urban phenomena including: housing dynamics [1], urban growth and residential location [2], gentrification [3] and traffic simulation [4]. At a more microscopic level, there are agent-based models that simulate pedestrians in urban centers [5] and crowd congestion [6]. These applications demonstrate a growing interest in linking agents to actual places and with geographic data (see [7] for a review) through the linking or coupling with geographical information systems (GIS) and agent-based models. This coupling allows agent-based modelers to simulate agents related to actual geographic locations and to think about how objects or agents and their aggregations interact and change in space and time. This focus of linking agents to real-world locations are often referred to as geosimulation models [8].

As agent-based models move evermore into the spatial domain, there is a need for new ways to explore, visualize and communicate such models especially to those who we seek to influence or where such models might inform decision makers. This has already been identified as one of the key challenges facing ABM [9]. Not only does this relate to the notion that good models, which generate spatial or physical predictions that can be mapped or visualized must 'look right' [10], but it also relates to one of the major purposes of agent-based models, which is to visually convey the behavior of the model clearly and quickly [11]. This is supported by [12] who write that "…visualization is one of the most effective ways to present key model information to decision-makers (p 280)." 2D visualization of agent-based and cellular automata models is common place such as the animation of model results of land-use change [13], [14] which allows users to see the dynamic behavior of recognizable characteristics of model results rather than just exploring models through data and statistics. However, just as in GIS, often the visualization of agent-based models is ineffective [11], which hinders their communication to those we seek to influence.

One potential way to visualize and communicate agent-based models is to utilize the third dimension using advances in computer hardware, software and networked communication. This third dimension is rarely ventured into academic agent-based models [15], [16][1]. We would argue that this is due to several reasons. First, it has to do with the nature of the discipline where the focus is on theory rather than outreach and end user visualization, unlike that of computer games and movies. Secondly, until recently there has been a lack of 3D data to construct such models

***Address correspondence to Andrew Crooks:** Department of Computational Social Science, Krasnow Institute for Advanced Study, 379 Research 1 Building, MS 6B2, George Mason University, Fairfax, VA 22030, USA; Tel: (001)(703) 993-4640; Fax: (001)(703) 993-9290; E-mail: acrooks2@gmu.edu

[1] Further information about this work can be seen at http://www.redfish.com/wildfire/ and http://www.redfish.com/stadium/

Danielle J. Marceau and Itzhak Benenson (Eds)

that are easily accessible to the modeler; however, this is changing (which we discuss below). Thirdly, most agent-based model builders outside of computer science are not fully taking advantage of the improvements in computer graphics and processes, networked communication and associated technology.

The intention of this chapter is to explore the recent advances in computer technology, software and associated techniques that allow for the creation of 3D agent-based models which can be used to simulate various aspects of city life, focusing on our own initial research of creating 3D cityscapes and 3D agent-based models. In the remainder of this chapter, we therefore explore our attempts at using digital data to create feature rich 3D cityscapes (Section 2), discuss why such cityscapes are important for ABM (Section 3), before moving into how advances in computer hardware allow for the creation of 3D agent-based models (Section 4); we then briefly explore a potential application domain, that of pedestrian modeling (Section 5). Section 6 presents techniques that we are currently utilizing to create 3D agent-based models through various linking and coupling approaches along with advantage and disadvantages of each approach before a discussion is presented (Section 7). However, a caveat is first needed, that is the communication of agent-based models is not only through visualization. For example, it can be done through the provision of modeling source code or by providing executables of models, as advocated by [17] and [18] but which is rarely done. Models can also be communicated through their logical and consistent descriptions [19]. Nevertheless the focus of this chapter is the communication of agent-based models through 3D visualization.

RISE IN DIGITAL DATA

ABM generally focuses on how micro-scale interactions of many individuals result in the emergence of more aggregate patterns. However, to create a real-world environment for these agents to inhabit, one needs fine scale and potentially extensive digital geometric data sources. Many of the applications highlighted in Section 1 utilize such data as a foundation for the environment (albeit in a 2D world). Such data might include a terrain for agents to walk over, buildings for the agents to live in, locations of business for them to work at, footpaths and roads for their travel, etc. The last decade has seen a proliferation of fine scale data sources becoming increasingly available at finer and larger extents and being coupled with height data as can be seen with the growing number of 3D digital city models (see [20] for further information). Such 3D models are a result of the integration of computer-aided design (CAD) software, GIS, computer graphics, web and aerial sensing technologies. Notwithstanding their application in ABM, such digital data sources have other applications ranging from urban planning, telecommunications, architecture, facilities and utilities management, property analysis, marketing, tourism and entertainment (see [21] for a review). The development of web and virtual globe technologies has given a massive boost to digital urban models, enabling widespread access and interaction by the public through geo-browsers such as the popular Google Earth.

Whilst the visualization capabilities of 3D city models are clear as we highlight in Fig. **1**, their analytical functionality is often underdeveloped [22]. Significant advances have been made in increasing the geometrical sophistication of 3D city models, yet many models remain 'empty shells' without any socio-economic data associated with the buildings, or the capability to analyze the role of the built environment in urban processes, which we consider a major hindrance for the creation of 3D ABM. However, we believe that future advances will explore how such models can be populated with socio-economic data and linked to transportation networks, thus moving from visualization to focus on policy applications and analysis, and acting as a foundation of 3D worlds for agents to inhabit. For example, [20] link the empty shells of the buildings with residential and commercial property information for all the buildings within the Greater London Authority which is similar to the work of [23]. There is much synergy in these aims with those of planning support systems (PSS), which provide tools to aid and enhance planning tasks [24]. However, incorporating multiple datasets into 3D city models also has its challenges in relation to handling and visualizing the shear amount of data that are required [25].

Such detailed data are important for the creation of agent-based models, but also for the visualization of model results. Fine scale social and built environment data have already been used for the creation of agent-based models [1]. This is often in 2D, but with the proliferation of 3D data and environments to visualize them, it is possible to use them as a backdrop for 3D agent-based models (this is discussed further in Section 6). This relates in a sense to urban modeling more generally. [26] defined urban models on a continuum between the iconic and the symbolic. Iconic models are physical versions of the 'real' thing but normally scaled down. Typical traditional examples include the architects' block model as we show in Fig. **2** and **2d** cartographic maps. Symbolic models represent systems in terms of the way they function, often through time and over space. Such models replace the physical or material system by some logical and/or mathematical formula, often in the form of algebraic equations within a digital form (*e.g.* a computer) such as in the case of land-use transport models (*e.g.* [27]). However, the distinction between the two is increasingly being blurred as technology advances. For example, the iconic representation of the city within a 3D GIS such as the Virtual London Model [28] is a digital manifestation of the architects' block model that we showed in Fig. **1** that can potentially acts as a container in which symbolic models are run, which we demonstrate in Section 6.

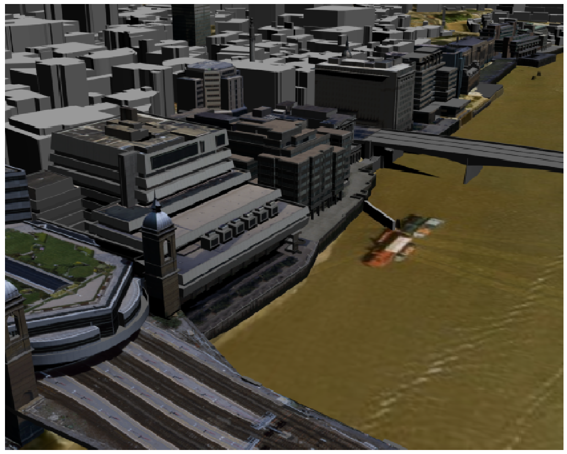


Figure 1: Blackfriars railway station and the river Thames, central London.

Combining the iconic with the symbolic models in a digital media gives us unprecedented power to understand, simulate, explore and visualize cities, especially when combined with agent-based models [29]. This was not possible hitherto and it coincides with the way we currently conceptualize and model cities. This has changed from the aggregate to disaggregate and from the static to the dynamic, taking ideas from complexity science. ABM provides tools to explore this change in approach. Specifically it allows us to explore the reasoning based on which individual decisions are made and how such decisions lead to more emergent structures evolving. This has potentially great benefits for urban design. Take for instance planners and architects who are increasingly being challenged to develop innovative uses for spaces but who do not know how people will use such spaces. Combining both the architect model with an agent-based model on how pedestrians might use space or exit a building in an emergency can potentially improve the design of buildings. Combining the symbolic and iconic models therefore could potentially improve the design process through embedding pedestrian dynamics in the related architectural structures [30], given that human movement behavior has deep implications on the design of effective pedestrian [31].

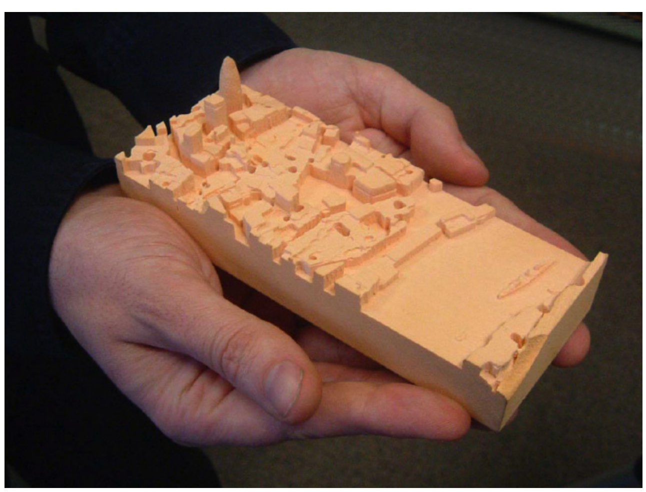


Figure 2: 3D architects block model portraying part of central London.

3D ENVIRONMENTS AS A BACKGROUND FOR SIMULATION

The previous section presented how there is an increasing amount of rich digital data to build 3D city models but only tentatively addressed why this might be of use for the creation of agent-based models. In this section, we will briefly sketch out why the use of the third dimension is a useful tool for ABM. Just as agent-based models can model real-world environments through the use of GIS [8] thus allowing us to map their outcomes spatially, digital 3D environments hold the ability to create a sense of place, and thus to mimic the real world for the purpose of digital planning [32]. This relates to the concept of 'legibility' in urban planning. Legibility is used to refer to the ease with which inhabitants of a city can develop a cognitive map over a period of time and so orientate themselves within it and navigate through it [33]. For example, [33] writes "nothing is experienced by itself, but always in relation to its surroundings (p 1)." This is the reality of the city, the built environment. Therefore, if for the purpose of digital planning and ABM, we are to replicate the built environment in digital space, the space itself must convey to the user a sense of location, orientation and identification; in short, it must convey a sense of place [32]. To gain an essential foothold, people have to be able to orientate themselves, to know where they are. But they also have to identify themselves within the environment, that is to know that they are in a certain place [34]. Through the rise of digital data, GIS and CAD technologies it is now possible to create such environments within computers (however, internal structures of buildings is problematic). We believe that there is therefore a clear outreach and knowledge creation mechanism to applications here. Moving into the 3D realm has the potential to provide 'windows' into the complexity of phenomena and environments under investigation [32]. This is particularly important as agent-based models occur in space; they are inherently structured in two (latitude and longitude), three (position above or below the Earth's surface), or four (time) dimensions.

Not only do 3D city models provide this sense of place but by animating the agents in these models, we can give them realistic appearances. This is particularity appealing when modeling the individual in such applications of pedestrian modeling [35]. The visualization of people moving allow us to better convey situations such as pedestrian movement and allow urban planners to improve the structure of the street network and communicate the idea of space to the public [36]. Additionally, allowing for the third dimension to be incorporated into models allows us to augment such models within the real world. For example, [37] explore the use of virtual cities as a test bed for examining the design of urban public spaces. Specifically, the authors combined an agent-based model with a virtual city model (in this case a platform at the Kyoto subway station) and used augmented reality to allow humans to interact with the agents (as if the agents and the humans were in the same crowd) through the use of positioning sensors around the station; then, they simulated an emergency. Combining agent-based models with 3D graphics not only allows us to carry out experiments which are not easy to do in reality, such as setting a building on fire, but also provides a sense of place which people can relate to. It can therefore potentially help communicate such models to people not familiar to ABM[2].

While there are clear benefits for linking agent-based models to 3D environments, many may consider 3D agent-based models as 'glorified' computer games or just 'eye-candy'. Perhaps this relates to the issue that many spatial problems can be treated in 2D such as finding the highest point on a terrain. Whilst essentially this is a 3D problem for each location has a height, a 2D surface is sufficient to find the highest point. However, there are problems such as line of sight or the spread of smoke which would benefit from the third dimension, not to mention giving agents more realistic appearances and giving them a sense of place (as discussed above). Even if we relate such models to computer games, 3D visualization in computer gaming has a lot of potential for agent-based models. Take for example, SimCity [39][3], a city-building simulation game whose objective as the name suggests is to build and design a city. The player can own land (*e.g.* commercial, industrial, or residential), add buildings, change taxes along with building transportation systems, and respond to disasters such as flooding or earthquakes. Scenarios within the simulation can be based on real cities and problems associated with them. For example, in the original model the Swiss capital, Bern in 1965 was one such scenario where at the time the roads were clogged with traffic; the mayor (*i.e.* a player) needed to reduce traffic and improve the city by installing mass transit systems. In a sense, such a game

---

[2] On a side note, it is not just urban systems that can benefit from moving into the third dimension but also ecosystems such as the study of river basin management, for example [38].

[3] SimCity is not really a true 3D model, but a 2.5D as it uses an isometric viewpoint. The player navigates the environment by scrolling left, right, up or down. It gives a sense of 3D without having a z axis (see [40] and [32] for more details).

provides a valuable teaching tool for urban geography, planners, designers and policy makers [41], because while it is a game it has business rules, ecosystem modeling, and social dependencies. The graphical user interface (GUI) of the game facilitates the learning about the complex, dynamic, and interrelated nature of urban problems.

However, there is a difference between agent-based models and such games. That is agent-based models have relatively simple visualization but deep behavioral content while within games, agents tend to have superficial behavior but very enriching graphics. Within this paper we are not advocating 3D agent-based models for just the sake of it but as a means for explaining the model to non modelers. For example, one solution is the release of SimCity under the name of Micropolis [42], an open source project therefore allowing developers to add more complex behaviors and rules to the model by editing and expanding the code base.

The true advantage of 3D in these models is however difficult to gain. User testing is obviously required and this will be part of our future research in a similar vain to that of those who test 2D and 3D user interfaces [43]. Nevertheless, some of the main simulation toolkits are starting to explore the third dimension for visualizing model outcomes. For example, some 2D toolkits are starting to integrate 3D authoring environments into the system, most notably StarLogo TNG [44]. However, there are limitations to such software in relation to geospatial research especially when the source code of the models is not available [7]. Others, while having a 3D component such as NetLogo [45] are essentially 3D visualization of objects on a 2D plain, therefore being still 2D models. Repast Simphony [46] allows for the integration of 3D objects and terrains such as the National Aeronautics and Space Administration (NASA) virtual globe Whirl Wind [47]. This move into 3D has been facilitated by developments in computing as associated software, specifically Java 3D and computer processing units which we now briefly turn to.

MOVING TOWARDS 3D: ADVANCES IN PROCESSING UNITS

The ability to visualize and model entire cities on a computer not only relates to availability of data but also to developments in computing in general, specifically to how computer processors have developed at exponential rates doubling approximately every two years [48]. This is especially the case for the central processing unit (CPU) which is at the heart of the computer, and whose job is to execute a collection of machine instructions that in turn tell what to do in terms of computation[4]. Generally speaking, 2D visualizations and simulations have been traditionally carried out on the CPU; however, moving to the third dimension requires moving the rendering of graphics onto the graphics processing unit (GPU). The GPU is a specialized processor optimized for accelerating graphics, and it offloads all the 3D graphics rendering from the CPU. The processor specifically uses most of its transistors to perform floating-point calculations which are fundamental to 3D graphics rendering. The development of the GPU and more generally graphic cards has been motivated by advances in games. It has become a fairly complex and specialized device that has allowed for the exploration of applications in different areas, which will be briefly explored below. While the graphics card plays an important role for graphic designers, and 3D animators, more recently the GPU has been utilized by scientists to perform computations that are beyond computer graphics. This technique is often referred to as general-purpose computing on graphics processing units or GPGPU for short [50]. In relation to ABM, not only does the GPU allow for the creation of 3D visualization and rendering of agent-based models but through the GPGPU there is a potential to run agent-based models containing millions of agents much faster than general purpose CPU programs such as those employed in ABM toolkits such as Repast [46] and NetLogo [45]. For more information pertaining to the use of the GPGPU for ABM the reader is referred to [51].

However, utilizing the GPU for ABM is not a trivial task. Users have two main options: the first is to write their own software code, the second is to use an existing piece of software. For the first option, programming languages such as C#, C++ or Java can be used to write software that can use graphic libraries such as DirectX [52] and OpenGL [53]. These are two of the main types of graphics libraries that can be used to write programs that provide instructions to the GPU to perform complex tasks. While the second lower entry option is to use existing software such as 3ds Max [54], a modeling, animation and rendering package developed by Autodesk that provides the capability to utilize the GPU for rendering outputs (which we discuss briefly in Sections 5 and 6).

[4] For further information about the CPU the reader is referred to its entry in Wikipedia [49].

3D ABM APPLICATION DOMAINS

2D agent-based models have been developed to study a wide range of phenomena; however, here we will focus on one application domain, that of pedestrian modeling, and trace how moving to the third dimension can aid such models. Pedestrian modeling itself is a wide domain. For example, there are models exploring the evacuation of buildings [55], to that of movement in shopping areas or art galleries, and crowds [56]. These models generate valuable insights into such events, and demonstrate how the action of many individuals results in more aggregate structures emerging. For example, in crowds, the agents themselves are often represented as dots or squares. We are able to validate such models through datasets of the real world equivalent to the agents to be modeled, or by using human observers to collect data.

Compare such models with those from first person video games such as Crysis [57] or fight scenes in Lord of the Rings between Orcs and Humans using purpose-built software called Massive [58] with which thousands of agents can be programmed to make decisions such as to defend, to retreat if out numbered, *etc.* While such models are highly visual representations in 3D and look semi-realistic, they do not focus on behavior *per se*. Within such models, be it from computer games or movies, behaviors are often homogeneous or have limited heterogeneity, while 2D social science applications tend to focus more on the latter. The reason for this, within movies, is that the realism of behavior is not a great priority, as compared to realism of the agents, such as characters in crowds. This kind of modeling is focused on agents in movies, computer game productions, or virtual worlds / environments. Here, the rendering of the agents is the point of focus, and animating in a believable manner is sufficient.

There are a host of tools available to create such crowd simulations, not only Massive but also Autodesk's 3ds Max, Maya [59], and Legion [60]. These tools have been used in the movie and games industry for years, and are high-end computer animation and artificial intelligent software packages. However, a recent trend has been the convergence of this high quality visualization along with the realism of behavior, where systems that are visualization oriented are trying to incorporate better behavior, and vice versa. While agent-based models need to incorporate behavior, adding high quality 3D visualization has its advantages over simple 2D in terms of representation. Specifically, as adapted from [61]:

- the model is easier to understand, and gives a visual reference of the location;
- it gives a general feel for the environment, and shows how the environment will look, in addition to the ABM simulation;
- it enhances communication of ideas; a good example is the redesign of the Oxford Circus interchange in central London [62];
- it helps spot obvious errors in the model;
- it makes ideas more accessible to others, that otherwise may not understand them; for example, planning proposals to councils; and
- it helps the user use his / her intuition in understanding a system.

More specific pedestrian applications that might benefit from the third dimension include the movement of pedestrians in complex structures, such as multi-floor buildings that are often spread over several floors with interconnecting stairways (as briefly discussed in Section 2), like office blocks and shopping malls, which are not 2D flat planes. The ability to move from the two to the third dimension allows us to simulate more realistic movement within such buildings but also to explore different land uses (mixes of land use) within the same model (in this example a building). For example, retail on the ground floor and residential on the upper floors cannot be easily visualized or modeled in 2D. This is not to say that 2D models of complex buildings are of no use, but that moving to the third dimension allows for more realism and flexibility within models especially when combined with advances in geometric and non geometric digital datasets (*e.g.* [63]). In relation to complex buildings, there is also the need to model evacuation scenarios in order to model overall evacuation performance. By taking into consideration the design of a building, in many cases, with several floors, a 3D simulation can highlight the impact the design of a building will have on individuals exiting it [64]. These types of models are seen with many of the commercial pedestrian modeling software packages such as Legion [60] and STEPS [65].

TECHNIQUES AND TOOLS FOR CREATING 3D AGENT-BASED MODELS

Thus far we have discussed how the rise in digital data can help create 3D models along with how advances in computer hardware are allowing us to create agent-based models within the third dimension. This section explores methods on how one goes about creating 3D agent-based models using existing software packages, or through linking/coupling agent-based models with 3D visualization software. These include creating an agent-based model within a specific software environment such as utilizing commercial game engines, which allow for the development of ones own models. Examples are provided by Unreal [66] and Crysis [57] that allow for the importation of 3D buildings and associated spatial data. For example, in Fig. **3** we show how the Virtual London 3D city model can be imported into the Crysis game engine. Alternatives to commercial game engines are their open source equivalents such as OGRE [67] or Panda3D [68], where models can be created using programming languages such as Python or C++. As with open source ABM toolkits such as Repast, the key advantage of open source games engines relates to the transparency of their inner workings. The user can explore the source code, permitting the modification, extension and correction of the system if necessary. This is particularly useful for verifying a model [7]. A further advantage of using this approach is the ability to focus on modeling rather than spending time developing parts of the model that are not content specific such as user interfaces.

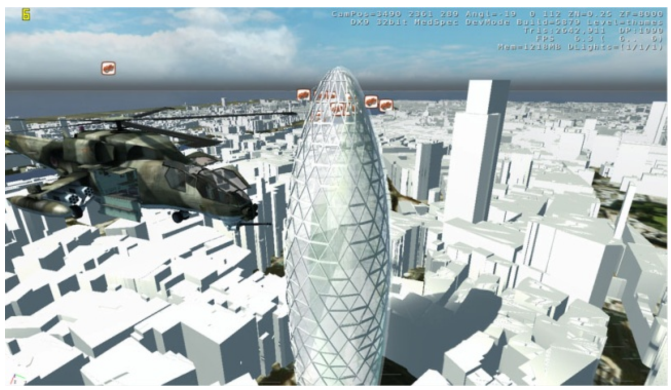


Figure 3: The Swiss Re building and the City of London imported into the Crysis game engine.

Rather than directly embedding the agent-based model into a software environment, the coupling of an agent-based model to 3D visualization software might be the preferred option, especially if the model has already been built. Such coupling can therefore be broadly defined as the linkage of two stand-alone systems by data transfer. This coupling can be seen on a continuum between loose and tight coupling as seen within GIS [7]. Loose coupling can be seen as the transferring of files, as we demonstrate in the NetLogo example in Section 6.2. Here, NetLogo is used to execute the model and store outputs of agent locations, which are then transferred to the 3D software for visualization. Tight or close coupling is characterized by the simultaneous operation of systems allowing direct inter-system communication during the program execution. For example, [69] have created agent-based models in Java, communicating with Second Life *via* remote procedures XML-RPC requests to the Second Life server. Second Life is only used to collect and display model information.

Tight-Coupling: Creating 3D Agent-Based Models in 3Ds Max

While not being a specialist agent-based visualization package *per se*, 3ds Max does have a built in simulation system. Known as ‘Crowd and Delegate’, it allows groups of 3D objects and characters to be animated using a system of delegates and behaviors. Using a series of simple rules, either those built in such as ‘avoid,’ ‘follow,’ ‘seek’ or custom written in the form of scripts, one can create crowds with highly complex behaviors. Fig. **4** illustrates the most basic level of agent-based model in 3ds Max using the ‘follow surface’ and ‘wander behavior’. Each delegate, represented as the blue triangles, forms part of a team assigned rule to wander for a set period of time in a random manner while following a 3D surface, in this case created to simulate a terrain. Once set in the scene, a time element is added for the simulation with the delegates location solved for each key frame (akin to a time step) of movement. Using the crowd and delegate method, it can also be extended to the simulation of schools of fish or people which can be quickly and easily achieved to a high level of visual realism [70].

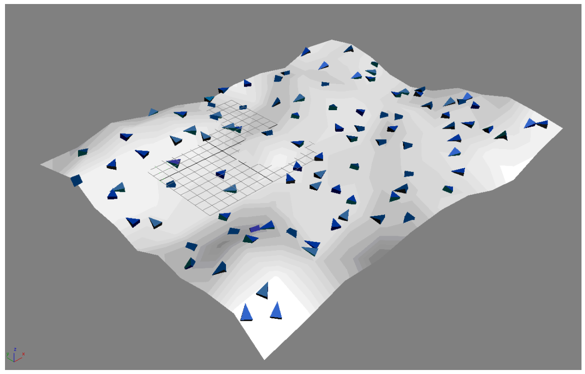


Figure 4: 3ds Max crowd and delegate system: follow surface and wander behavior.

Models created in such software not only utilize advances in graphic card technology, but also advances in physics based engines (such as Havok, [71]) which allow us to easily add additional elements such as mass and gravity to influence the agents' behavior. For example, in Fig. **5** we extend a basic flocking algorithm to model pedestrian movement that also includes avoidance of vehicles, in this case a bus, and frame the model in a 3D cityscape. Other simple agent-based models within 3ds Max include simple ant like behaviors to simulate shockwaves within traffic akin to [72] traffic example. The various built-in components of 3ds Max enables high quality graphic outputs as well as real time previews and outputs to game engines such as Crysis. This allows researchers to achieve 'semi movie like' results. However, as with all movie clips and demonstrations out of standard 3D packages they need to be taken with a pinch of salt. The science is there and the simulations are realistic but the science is hidden and not produced by the agent-based modeler but by the package itself as it is essentially a 'black box.' Since the inner workings are often hidden, it potentially makes these packages of limited value on their own, which makes us turn our attention to loose coupling approaches.

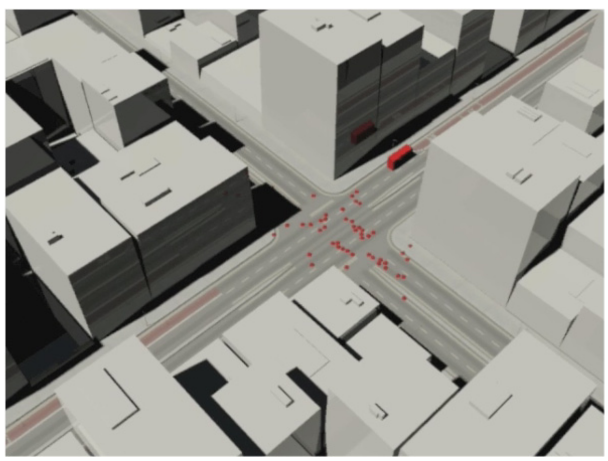


Figure 5: Pedestrian agents and a vehicle agent within a cityscape created in 3ds Max.

Loose Coupling Netlogo with 3Dds Max

Loose coupling provides an attractive alternative, in the sense that we can create an agent-based model using a specific programming library or use a dedicated simulation/modeling toolkit designed specifically for ABM and then visualize the outputs from the model in a 3D environment (thus the 3D scene is purely for visualization purposes unless there are x, y and z coordinates directly incorporated in the modeling process). In this instance we use NetLogo, a simulation/modeling system for the modeling and 3ds Max for the visualization. As a proof of concept we take a simple traffic model, [73] from NetLogo as shown in Fig. **6** (top panel), which models the movement of cars over a street network. Movement is restricted by traffic lights; agents stop at red lights and move on green. In order to achieve a physical three dimensional representation of the environment, the movement of the cars in NetLogo is translated into text files by recording their movement at each iteration (tick) of the model. Along with recording the coordinates of the cars, the coordinates of the road patches, and the green and red turtles (traffic lights) are stored for each tick for a total of 500 ticks. These data are then read into 3ds Max through a script. The script takes all the coordinate information from the cars movement, the traffic light states and the road patches. Key frames are first created, steps are then taken to animate and render the scene as we show in Fig. **6** (bottom panel). The process of linking NetLogo to 3ds Max is shown in Fig. **7**. Further information including a tutorial can be found in [74].

This approach has the potential of creating high end visualizations of geographically explicit agent-based models especially as NetLogo supports the integration of geographic datasets. Another example of this loosely coupled approach is by [75] who combined a pedestrian model from NetLogo, which explored human behavior to room configurations, where the outputs were visualized with 3ds Max. The ability to import coordinates into these systems means that 3D cityscapes created in CAD and GIS packages can be populated with agents from other models. For example, outputs from large scale traffic models such as MATSim [76] could be visualized in 3ds Max providing a sense of location and place, which non modelers could relate to.

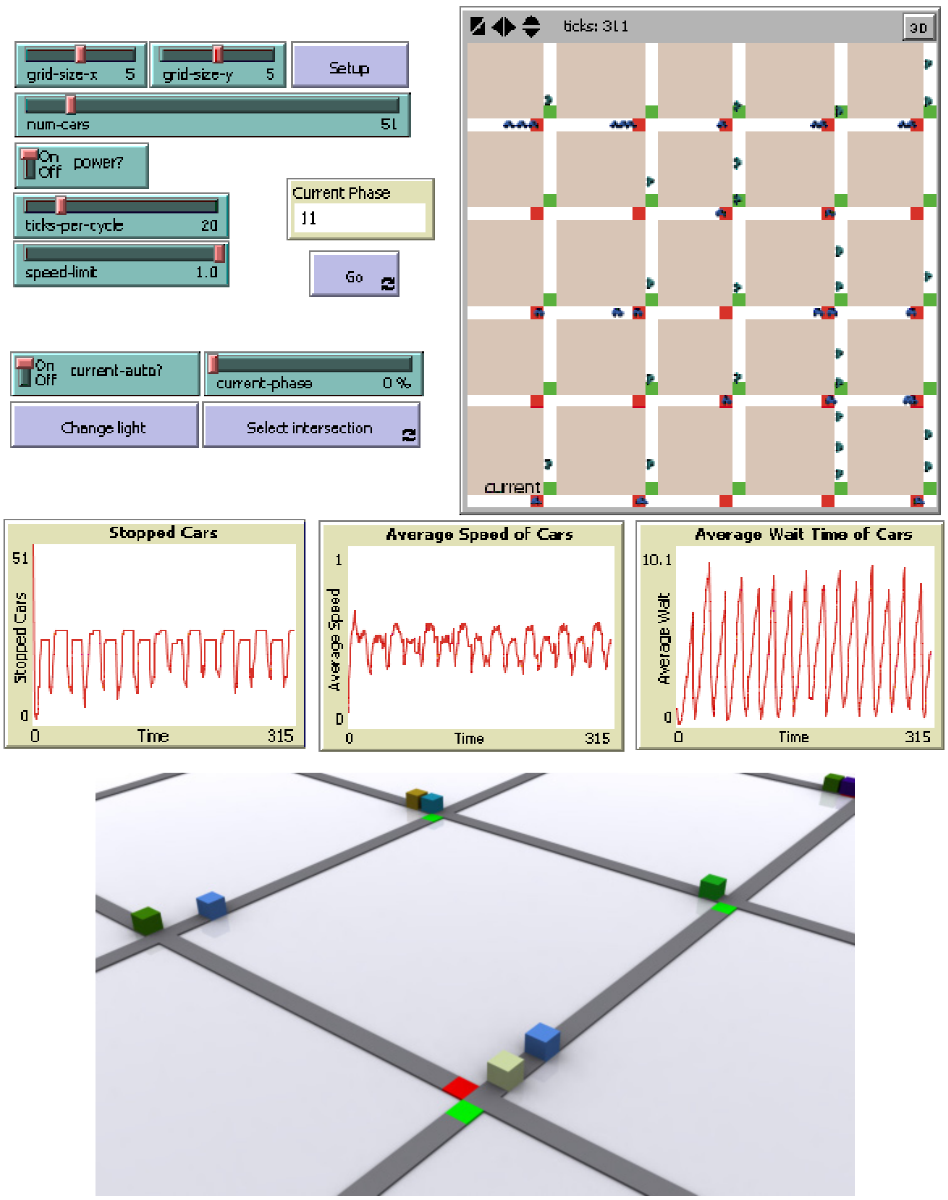


Figure 6: 3D Visualization of NetLogo traffic simulation: (top panel) NetLogo Traffic Simulation; (bottom panel) Cars on top of roads, with red and green traffic light within 3ds Max.

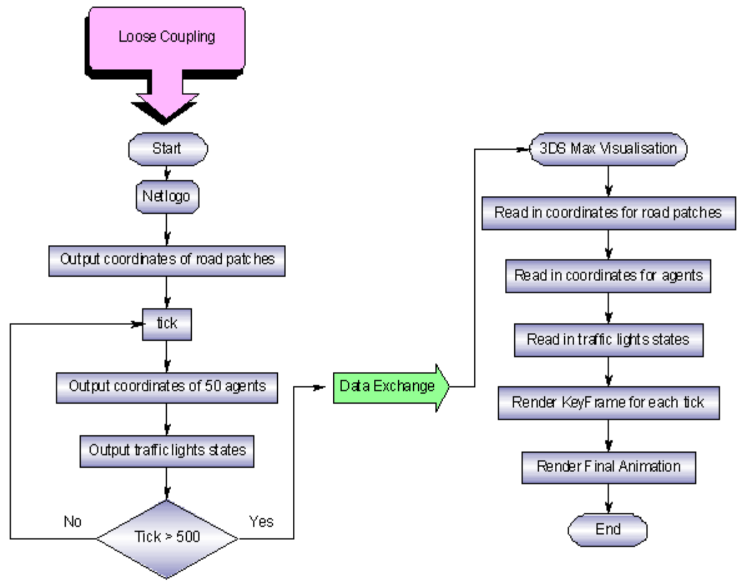


Figure 7: Loose coupling between the two standalone systems.

Agent-Based Models in Virtual Environments

With respect to visualizations, the models presented above can be broken down into two categories. The first is fly-through, where the creator has set up a prescribed flight path (viewing angle) that the viewer cannot deviate from (for example, those in 3ds Max mentioned in Sections 6.1 and 6.2. The second is interactive visualizations, where the viewer has control on how and where to view the simulation from (such as shown in Fig. **3**). It is to these we now turn to as this represents immersing the user into the modeling environment.

With improved graphics and processes, networked communication and associated technology has led to the rise of interactive content through Web 2.0 technologies. These technologies have lead to users expecting a more interactive experience over the internet [77] and we would argue that this is the case for modeling as well especially when embedding such models in virtual worlds. Virtual worlds offer such an experience as they allow users to explore areas and interact with the content that interest them. However, finding ones way through such worlds can be a difficult task [78]; these problems can be overcome if one considers legibility (as discussed in Section 3), which refers to the ease at which inhabitants can develop a cognitive map over a period of time and thus can orientate themselves and navigate through space. The use of actual buildings and 3D cityscapes could greatly facilitate such navigation and understanding of agent-based models directly related to spatial locations. For example, the Unity [79] multi-platform game development tool allows us to embed models into rooms as highlighted in Fig. **8**. The room itself is created in SketchUp [80] and the models in 3ds Max. In essence these are just table-top models but one can extend such models into virtual environments, for example, in the virtual world of Second Life to which we now turn.

Figure 8: Agent-based models displayed in unity.

Agent-based models are usually considered as forming a miniature laboratory where the attributes and behavior of agents, and the environment in which they are housed, can be altered, and experimented with, where their repercussions are observed over the course of multiple simulation runs. Virtual worlds such as Second Life act in a similar way to agent-based models *i.e.*, they are artificial worlds populated by agents. The idea behind such systems is to engage a community of users where people represented as avatars can be active users contributing to sites and participating in site content in real time through the world wide web which opens their use to whoever is connected. Such worlds can potentially be used as online laboratories – collaboratories [81], for example, where model building and users engage in mutual and shared development activities, although their infancy are very much on the horizon.

Virtual worlds such as Second Life have great potential for research in the social and behavioral sciences along with offering an environment for education and outreach [82]. These systems allow people to discuss and visualize models in real time; they provide an effective medium to clearly communicate models and results between the developer and the decision maker which in the past was the sole province of powerful scientific workstations. For agent-based modelers it offers a unique way for the exploration and understanding of social processes by means of computer simulation. Researchers have used agents within virtual worlds to study a variety of phenomena from human-to-agent interaction [83], the study of norms between agents and avatars [84], healthcare issues [85], to herding behavior [69]. We are using Second Life as a collaborative geographic space [86] for the dissemination of geographic content and for the exploration of agent-based models in an interactive 3D media.

Within this world we have created a number of agent-based models using the Linden Scripting Language [87] as we show in Fig. **9**. It is the purpose of these models to act as pedagogic demonstrators and as a 'proof-of-concept', thus we have chosen Conway's Game of Life and Schelling's [88] segregation model. These models were selected to highlight how classical automata styles of models that have inspired a generation of modelers can be created and explored in Second Life [89]. The third model we have created is a prototype pedestrian evacuation model, which is not only more complex than the first two, but highlights how more complex models can be created and be linked to actual buildings as we show in Fig. **10**.

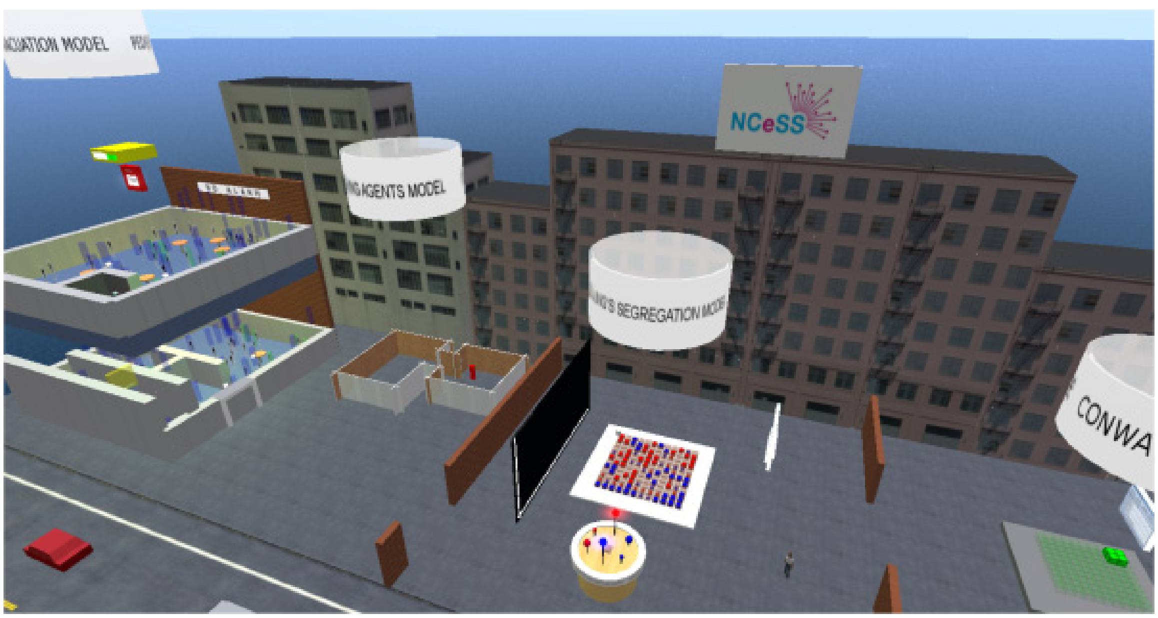


Figure 9: Agent street: Agent-based models in Second Life.

Agents within the evacuation model have been designed to mimic 'real' people with realistic anthropomorphic dimensions that exit a building when an alarm is sounded. We represent the building (enclosure) as a continuous space opposed to the more common regular lattice (grid) or coarse network enclosure representations [90] which are common in 2D pedestrian models. Therefore agents are not restricted to discrete cells nor represented as flows thus enabling us to simulate pedestrian movement more explicitly in the x, y and z dimensions. The agents within the model interact with each other and their environment (*e.g.* obstacle avoidance) both of which can have an effect on occupant movement. For example, agents adjust their walking speed when approaching congestion. Users can explore several room configurations that allow one to study exit route choice and way finding, and identify bottlenecks in building design. This model relates to the genius of such models of which the social forces model developed and popularized by [91] is typical. Additionally the agents within the model can also be influenced by the presence of avatars –digital representations of actual people (*i.e.* users of the model)[5].

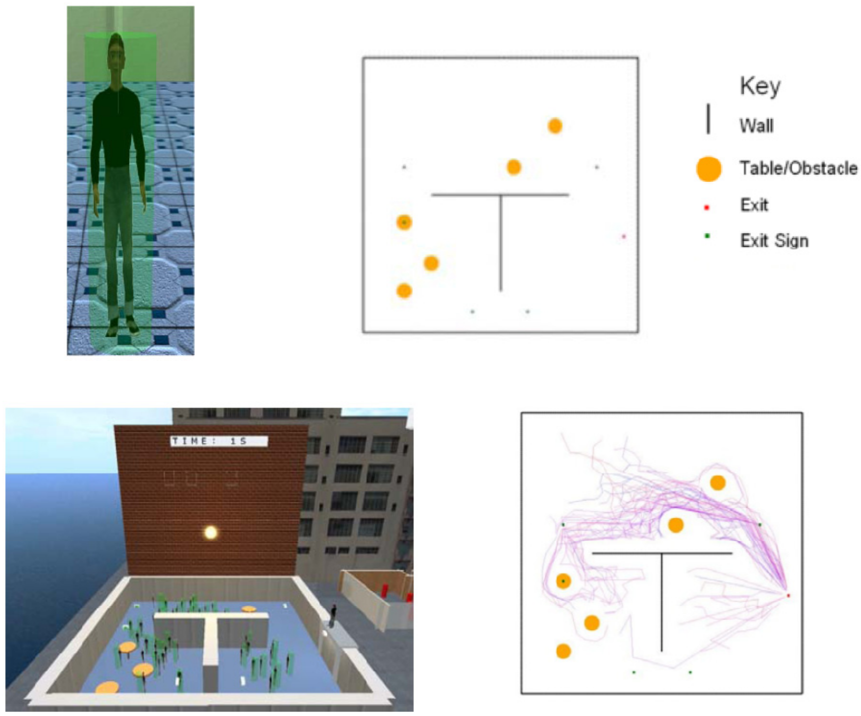


Figure 10: Pedestrian modeling within Second Life: (top left) a pedestrian agent within the model in green; (top right) room configuration; (bottom left) pedestrians and their environment; (bottom right) tracing the pedestrians routes to the exit (red dot).

DISCUSSION

In the past, the communication of models was mainly done through discussion of model results. However, increasing amounts of digital data and advances in GIS and CAD software enables us to not only create geographic explicit agent-based models, but also detailed 3D cityscapes in which to embed such model results. Advances in computer technology specifically the CPU and the GPU, and networked communications allow us to analyze and communicate such data and models to anyone who is connected to the internet. Nevertheless, combining ABM and 3D cityscapes is still much in its infancy and their combined potential is still unknown for scientific research. This chapter has attempted to explore this potential outlining some of our initial research, why it might be important and how 3D agent-based models can be created utilizing advances in computer technology. While there are many software environments that support the development of 3D agent-based models, many are commercial applications originating from computer gaming and the entertainment industry, and are to some extent black boxes. However, programming 3D agent-based models from scratch using the GPU is a non trivial task but perhaps in the future, toolkits might be developed to do this, just as Repast and NetLogo have developed functionality to deal with geospatial data.

One question this paper attempts to address is why do we need the 3D component in agent-based models? In Section 3 we discussed how people make cognitive maps of their environment. By relating models to actual places we would

[5] The website accompanying this work can be found at http://www.casa.ucl.ac.uk/abm/secondlife/.

argue that people can more easily relate to such models (*i.e.* gain a sense of location and place). This is perhaps one of the most important roles of 3D agent-based models. If the role of the model is to portray some complex behavior or problem to those that we seek to influence by relating it to actual places, people may more easily understand what is occurring within the model. We believe such an approach allows us to share modeling processes and its outcomes with various non-expert participants and potentially allow non-experts to participate in actual model construction in the case of virtual worlds. However, to truly understand the utility of 3D agent-based models over their 2D counterparts, we need to carry out user testing, which we see as a future avenue of research. We do not want to simply state that by moving agent-based models from the 2D to the 3D will further enhance their communication, usability and persuasive powers without just cause. The tools and techniques presented show the potential of virtual worlds, CAD packages and game engines to act as portals for allowing modelers, policy makers and citizens to communicate, share and visualize 3D spatial agent-based models which tentatively further our understanding of how these models work. By making these models available to whoever is connected to the internet allows them to go under greater scrutiny than was possible in the past, thus aiding the use of agent-based models as a tool for decision support.

However, a note of caution is also needed, that is 3D visualization of agent-based models does not replace the need for good models. Just as their 2D counterparts, 3D agent-based models that attempt to tackle the real-world problems need to be based on theory or insights gained from the phenomena under investigation. If this is not the case, 3D agent-based models are no better than 'eye-candy' and tell us little about the phenomena under investigation. We envisage 3D visualization as a tool for conveying the complexities of agent-based models to those we seek to influence. Both in relation to how people relate to space (as discussed in Section 3) but also how human spatial behavior within the built environment may be related to some simple physical properties of the urban environment [92]. Without incorporating the complexities of the third dimension into these models this may be missed. For example, the use of 3D models allows one to evaluate potential visual impacts of the existing and proposed urban form before urban design decisions are made. Furthermore, in the introduction we discussed how agent-based models exploring spatial patterns must look right. Combining models to 3D digital environments might therefore be of benefit here. However, maybe the biggest benefit of the development of 3D city models and game engines is the development of GPUs that allow us to simulate millions of agents as discussed in Section 4.

Looking towards the future, it is clear that cities are composed of many individuals and objects. Such objects interact with each other over varying scales both spatially and temporally, from the movement of pedestrians, to the hourly flows of traffic, to urban growth and change over months, to that of migration over years, to the rise and fall of civilizations over eons. What these processes all have in common is that they are composed of individual actors and to some extent, progress in exploring these using agent-based models is being made. The potential of combining these different processes within a single 3D modeling environment is highly appealing in the sense of giving a picture of city life. With the growth in computational power in the not so distant future it may be possible to use virtual worlds such as Second Life or OpenSim [93] to model whole cities, combining various types of models from iconic to symbolic in a single environment. Models exploring issues such as pedestrian movement, traffic, residential location, employment, gentrification could be merged in a single environment whose interactions feed back into each other and to the overall character of the city. A SimCity for real if you like, but where the focus is not just on end-user visualization but on understanding the behavior and interactions of all the agents and processes that underpin a city. However, to do this we need to improve our understanding of these complex processes.

REFERENCES

[1] I. Benenson, I. Omer, and E. Hatna, "Entity-based modelling of urban residential dynamics: the case of Yaffo, Tel Aviv", Environment and Planning B, vol. 29, pp. 491-512, 2002.

[2] P. M. Torrens, "Simulating sprawl", Annals of the Association of American Geographers, vol. 96, pp. 248-275, 2006.

[3] J. Jackson, B. Forest, and R. Sengupta, "Agent-based simulation of urban residential dynamics and land rent change in a gentrifying area of Boston", Transactions in GIS, vol. 12, pp. 475-491, 2008.

[4] C. L. Barrett, R. J. Beckman, K. P. Berkbigler, K. R. Bisset, B. W. Bush, K. Campbell, S. Eubank, K. M. Henson, J. M. Hurford, D. A. Kubicek, M. V. Marathe, P. R. Romero, J. P. Smith, L. L. Smith, P. E. Stretz, G. L. Thayer, E. Van Eeckhout, and M. D. Williams, "TRansportation ANalysis SIMulation System (TRANSIMS). Portland study reports 1, LA-UR-01-5711", Los Alamos National Laboratory, Los Alamos, NM 2002.

[5] M. Haklay, D. O'Sullivan, M. Thurstain-Goodwin, and T. Schelhorn, ""So go downtown": simulating pedestrian movement in town centres", Environment and Planning B, vol. 28, pp. 343-359, 2001.

[6] M. Batty, J. Desyllas, and E. Duxbury, "Safety in numbers? Modelling crowds and designing control for the Notting Hill carnival", Urban Studies, vol. 40, pp. 1573-1590, 2003.

[7] C. J. E. Castle and A. T. Crooks, "Principles and concepts of agent-based modelling for developing geospatial simulations", Centre for Advanced Spatial Analysis (University College London): Working Paper 110, London, UK, 2006.

[8] I. Benenson and P. M. Torrens, Geosimulation: Automata-based modelling of urban phenomena. London, UK: John Wiley & Sons, 2004.

[9] A. T. Crooks, C. J. E. Castle, and M. Batty, "Key challenges in agent-based modelling for geo-spatial simulation", Computers, Environment and Urban Systems, vol. 32, pp. 417-430, 2008.

[10] B. B. Mandelbrot, The Fractal Geometry of Nature. New York, NY: W.H. Freeman, 1983.

[11] D. Kornhauser, U. Wilensky, and D. Rand, "Design guidelines for agent based model visualization", Journal of Artificial Societies and Social Simulation, vol. 12, no. 2, 2009. [Online] Available: http://jasss.soc.surrey.ac.uk/12/2/1.html. [Accessed Feb. 8, 2010].

[12] M. J. North and C. M. Macal, Managing business complexity: discovering strategic solutions with agent-based modelling and simulation. New York, NY: Oxford University Press, 2007.

[13] W. Tobler, "A computer movie simulating urban growth in the Detroit region", Economic Geography, vol. 46, pp. 234-240, 1970.

[14] K. C. Clarke, N. Gazulis, C. K. Dietzel, and N. C. Goldstein, "A Decade of SLEUTHing: lessons learned from applications of a cellular automaton land use change model", in Classics from IJGIS: Twenty Years of the International Journal of Geographical Information Science and Systems, P. Fisher, Ed. Boca Raton, FL: Taylor & Francis, 2007, pp. 413-426.

[15] C. Dibble and P. G. Feldman, "The GeoGraph 3D computational laboratory: network and terrain landscapes for Repast", Journal of Artificial Societies and Social Simulation, vol. 7, no. 1, 2004. [Online] Available: http://jasss.soc.surrey.ac.uk/7/1/7.html. [Accessed Feb. 8, 2010].

[16] J. Thorp, S. Guerin, F. Wimberly, M. Rossbach, O. Densmore, M. Agar, and D. Roberts, "Santa Fe on fire: agent-based modelling of wildfire evacuation", in Proceedings of the Agent 2006 Conference on Social Agents: Results and Prospects, University of Chicago and Argonne National Laboratory, Chicago, IL, 2006.

[17] D. C. Parker, S. M. Manson, M. A. Janssen, M. J. Hoffmann, and P. Deadman, "Multi-agent systems for the simulation of land-use and land-cover change: a review", Annals of the Association of American Geographers, vol. 93, pp. 314-337, 2003.

[18] V. Grimm, "Visual debugging: a way of analyzing, understanding, and communicating bottom-up simulation models in ecology", Natural Resource Modeling, vol. 15, pp. 23-38, 2002.

[19] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. Heinz, G. Huse, A. Huth, J. Jepsen, C. Jorgensen, W. Mooij, B. Muller, G. Pe'er, C. Piou, S. Railsback, A. Robbins, M. Robbins, E. Rossmanith, N. Ruger, E. Strand, S. Souissi, R. Stillman, R. Vabo, U. Visser, and D. Deangelis, "A standard protocol for describing individual-based and agent-based models", Ecological Modelling, vol. 198, pp. 115–126, 2006.

[20] D. A. Smith and A. T. Crooks, "From buildings to cities: techniques for the multi-scale analysis of urban form and function", Centre for Advanced Spatial Analysis (University College London): Working Paper 155, London, UK, 2010.

[21] M. Batty, D. Chapman, S. Evans, M. Haklay, S. Kueppers, N. Shiode, A. Hudson-Smith, and P. M. Torrens, "Visualizing the city: communicating urban design to planners and decision-makers", in Planning Support Systems: Integrating Geographic Information Systems, Models and Visualisation Tools, R. K. Brail and R. E. Klosterman, Eds. Redlands, CA: ESRI Press, 2001, pp. 405-443.

[22] M. Batty and A. Hudson-Smith, "Virtuality and cities: definitions, geographies, designs", in Virtual Reality in Geography, P. Fisher and D. Unwin, Eds. London, UK: Taylor & Francis, 2002, pp. 270-291.

[23] S. Orford, "Towards a data-rich infrastructure for housing market research: deriving floor area estimates for individual properties from secondary data sources", Environment and Planning B, vol. 37, pp. 248-264, 2010.

[24] R. K. Brail and R. E. Klosterman (Eds), Planning support systems: integrating geographic information systems, models and visualisation tools. Redlands, CA: ESRI Press, 2001.

[25] A. M. MacEachren and M. Kraak, "Research challenges in geovisualization", Cartography and Geographic Information Science, vol. 28, pp. 3-12, 2001.

[26] I. S. Lowry, "A short course in model design", Journal of the American Institute of Planners, vol. 31, pp. 158-165, 1965.

[27] M. Batty, Urban modelling: algorithms, calibrations, predictions. Cambridge, UK: Cambridge University Press, 1976.

[28] M. Batty and A. Hudson-Smith, "Urban simulacra", Architectural Design, vol. 75, pp. 42-47, 2005.

[29] M. Batty, "Models in planning: technological imperatives and changing roles", International Journal of Applied Earth Observation and Geoinformation, vol. 3, pp. 252-266, 2001.

[30] G. Vizzari, G. Pizzi, and M. Federici, "Visualization of discrete crowd dynamics in a 3D environment", in 7th International Conference on Cellular Automata for Research and Industry, Perpignan, France, 2006, pp. 720-723.

[31] A. Willis, N. Gjersoe, C. Havard, J. Kerridge, and R. Kukla, "Human movement behaviour in urban spaces: implications for the design and modelling of effective pedestrian environments", Environment and Planning B, vol. 31, pp. 805-828, 2004.

[32] A. Hudson-Smith, "Digitally distributed urban environments: the prospects for online planning", London, England: PhD Thesis, University College London, 2003.

[33] K. Lynch, The image of the city. Cambridge, MA: MIT Press, 1960.

[34] L. N. Groat, Giving places meaning. London, UK: Academic Press, 1995.

[35] N. Pelechano, J. Allbeck, and N. Badler, "Controlling individual agents in high-density crowd simulation", in ACM SIGGRAPH / Eurographics Symposium on Computer Animation, San Diego, CA, 2007, pp. 99-108.

[36] R. Burkhard, S. Bischof, and A. Herzog, "The potential of crowd simulations for communication purposes in architecture", in 12th International Conference Information Visualisation, London, UK, 2008, pp. 403-408.

[37] H. Nakanishi, T. Ishida, and S. Koizumi, "Virtual cities for simulating smart urban public spaces", in Handbook of research on urban informatics: the practice and promise of the real-time city, M. Foth, Ed. Hershey, PA: IGI Global, 2009, pp. 257-269.

[38] M. Volk, J. Hirschfeld, A. Dehnhardt, G. Schmidt, C. Bohn, S. Liersch, and P. W. Gassman, "Integrated ecological-economic modeling of water pollution abatement management options in the Upper Ems River Basin", Ecological Economics, vol. 66, pp. 66-76, 2008.

[39] SimCity, "Simcity Societies." 2009. [Online] Available: http://simcitysocieties.ea.com/. [Accessed June. 6, 2009].

[40] J. Krikke, "Axonometry: A matter of perspective", IEEE Computer Graphics and Applications, vol. 20, pp. 7-11, 2000.

[41] P. C. Adams, "Teaching and learning with SimCity 2000", Journal of Geography, vol. 97, pp. 45-55, 1998.

[42] Micropolis, "Micropolis: binary and source code." 2009. [Online] Available: http://www.donhopkins.com/home/micropolis/. [Accessed June. 6, 2009].

[43] A. Cockburn and B. McKenzie, "Evaluating the Effectiveness of Spatial Memory in 2D and 3D physical and virtual environments", in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Changing Our World, Changing Ourselves, Minneapolis, MN, 2002, pp. 203-210.

[44] StarLogo TNG, "StarLogo TNG homepage." 2008. [Online] Available: http://education.mit.edu/drupal/starlogo-tng/. [Accessed Jan. 21, 2009].

[45] NetLogo, "NetLogo: multi-agent programmable modeling Environment." 2009. [Online] Available: http://ccl.northwestern.edu/netlogo/. [Accessed June. 6, 2009].

[46] Repast Simphony, "Recursive porous agent simulation toolkit: simphony." 2008. [Online] Available: http://ccl.northwestern.edu/netlogo/. [Accessed June. 6, 2009].

[47] NASA, "World Wind." 2009. [Online] Available: http://worldwind.arc.nasa.gov/java/. [Accessed Sept. 18, 2009].

[48] G. E. Moore, "Cramming more components onto integrated circuits", Electronics, vol. 38, pp. 114-117, 1965.

[49] Wikipedia, "Central processing unit." 2009. [Online] Available: http://en.wikipedia.org/wiki/CPU. [Accessed June. 12, 2009].

[50] J. D. Owens, D. Luebke, N. Govindaraju, M. Harris, J. Krüger, A. E. Lefohn, and T. Purcell, "A survey of general-purpose computation on graphics hardware", in Eurographics, State of the Art Reports, Dublin, Ireland, 2005, pp. 21-51.

[51] M. Lysenko and R. M. D'Souza, "A framework for megascale agent based model simulations on graphics processing units", Journal of Artificial Societies and Social Simulation, vol. 11, no. 4, 2008. [Online] Available: http://jasss.soc.surrey.ac.uk/11/4/10.html. [Accessed Feb. 8, 2010].

[52] DirectX, "DirectX: application programming interfaces." 2009 [Online] Available: http://www.microsoft.com/games/en-US/aboutgfw/Pages/directx10.aspx. [Accessed June. 19, 2009].

[53] OpenGL, "OpenGL: 2D and 3D graphics application programming interface." 2009. [Online] Available: http://www.opengl.org/. [Accessed June. 19, 2009].

[54] 3ds Max, "3ds Max: 3D modeling, animation, and rendering software." 2009 [Online] Available: http://usa.autodesk.com/adsk/servlet/index?siteID=123112&id=5659302. [Accessed June. 22, 2009].

[55] C. J. E. Castle, "Agent-based modelling of pedestrian evacuation: a study of London's King's Cross underground station", London, UK: PhD Thesis, University College London, 2007.

[56] M. Batty, R. Conroy, B. Hillier, B. Jiang, J. Desyllas, C. Mottram, A. Penn, A. Smith, and A. Turner, "The virtual Tate", Centre for Advanced Spatial Analysis (University College London): Working Paper 5, London, UK, 1998.

[57] Crysis, "Crysis: EA Games official crysis website." 2009. [Online] Available: http://games.ea.com/crysis/. [Accessed June. 4, 2009].
[58] Massive, "Simulation and visualization system for generating and visualizing realistic crowd behaviors." 2009. [Online] Available: http://games.ea.com/crysis/. [Accessed June. 4, 2009].
[59] Maya, "Maya: Integrated 3D Modeling, Animation and Rendering Software." 2009 [Online] Available: http://usa.autodesk.com/adsk/servlet/index?id=7635018&siteID=123112. [Accessed June. 22, 2009].
[60] Legion, "Legion: pedestrian simulation software." 2009 [Online] Available: http://www.legion.com/. [Accessed June. 22, 2009].
[61] D. Thalmann and S. R. Musse, Crowd simulation. London, UK: Springer-Verlag, 2007.
[62] Designhive, "Oxford Circus redesign." Nov 2008. [Online] Available: http://www.designhive.co.uk/Press_and_PR/News_London_Oxford_Circus. [Accessed June. 12, 2009].
[63] Y. Akiyama and R. Shibaski, "Development of detailed spatio-temporal urban data through the integration of digital maps and yellow page data and feasibility study as complementary data for existing statistical information", in 11th International Conference on Computers in Urban Planning and Urban Management Hong Kong, PRC, 2009.
[64] S. Gwynne, E. R. Galea, P. J. Lawrence, and L. Filippidis, "Modelling occupant interaction with fire conditions using the buildingEXODUS evacuation model", Fire Safety Journal, vol. 36, pp. 327-357, 2001.
[65] STEPS, "Pedestrian movement software." 2009. [Online] Available: http://www.mottmac.com/skillsandservices/software/stepssoftware/. [Accessed June. 23, 2009].
[66] Unreal, "Unreal developer network." 2009. [Online] Available: http://udn.epicgames.com/Main/WebHome.html. [Accessed June. 4, 2009].
[67] OGRE, "Ogre 3D: open source game engine." 2009. [Online] Available: http://www.ogre3d.org/. [Accessed June. 4, 2009].
[68] Panda3D, "Panda3D game engine." 2009. [Online] Available: http://www.panda3d.org/. [Accessed June. 4, 2009].
[69] K. Merrick and M. Maher, "Motivated reinforcement learning for adaptive characters in open-ended simulation games", in Proceedings of the International Conference on Advances in Computer Entertainment Technology, Salzburg, Austria, 2007, pp. 127-134.
[70] K. Stephens, B. Pham, and A. Wardhani, "Modelling fish behaviour", in Proceedings of the 1st International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia, Melbourne, Australia, 2003, pp. 71-78.
[71] Havok, "Event-driven character behaviors system." 2009. [Online] Available: http://www.havok.com/. [Accessed June. 26, 2009].
[72] K. Nagel and M. Schreckenberg, "A cellular automaton model for freeway traffic", Journal de Physique 1, vol. 2, pp. 2221-2229, 1992.
[73] U. Wilensky, "NetLogo traffic grid model." 2003. [Online] Available: http://ccl.northwestern.edu/netlogo/models/TrafficGrid. [Accessed June. 12, 2009].
[74] A. Patel, "Crowd simulation, May, 2009. [Online] Available: http://crowdsimulation.blogspot.com/ [Accessed June, 4, 2009].
[75] T. Narahara, "Enactment software: spatial designs using agent-based models", in Proceedings of the Agent 2007 on Complex Interaction and Social Emergence, Chicago, IL, 2007, pp. 465-474.
[76] MATSim, "Large-scale agent-based transport simulations toolbox." 2009. [Online] Available: http://www.matsim.org/. [Accessed June. 23, 2009].
[77] D. E. Baird and M. Fisher, "Neomillennial user experience design strategies: utilizing social networking media to support "always on" learning styles", Journal of Educational Technology Systems, vol. 34, pp. 5-32 2005.
[78] R. J. Ingram and S. D. Benford, "The application of legibility techniques to enhance information visualisations", The Computer Journal, vol. 39, pp. 819-836, 1996.
[79] Unity, "Multiplatform game development tool." 2009. [Online]Available: http://unity3d.com/unity/. [Accessed June. 8, 2009].
[80] SketchUp, "Create, modify and share 3D models." 2009. [Online] Available: http://sketchup.google.com/. [Accessed June. 29, 2009].
[81] Science of Collaboratories, "The science of collaboratories." 2009. [Online] Available: http://www.scienceofcollaboratories.org/. [Accessed Oct. 1, 2009].
[82] W. S. Bainbridge, "The scientific research potential of virtual worlds", Science, vol. 317, pp. 472-476, 2007.
[83] H. Berger, M. Dittenbach, D. Merkl, A. Bogdanovych, S. Simoff, and C. Sierra, "Opening New Dimensions for E-Tourism", Virtual Reality, vol. 11, pp. 75-87, 2007.

[84] A. Bogdanovych, M. Esteva, S. Simoff, C. Sierra, and H. Berger, "A methodology for 3D electronic institutions", in Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems, Honolulu, HI, 2007, pp. 358-360.

[85] E. Dieterle and J. Clarke, "Multi-user virtual environments for teaching and learning", in Encyclopaedia of Multimedia Technology and Networking, 2nd Edition ed, M. Pagani, Ed. Hershey, PA: Idea Group, Inc, 2009, pp. 1033-1041.

[86] A. Hudson-Smith and A. T. Crooks, "The renaissance of geographic information: neogeography, gaming and Second Life", Centre for Advanced Spatial Analysis (University College London): Working Paper 142, London, UK, 2008.

[87] M. Rymaszewski, W. J. Au, M. Wallace, C. Winters, C. Ondrejka, and B. Batstone-Cunningham, Second Life: the official guide. Hoboken, NJ: John Wiley & Sons, Inc, 2007.

[88] T. C. Schelling, "Dynamic models of segregation", Journal of Mathematical Sociology, vol. 1, pp. 143-186, 1971.

[89] A. T. Crooks, A. Hudson-Smith, and J. Dearden, "Agent Street: an environment for exploring agent-based models in Second Life", Journal of Artificial Societies and Social Simulation vol. 12, no. 4, 2009. [Online] Available: http://jasss.soc.surrey.ac.uk/12/4/10.html. [Accessed Feb. 8, 2010].

[90] C. J. E. Castle, "Guidelines for assessing pedestrian evacuation software applications", Centre for Advanced Spatial Analysis (University College London): Working Paper 115, London, UK, 2007.

[91] D. Helbing and P. Molnár, "Social force model for pedestrian dynamics", Physical Review E, vol. 51, pp. 4282-4286, 1995.

[92] P. P. Yang, S. Y. Putra, and W. Li, "Viewsphere: a GIS-Based 3D visibility analysis for urban design evaluation", Environment and Planning B, vol. 34, pp. 971-992, 2007.

[93] OpenSim, "OpenSimulator: a 3D application server." 2009. [Online] Available: http://opensimulator.org/. [Accessed Dec. 23, 2009].

CHAPTER 5

Semantically-Enhanced Virtual Geographic Environments for Multi-Agent Geo-Simulation

Mehdi Mekni*, Bernard Moulin[1] and Sébastien Paris[2]

[1]*Cognitive Informatics Laboratory (CIL), Laval University, Quebec and* [2]*Graphics Vision Visualization (GV2), Trinity College, Dublin, Ireland*

Abstract: Multi-Agent Geo-Simulation (MAGS) is a modeling and simulation paradigm which aims to study various phenomena in a variety of domains involving a large number of heterogeneous actors (implemented as software agents) evolving in, and interacting with, a Virtual representation of the Geographic Environment (VGE). A critical step towards the development of advanced MAGSs is the creation of semantically-enhanced and geo-metrically accurate virtual geographic environments called Informed VGE (IVGE). In this chapter we propose a novel approach to automatically build an accurate IVGE using an exact decomposition of realistic spatial data provided by Geographic Information System (GIS). The IVGE model relies on a hierarchical topologic graph structure built using geometric, topologic, and semantic abstraction processes and enhanced by spatial seman-tic information represented using Conceptual Graphs (CGs). We demonstrate how we take advantage of the IVGE description in order to provide an accurate 2D/3D visualization tool of the space partitioning as well as to support situated reasoning such as path planning with respect to both agents and environments' characteristics.

INTRODUCTION

During the last decade, Multi-Agent Geo-Simulation (MAGS) [4] has attracted a growing interest from researchers and practitioners to simulate various phenomena in a variety of domains including traffic simulation, crowd simulation, urban dynamics, and changes of land use and cover, to name a few. Such approaches are used to study various phenomena (*i.e.* car traffic, crowd behaviors, *etc.*) involving a large number of simulated actors (implemented as software agents) evolving in, and interacting with, an explicit description of the geographic environment called *Virtual Geographic Environment* (VGE). A critical step towards the development of MAGS is the creation of a VGE, using appropriate representations of the geographic space and of the objects contained in it (also called "situated objects"), in order to efficiently support the agents' situated reasoning. Since a geographic environment may be complex and large scale, the creation of a VGE is difficult and needs large quantities of geometrical data originating from the environment characteristics (terrain elevation, location of objects and agents, *etc.*) as well as semantic information that qualifies space (building, road, park, *etc.*). In order to yield realistic MAGSs, a VGE must precisely represent the geometrical information which corresponds to geographic features. It must also integrate several semantic notions about various geographic features. To this end, we propose to enrich the VGE data structure with semantic information that is associated with the geographic features. A number of challenges arise when creating such a semantically-enriched and geometrically-accurate representation of a VGE, among which we mention: 1) to automatically create an accurate geometric representation of a 3D VGE; 2) to automatically integrate in the geometric representation several types of semantic information; 3) to make use of this representation in "situated reasoning" algorithms (such as path finding) which are required for MAGS.

In this chapter we present a novel approach that addresses these challenges toward the creation of such a semantically-enriched and geometrically-accurate VGE, which we call an *Informed VGE* (IVGE). Fig. **1** presents an overview of the proposed approach which aims at producing an exact representation of the geographic environment based on realistic data provided by *Geographic Information Systems* (GIS) and which uses the *Constrained Delaunay Triangulation* (CDT) technique for an accurate spatial decomposition. This representation is organized as a topological graph enhanced with data integrating both quantitative (like the geometry) and qualitative information (like the types of areas such as roads and buildings). The qualitative information is organized as expressions

***Address correspondence to Mehdi Mekni:** Institut national de la recherche scientifique (INRS), Centre Eau-Terre-Environnement, 490, rue de la Couronne, Québec (Québec), G1K 9A9, Canada; Tel: 418-262-7601; E-mail: mmekni@gmail.com

Danielle J. Marceau and Itzhak Benenson (Eds)

composed of concepts and relations using Conceptual Graphs (CG) [36] which enable us to create a semantic description of the virtual geographic environment. In addition, the topological graph is semantically abstracted in order to provide different views of the same IVGE to specific agents' behavioral activities (navigable areas for pedestrians, rolling areas for vehicles, *etc.*). Finally, the IVGE description can be visualized for spatial analysis purposes. Its topological graph structure can either be saved for future use in a binary proper format or exported to a standard GIS vector file.

The rest of this chapter starts with a review of works on GIS and agent-based modeling and an overview of GIS data formats (Section 2) and space decomposition techniques (Section 3). In Section 4, we present our approach to automatically create an accurate Informed VGE (IVGE) from GIS data. Section 5 outlines two ways to enhance the IVGE description: 1) using a geometric abstraction process in order to qualify the terrain's elevation with elevation semantics; and 2) using a topological abstraction that reduces the size of the topological graph and enables building a hierarchical topological graph. Section 6 points out the way we represent, propagate, and exploit the semantic information using Conceptual Graphs (GCs) in order to overlay hierarchical topological graphs with semantic abstraction. Section 7 presents how we leverage the hierarchical graph structure of the IVGE model in order to support situated reasoning algorithms such as hierarchical path planning. Section 8 highlights some results obtained by applying our approach to an urban environment in order to address the issue of path planning with respect to both agents' and environments' characteristics. Finally, Section 9 concludes this chapter and outlines some future works.

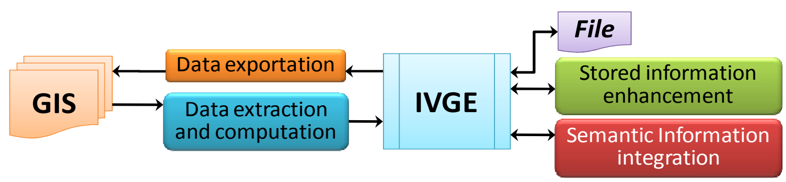


Figure 1: Global architecture for IVGE management: in blue, GIS data extraction and associated processes; in orange, IVGE data exportation to GIS format; in green, optional stages for IVGE information enhancement; in red, the semantic information representation using the CG formalism.

GIS AND AGENT-BASED MODELING

Several approaches have been proposed to create virtual scenes (virtual cities and urban contexts [10, 12, 13, 26, 40], virtual train stations [33, 35], virtual shopping malls [1]) in the field of behavioral and human animations. However, as the focus of these approaches is computer animation and virtual reality, the virtual environment usually plays the role of a simple background scene in which agents mainly deal with its geometric characteristics. Moreover, data used to build such virtual scenes are mostly provided by Computer Aided Design and Graphics (CADG) systems. In contrast to GIS data, CADG data are used to create precision drawings or technical illustrations for animation and visualization purposes, without taking into account geo-referenced spatial data. Moreover, interactions between agents and the environment are usually simple, permitting agents to only perceive and avoid obstacles in 2D or 3D virtual scenes [10]. This is due to the fact that the description of virtual environments is often limited to the geometric level, while it should also contain topological and semantic information for other types of applications.

In [30] Najlis and North discuss the interest of integrating GIS and agent-based modeling systems ([3, 6, 34, 41]; to name a few). Examples of recent applications include pedestrian dynamics, urban growth models and land use models. For agent-based modelers, this integration provides the ability to manipulate agents that are related to actual geographic locations. For GIS users, it provides the ability to model the emergence of phenomena through individual interactions of features over time and space [30].

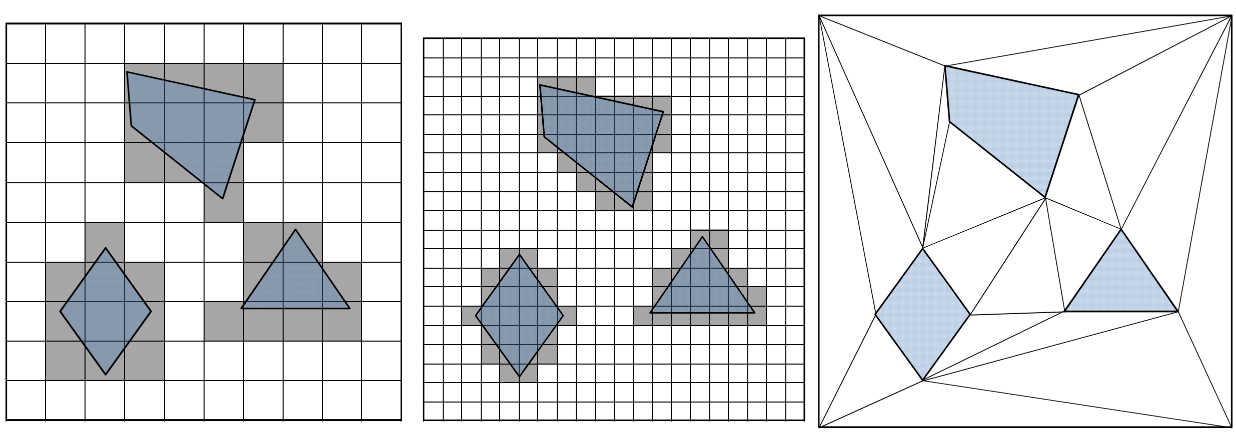
Several levels of integration of agent-based models and GIS are possible. Models may run on grid-based representation of geographic environments, implementing simple spatial functionalities within the agent model. Models may read spatial data from a GIS and usually can write output into a format readable by a GIS. When required at run time, integrated models can dynamically implement both agent-based models and GIS functionalities. Three approaches to such models are identified in [18]: (1) models that use separate GIS and agent-based programs/libraries and communicate via files written to disk; (2) models that use separate programs but communicate via a shared database or virtual memory, (3) and stand-alone models that implement GIS functionalities within the agent-based model. In the latter category, agents evolve and interact with an explicit representation of the geographic environment built using GIS data. The semantically-enhanced virtual geographic environment model that we propose in this chapter falls in this category.

GIS DATA AND SPATIAL DECOMPOSITION

GIS data are usually available in either raster or vector formats [32]. The raster format subdivides the space into a grid of regular cells associated with space related attributes. The cell size used for a raster layer affects the accuracy of the spatial decomposition. The cell size should be based on the original scale of the map because a too large cell size causes a loss of spatial information. Alternatively, using a cell size that is too small requires a lot of storage space and takes longer to process without adding additional precision to the spatial description. In contrast, the vector format exactly describes geographic information without constraining geometric shapes, and generally associates one qualitative data with each shape.

Such data, either raster or vector, is usually exploited by a VGE in two ways [15]: the approximate geometric subdivision and the exact geometric subdivision methods. The approximate geometric subdivision method is the direct mapping of the raster format [27], but it can also be applied to the vector format (Fig. **2a**). This discrete representation produces a grid-based data structure which can be used to merge multiple semantic data [42], the locations where to store these data being organized using the grid cells. The main drawback of the grid method is related to a loss of location accuracy [2], making it difficult to accurately position any information which is not aligned with the subdivision. Another drawback arises when trying to precisely represent large environments using a grid: the number of cells tends to increase dramatically, which makes the environment exploitation very costly. The grid method is easy to compute, to update and to maintain. It is suitable for the representation of continuous spaces. The grid-based method is mainly used for animation purposes [39] and large crowd simulations [29] because of the fast data access it provides.

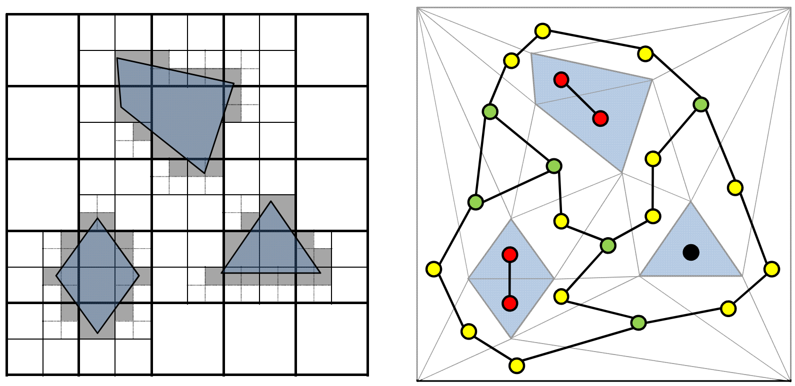
The second method, called exact geometric subdivision, consists of subdividing the environment in convex cells using the vector format as an input. The convex cells can be generated by several algorithms, among which the most popular is the Constrained Delaunay Triangulation (CDT) [20]. The CDT produces triangles while keeping the original geometry segments which are named constraints (Fig. **2b**). The first advantage of the exact subdivision method is to preserve the geometry of the input data, allowing to accurately manipulate and visualize the environment at different scales. Another advantage of this approach is that the number of produced cells only depends on the complexity of the input shapes, but not on the environment's size and scale as it is the case with the grid method. The main drawback of this approach is the difficulty to merge multiple semantic data for overlapping shapes. Moreover, this method is generally used to represent planar environments because the CDT can only handle 2D geometries. This method tends to be used for applications in which motion accuracy is essential, as for example in crowd microscopic simulations[23].



2-a-1 **2-a-2** **2-b**

Figure 2: Two common cell decomposition techniques; (a) approximate decomposition by grids considering two resolutions; (b) exact decomposition using CDT. White boxes are free, grey are obstacles.

Both VGE representations, approximate and exact, can be enhanced by an abstraction process [33]. The first goal of an abstraction is to improve the performance of the algorithms based on the environment description, such as path planning, by reducing the number of elements used to describe the environment. The usual abstraction model for grids is mainly geometric (Fig. **3a**). For example, the *quadtree* structure groups four cells of the same kind to create a higher level cell [35]. When considering the exact decomposition, an abstraction is usually based on topological properties rather than on purely geometric ones. Indeed, the exact cell subdivision generates connected triangles which can be manipulated as the nodes of a topological graph. This graph can then be abstracted by grouping the nodes, producing a new graph with a smaller number of nodes [33]. For example, Fig. **3b**) shows an abstraction which is only based on the nodes' number of connections c: isolated (c = 0), dead-end (c = 1), corridor (c = 2), and crossroad (c = 3). A topological graph can be used for spatial reasoning, like path planning, using traversal algorithms. These algorithms benefit from the abstraction by traversing first the more abstracted graph, and then by refining the computation in the sub-graphs until reaching the graph of the spatial subdivision. However, such algorithms raise a new need for an abstracted graph which is less addressed in the literature: it must contain the minimal information necessary to make decisions. For example, if the width of a path is relevant to a path planning algorithm, this information must be accessible in all the abstracted graphs; if not, the evaluation would be greatly distorted comparatively to a graph which is not abstracted.



(a) **(b)**

Figure 3: Abstraction examples for two kinds of environment descriptions; (a) Grid abstraction using a three levels *quatree*. (b) Left: topological graph representation of a CDT. Top right: arbi-trary abstraction example of this graph.

GENERATION OF INFORMED VIRTUAL GEOGRAPHIC ENVIRONMENTS

We propose an automated approach to compute the IVGE data directly from vector GIS data. This approach is based on four stages which are detailed in this section (Fig. **4**): input data selection, spatial decomposition, maps unification, and informed graph generation.

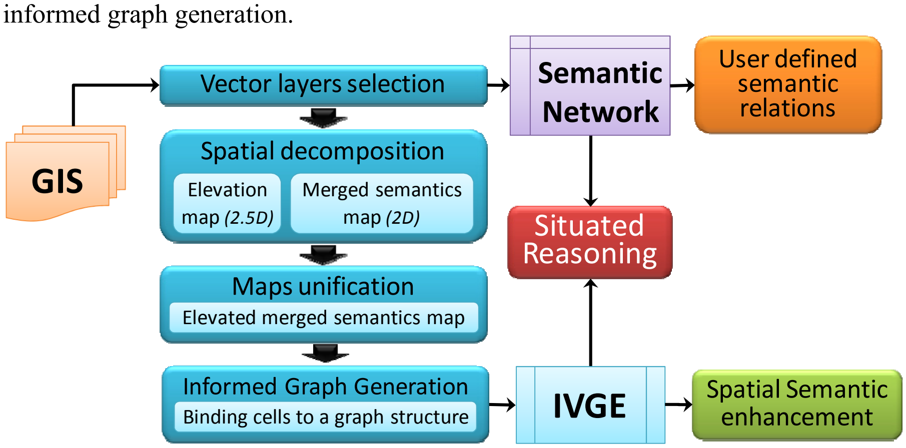


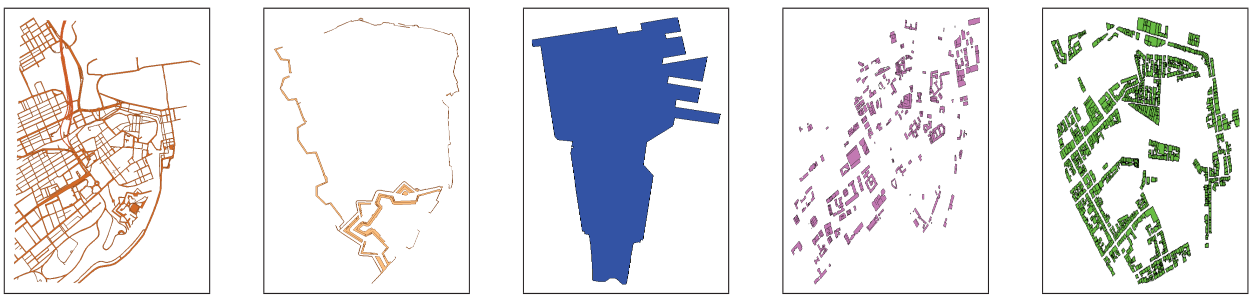
Figure 4: Left: the four stages to obtain an IVGE from GIS data; Right: the enhancements of the IVGE's description using semantic information.

Input Data Selection

The first step of our approach is the only one requiring human intervention. It consists of selecting the different vector data sets which are used to build the IVGE. The only restriction concerning these vector data layers is that they must belong to maps which respect the same scale. Vector data of different scales are not supported in our approach. Such data must be converted into a common scale before being used as input to the IVGE generation process.

The input data can be organized into two categories. First, elevation layers contain geographical marks indicating absolute terrain elevations. Since we consider 2.5D IVGE, a given coordinate cannot have two different elevations, making it impossible to represent tunnels for example. In addition, our model is able to automatically manage and merge multiple elevation layers.

Second, semantic layers are used to qualify various types of data in space. Each layer indicates the physical or virtual limits of a given set of features with identical semantics in the geographic environment, such as roads and buildings. The limits can overlap between two layers, our model being able to merge the information. Fig. **5** presents various semantic layers characterizing the city of Quebec in Canada.



(a) (b) (c) (d) (e)

Figure 5: Various semantic layers related to Quebec City in Canada: (a) road network; (b) old city wall; (c) marina; (d) governmental buildings; (e) houses

Spatial Decompostion

The second step consists of obtaining an exact spatial decomposition of the input data in cells. This process is entirely automatic, using Delaunay's triangulation. It can be divided into two parts in relation to the previous phase.

First, an elevation map is computed, corresponding to the triangulation of the elevation layers. All the elevation points of the layers are injected in a 2D triangulation, the elevation being considered as an additional data layer. This process produces an environment subdivision composed of connected triangles (Fig. **6a**). Such a subdivision provides information about coplanar areas: the elevation of any point inside a triangle can be deduced using the elevation of the three original points.

Second, a merged semantics map is computed, corresponding to a constrained triangulation of the semantic layers. Indeed, each segment of a semantic layer is injected as a constraint which keeps track of the original semantic data, using an additional datum. The obtained map is then a constrained triangulation merging all input semantics (Fig. **6b**): each constraint represents as many semantics as the number of input layers containing it.

Maps Unifiication

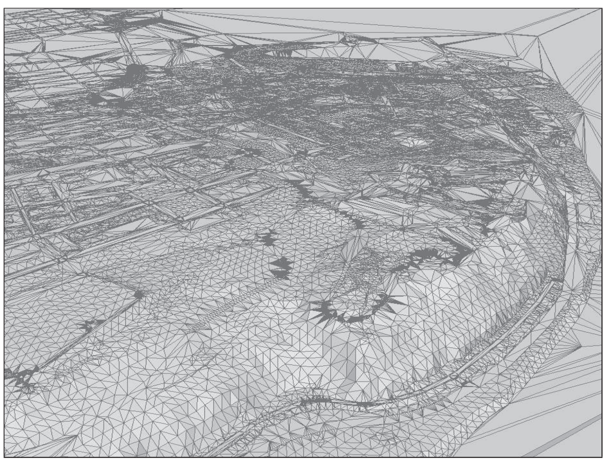
The third step consists of the unification of the two maps previously obtained. This phase can be depicted as the mapping of the 2D merged semantic map (Fig. **6b**) on the 2.5D elevation map (Fig. **6a**) in order to obtain the final 2.5D elevated merged semantics map (Fig. **6c**).

First, a preprocessing is carried out on the merged semantics map in order to preserve the elevation precision inside the unified map. Indeed, all the points of the elevation map are injected in the merged semantics triangulation, creating new triangles. This first process can be dropped if the elevation precision is not important.

Then, a second process elevates the merged semantics map. The elevation of each merged semantics point P is computed by retrieving the corresponding triangle T inside the elevation map, *i.e.* the triangle whose 2D projection contains the coordinates of P. Once T is obtained, the elevation is simply computed by projecting P on the plane defined by T using the Z axis. If P is outside the convex hull of the elevation map, then no triangle can be found and the elevation cannot be directly deduced. In this case, we use the average height of the points of the convex hull which are visible from P.

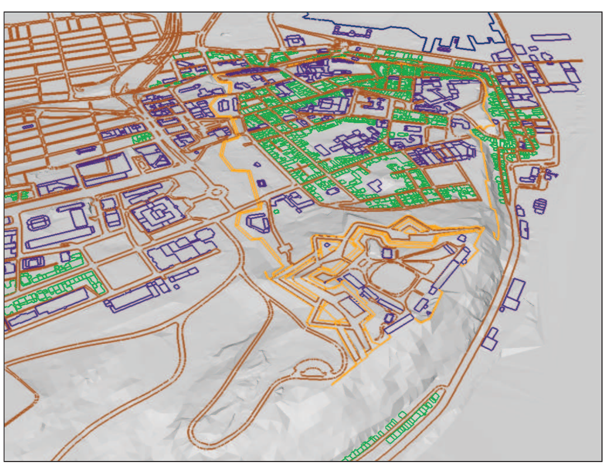
Informed Graph Generation

The obtained unified map now contains all the semantic information of the input layers, along with the elevation information. This map can be used as an Informed Graph (IG), where each node corresponds to the map's triangles, and each arc to the adjacency relations between these triangles. Nodes in the informed graph encompass not only geometric data related to the convex cells produced by the spatial decomposition but also semantic information related to the input GIS data. Then, common graph algorithms can be applied to this topological graph, and especially graph traversal ones. One of these algorithms allows to retrieve the node, and so the triangle, corresponding to given 2D coordinates. Once this node is obtained, it is possible to extract the data corresponding to the position, such as the elevation using the 2.5D triangle, and the semantics information.

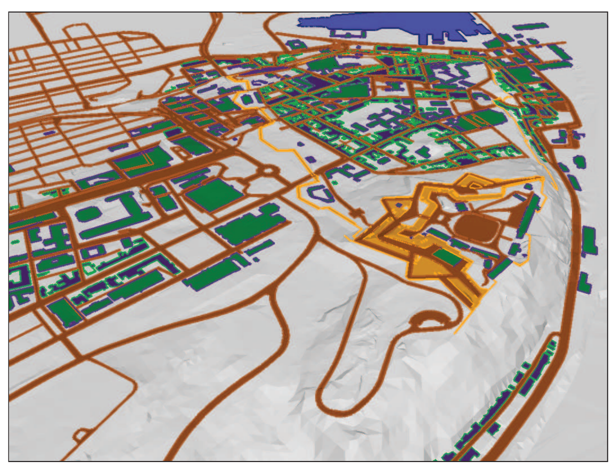


(a) Triangulated elevation map (2.5*D*).

(b) Merged semantics map (2*D*).



(c) Borders semantics.



(d) Borders and cells semantics.

Figure 6: The two processed maps (a, b), the unified map (c). The semantic colors are the same as in figure 5. In addition, the grey lines represent unconstrained triangulation segments. (d) Illustrates the unified maps after semantics propagation.

To conclude, we presented in this section a four-stage automated approach for the generation of accurate informed VGEs. As we will see in the following sections, the informed graph still needs enhancements particularly regarding its size when dealing with large scale and complex geographic environments and regarding the qualification of the terrain's elevation.

ENHANCEMENTS OF THE IVGE

The IVGE description resulting from the exact spatial subdivision and semantics merging provides an accurate informed virtual geographic environment. However, this description needs to be efficiently exploited in MAGS. The first enhancement is related to the qualification of terrain's elevation (Section 5.1). The second enhancement aims at optimizing the size of the informed graph structure using a topological abstraction process (Section 5.2).

Geomeric Abstraction

The spatial decomposition subdivides the environment into convex cells. Such cells encapsulate various quantitative geometric data which are suitable for accurate computations. Since geographic environments are seldom flat, it is important to consider the terrain's elevation. Elevation data are stored in a quantitative way which may suit to exact calculations, but spatial reasoning often needs to manipulate qualitative information. Indeed, when considering a slope, it is obviously simpler and faster to qualify it using an attribute such as *light* and *steep* rather than using numerical values. However, when dealing with large-scale steep geographic environments, handling the terrain's elevation, including its light variations, may be a complex task. To this end, we propose an abstraction process that uses geometric data to extract the average terrain's elevation information from spatial areas. The objectives of this *geometric abstraction* are twofold. First, it aims to reduce the amount of data used to describe the terrain's elevation in the IVGE. Second, it enhances the environment description by integrating qualitative information characterizing the terrain's elevations.

As presented in Section 4, the geographic environment is subdivided in cells of different slopes and sizes. The geometric abstraction process gathers cells in groups according to a geometric criterion: we chose the coplanarity of adjacent cells in order to obtain uniform elevation areas. The algorithm takes advantage of the graph structure obtained using the IVGE extraction process (see sub-section 4.4). A cell corresponds to a node in the informed graph. A node represents a triangle generated by the CDT spatial decomposition technique. A cell is characterized by its boundaries, its neighboring cells, its surface as well as its normal vector which is a vector perpendicular to its plan. A group is a container of adjacent cells. The grouping strategy of this algorithm is based on a coplanarity criterion which is assessed by computing the difference between the normal vectors of two neighboring cells or groups of cells. Since a group is basically composed of adjacent cells, it is obvious to characterize a group by its boundaries, its neighboring groups, its surface, as well as its normal vector. However, the normal vector of a group must rely on an interpretation of the normal vectors of its composing cells. In order to compute the normal vector of a group, we adopt the area-weight normal vector [7] which takes into account the unit normal vectors of its composing cells as well as their respective surfaces. Let S_c denote the surface of a cell c and N_c be its unit normal vector, the normalized area-weight normal vector N_G of a group G is computed as follows:

$$\underset{N_G}{\longrightarrow} = \sum_{c\in G}\left(S_c.\underset{N_c}{\longrightarrow}\right)\Big/\sum_{c\in G} S_c$$

The geometric abstraction algorithm uses two input parameters: 1) a set of starting cells which act as access points to the graph structure, and 2) a gradient parameter which corresponds to the maximal allowed dif-ference between cells' inclinations. Indeed, two adjacent cells are considered coplanar and hence grouped, when the angle between their normal vectors is inferior or equal to the gradient. The recursive geometric abstraction algorithm is composed of five steps:

- **Step 1:** For each cell c of the *starting cells*, create a new group G and proceed to *step 2*.
- **Step 2:** For each neighboring group or cell *n* of *G*, depending if the neighbor has already been processed, proceed to *step 3*, else proceed to *step 1*.
- **Step 3:** if angle ($\underset{N_g}{\longrightarrow}\underset{N_n}{\longrightarrow}$) ≤ gradient then proceed to *step 4*. Otherwise proceed to *step 5*.
- **Step 4:** Merge n in group G, then evaluate $\underset{N_g}{\longrightarrow}$ using equation (1). Proceed to *step 2* again for G.
- **Step 5:** If n is an unprocessed cell, create a new group G with n and proceed to *step 2*.

The terrain's elevation which characterizes each group is still a quantitative data described using area-weighted normal vectors. Such quantitative data are too precise to be used by qualitative spatial reasoning systems. Hence, a qualification process would greatly simplify spatial reasoning mechanisms. The geometric abstraction is used to enhance the IVGE by qualifying the terrain's elevation with semantic information and integrating it in the description of the geographic environment.

Qualification of Terrain's Elevation

The geometric abstraction algorithm computes quantitative geometric data describing the terrain's inclination. Such data is stored as numerical values which allow to accurately characterizing terrains' elevations. However, handling and exploiting quantitative data is a complex task since the number of values may be too large and consequently difficult to transcribe and analyze. Therefore, we propose to interpret the quantitative data of terrain's inclination by qualifying areas' elevations. Semantic labels, which are called semantics elevation, are associated with quantitative intervals of values that represent the terrain's elevation. In order to obtain elevation semantic we propose a process in two steps that take advantage of the geometric abstraction: 1) discretise the angle a between the weighted normal vector Ng of a group g and the horizontal plane. 2) assign to each discrete value a semantic information which quali.es it. The discretisation process can be done in two ways: a customised and automated approach.

The customized approach qualifies the terrain's elevation and requires the user to provide a complete specification of the discretisation. Indeed, a list of angle intervals, as well as their associated semantic attributes must be specified. The algorithm iterates over the groups obtained by the geometric abstraction process. For each group G, it retrieves the terrain inclination value I. Then, this process checks the list of angle intervals' bounds and determines the one to which one belongs the inclination value *I*. Finally, the customized discretisation extracts the semantic elevation from the selected elevation interval and assigns it to the group

G. For example, let us consider the following inclination interval and the associated semantic elevations: {([10°, 20°], light slope),([20°, 25°], steep slope)}. Such a customized specification associates the semantic elevation "*light slope*" with inclination values included in the interval [10°, 20°] and the semantic elevation "*steep slope*" with inclination values included in the interval [20°, 25°].

The automated approach only relies on a list of semantic elevations representing the elevation qualifications. Let N be the number of elements of this list, and T be the total number of groups obtained by the geometric abstraction algorithm. First, the automated discretisation orders groups, using their terrain inclination. Then, it iterates over these groups and uniformly associates a new semantic elevation label from the semantics set, each T /N processed groups. For example, let us consider the following semantic elevations: {*light*, *medium*, *steep*}. Besides, let us consider an ordered set S of groups as follows:

$S = \{G_i \mid i \in \{1,2,..,6\}\}$ with respectively the following terrain's inclination values: {5°, 10°, 15°, 20°, 25°, 30°}. For every two groups (as T = 6 and N = 3, T/N = 2), the automated discretisation assigns a new semantic elevation.

Let us compare these two discretisation approaches. On the one hand, the customized discretisation process allows users to freely specify the qualification of the terrain's elevations. Such qualifications are used by situated agents for spatial reasoning purposes. However, qualifications resulting from such a flexible approach deeply rely on the correctness of the interval bounds' values provided by users. Therefore, the customized discretisation method requires that users have a good knowledge of the terrain characteristics in order to guarantee a valid specification of inclination intervals. On the other hand, the automated discretisation process is also able to qualify groups' elevations without the need to specify elevation intervals' bounds. Such a qualification usually produces a visually uniformed semantic assignment. This method also guarantees that all the specified semantic attributes will be assigned to the groups without a prior knowledge of the environment characteristics.

Improving the Geometric Abstraction

The geometric abstraction algorithm produces groups that are built on the basis of their terrain's inclination characteristics. Using the extraction of elevation semantics, terrain's inclination is qualified using semantic attributes and associated with groups and with their cells. Depending on the classification intervals, adjacent groups with different area-weighted normal vectors may obtain the same elevation semantic. In order to improve the results provided by the geometric abstraction, we propose a process that merges adjacent groups which share the same semantic elevation (Fig.

7). This process starts by iterating over groups. Then, every time it finds a set of groups sharing an identical semantic elevation, it creates a new group. Next, cells composing the adjacent groups are registered as members of the new group. Finally, the area-weighted normal vector is computed for the new group. Hence, this process guarantees that every group is only surrounded by groups which have different semantic elevations.

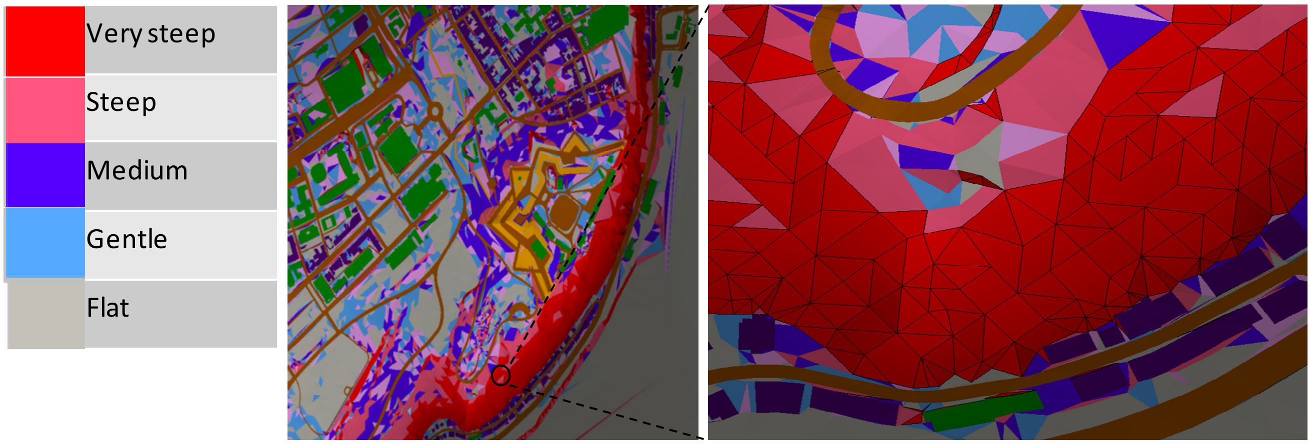


Figure 7: Extraction of elevation semantics: the left most figure presents the elevation qualification, the middle figure illustrates an overview of the initial geometric abstraction, and the right most figure shows a zoom in depicting adjacent groups with identical semantic elevations.

TOPOLOGICAL ABSTRACTION

In Section 4, we presented the generation of informed virtual geographic environments using an exact spatial decomposition scheme which subdivides the environment into convex cells organized in a topological graph structure. However, for large-scale and complex geographic environments (such as an entire city for example), such topological graphs can become very large. The size of such a topological graph has a direct effect on paths' computation time. In order to optimize the performance of path computation, we need to reduce the size of the topological graph representing the IVGE. The aim of the topological abstraction is to provide a compact representation of the topological graph suitable to situated reasoning and enabling fast path planning. However, in contrast to the geometric abstraction which only enhances the description of the IVGE with elevation semantics, the topological abstraction extends the topological graph with new layers. In each layer (except for the initial layer which is called level 0), a node corresponds to a group of nodes of the immediate lower level (Fig. **8a**). Indeed, the topological abstraction simplifies the IVGE description by combining cells (triangles) in order to obtain convex groups of cells. Such a hierarchical structure evolves the concept of Hierarchical Topologic Graph in which cells are fused in groups and edges are abstracted in boundaries (Fig. **8b**).

To do so, convex hulls are computed for every node of the topological graph. Then, the coverage ratio of the convex hull is evaluated as the surface of the hull divided by the actual surface of the node. The topological abstraction finally performs groupings of a set of connected nodes if and only if the group ratio is close to one. Let C be the convexity rate and CH (gr) be the convex hull of the polygon corresponding to G. C is computed as follows:

and $0 < C(G) \leq 1$

Indeed, the convex property of groups needs to be preserved after the topological abstraction. This ensures that an entity can move freely inside a given cell (or group of cells), and that there exists a straight path linking edges belonging to the same cell (or group of cells). Fig. **9** illustrates an example of the topological abstraction process and the way it reduces the number of cells representing the environment. In Fig. **9a**, we present the initial vector format GIS data of a complex building. Fig. **9b** depicts the initial exact spatial decomposition (sub-section 4.2) which yields 63 triangular cells. Fig. **9c** presents 28 convex polygons generated by the topological abstraction algorithm. The optimization rate of the number of cells representing the environment is around 55%.

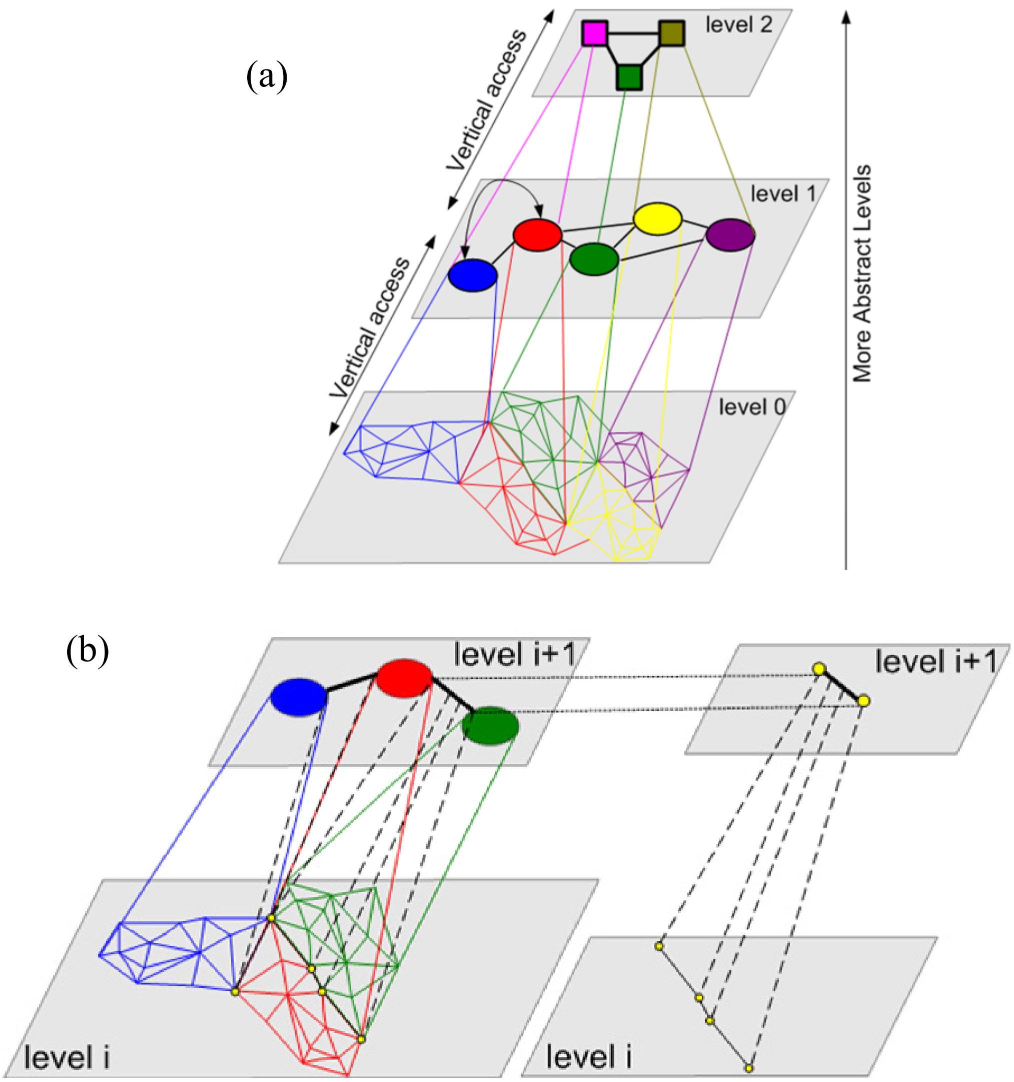
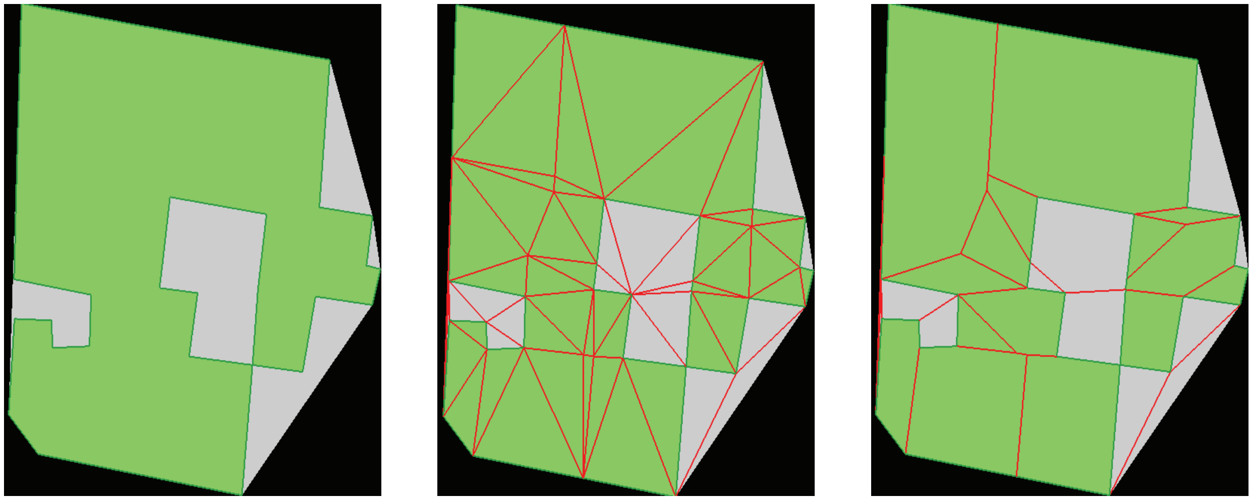


Figure 8: The topological graph extraction from space decomposition and extension into different levels using the topological abstraction; (a) topologic abstraction and convex cells; (b) abstraction of the cells' edges into groups' boundaries grouping.

To conclude, we proposed in this section two approaches aiming at enhancing the description of the IVGE. The first approach allows for qualifying the terrain's elevation with semantic information which is integrated in the IVGE. The second approach aims at simplifying large informed graphs corresponding to large scale and complex geographic environments. This approach reduces the number of convex cells by overlaying the informed graph with a topologically abstracted graph produced by a topological abstraction. The resulting IVGE is hence based on a hierarchical graph whose lowest level corresponds to the informed graph initially produced by the spatial decomposition. In the following section, we show how we use a well-known knowledge representation formalism to represent the semantic information in order to further enhance the IVGE description with respect to agents' and environments' characteristics.

SEMANTIC INFORMATION REPRESENTATION AND PROPAGATION

Two kinds of information can be stored in the description of an IVGE. Quantitative data are stored as numerical values which are generally used to depict geometric properties (like a path's width of 2 meters) or statistical values (like a density of 2.5 persons per square meter). Qualitative data are introduced as identifiers which can range from a word with an arbitrary semantic, called a label, to a reference to an external database or to a specific knowledge representation. Such semantic information can be used to qualify an



(a) (b) (c)

Figure 9: Illustration of the topological abstraction process with a strict convex property (C(gr) = 1); (a) the GIS data of a complex building; (b) the exact space decomposition using CDT techniques (63 triangular cells) ; (c) the topological abstraction (28 convex polygons)

area (like a road or a building) or to interpret a quantitative value (like a narrow passage or a crowded place). An advantage of interpreting quantitative data is to reduce a potentially infinite set of inputs to a discrete set of values, which is particularly useful to condense information in successive abstraction levels to be used for reasoning purposes. Furthermore, the semantic information enhances the description of the IVGE which, in turn, extends the agents' knowledge about their environment. However, the integration of the semantic information raises the issue of its representation. Therefore, we need a standard formalism which allows to precisely represent the semantic information that qualifies space and which is computationally tractable in order to be used by spatial reasoning algorithms. In the following sub-section, we present the Conceptual Graphs (CGs) formalism which is used to represent the semantic information in the IVGE. Next, we introduce the semantic abstraction process which takes advantage of both the CG representation and the hierarchical topological graph. This semantic abstraction aims at extracting a specialized view of the IVGE which takes into account the characteristics of the involved agent archetypes, the performed actions, and the locations where these actions occur.

Representation of the Semantic Information Using Conceptual Graphs

Sowa [36] introduced Conceptual Graphs (CGs) as a system of logic based on Peirce's existential graphs and on semantic networks which are widely used in artificial intelligence. We chose to use CGs in order to represent spatial semantic information for several reasons. First, CGs are known to express meaning in a form that is logically precise and computationally tractable [36, 37]. Second, CGs are proved to be efficient for spatial semantic information (spatial semantics) and geographic knowledge representation [16, 21, 22]. Third, CGs serve as an intermediate language to translate computer-oriented formalisms to and from natural language. This will help MAGS designers and analysts to express and to interpret spatial semantics and space qualification in a formalism that is close to natural language. Fourth, they provide extensible means to capture and represent real-world knowledge and have been implemented in a variety of applications for information retrieval [19], natural language processing [8], and qualitative simulations [16]. However, our proposal to use CGs to represent spatial semantics in virtual geographic environments and to reason about them is an innovative issue. Syntactically, a conceptual graph is a network of concept nodes linked by relation nodes. Concept nodes are represented by the notation [Concept Type: Concept instance] and relation nodes by (Relationship-Name). A concept instance can either be a value, a set of values or even a CG. The formalism can be represented in either graphical or character-based notations. In the graphical notation, concepts are represented by rectangles, relations by circles and the links between concept nodes and relation nodes by arrows. The most abstract concept type is called the universal type (or simply Universal) denoted by the symbol.

A MAGS usually involves a large number of situated agents of different types (human, animal, static, mobile, *etc.*) performing various actions (moving, perceiving, *etc.*) in virtual geographic spaces of various extents. Using CGs greatly simplifies the representation of complex situated interactions occurring at different locations and involving various agents of different types. In order to create models for MAGS, we propose to consider three fundamental abstract concepts: 1) agents; 2) actions; and 3) locations. Taking advantage of the abstraction capabilities of the CGs formalism (through the Concept Type Lattice (CTL), instead of representing different situated interactions of various agents in distinct locations, we are able to represent abstract actions performed in abstract locations by what we call agent archetypes. Moreover, we first need to specify and characterize each of the abstract concepts. The concept type lattice enables us to specialize each abstract concept in order to represent situated behaviors such as path planning of agents in space. Indeed, concept types are organized in a hierarchy according to levels of generality. However, this hierarchy is not a tree, since some concept types may have more than one immediate super-type [37].

Fig. **10** presents the first level of the concept type lattice refining the agent, action and location concepts that we propose. Fig. **11a**,**b** and **c** present the expansion of the concept type lattice presented in Fig. **10**. Indeed, Fig. **11a** illustrates some situated actions that can be performed by agents in the IVGE such as sailing for maritime vehicles, rolling for terrestrial vehicles, walking for humans, and access for humans to enter or exit buildings (we assume that buildings are not navigable locations). Besides, the location concept may be specialized into Navigable and Not Navigable concepts. Fig. **11b** depicts how the *Navigable* concept may also be specialized into *Terrestrial Vehicle Navigable*, *Pedestrian Navigable*, *Marine Vehicle Navigable*, and *Bike Navigable* which are dedicated navigable areas with respect to agent archetypes and environment's characteristics as specified by the elementary semantics. Finally, Fig. **11c** illustrates a few agent archetypes that are relevant to our geo-simulation example, including *pedestrians*, *cars*, *trucks*, and *bikes*.

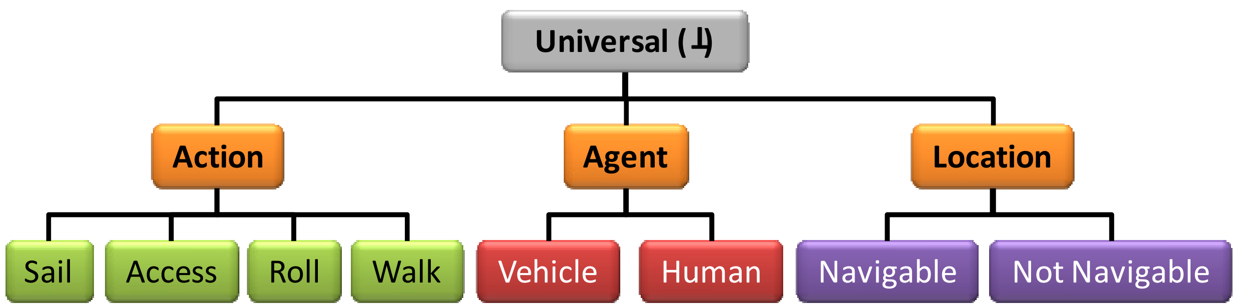


Figure 10: Illustration of the action, agent and location concepts using a concept type lattice.

In order to show how powerful such a representation may be, let us consider the following example. We want to build a MAGS simulating the navigation of three human agents (a man, a woman, and a child), two bike riders (a man and a woman), and three vehicles (a car, a bus, and a boat) in a coastal city. The navigation behaviors of these different agent archetypes must respect the following constraints (or rules): 1) pedestrian agents can only move on sidewalks, on pedestrian streets, and eventually on crosswalks if needed; 2) vehicles can move on roads and stop at parking lots; 3) boats sail on the river and stop at the harbor; and 4) bikes move on bikeways, roads, and pedestrian streets. Using standard programming languages, it might be difficult to represent or develop the functions related to such simple navigation rules which take into account both agents' and locations' characteristics. However, the representation of these navigation rules becomes an easy task when using CGs and our concept type lattice. Here are their expressions in CGs:

[PEDESTRIAN:*p]<-(agnt)<-[WALK:*w1]->(loc)->[PEDESTRIAN NAVIGABLE:*pn]

[VEHICLE:*v]<-(agnt)<-[ROLL:*r1]->(loc)->[TERRESTRIAL NAVIGABLE:tn]

[BOAT:*bo]<-(agnt)<-[SAIL:*s1]->(loc)->[MARINE NAVIGABLE:*mn]

[BIKE:*bi]<-(agnt)<-[RIDE:*r2]->(loc)->[BIKE NAVIGABLE:*bn]

The arrows indicate the expected direction for reading the graph. For instance, the first example may be read: an agent *p, which is a "pedestrian", walks on a location *pn which is "pedestrian navigable". Since this expression involves the concepts Pedestrian, Walk and Pedestrian Navigable, this rule remains valid for every sub-type of these concepts. Therefore, using CGs and the concept type lattice, there is no need to specify the navigation rules for men, women, and children if they act as pedestrians in locations such as pedestrian streets, sidewalks, or crosswalk.

Indeed, these agent archetypes are subtypes of the Pedestrian concept and pedestrian streets, sidewalks, and crosswalks are subtypes of the Pedestrian Navigable concept. To conclude, CGs offer a powerful formalism to easily describe different concepts involved in MAGS including agents, actions, and environments. It is important to mention that only elementary semantics are spatially situated in our IVGE, the remaining abstract and specialized semantics are conceptual semantics used by the semantic abstraction process described in Section 6.3.

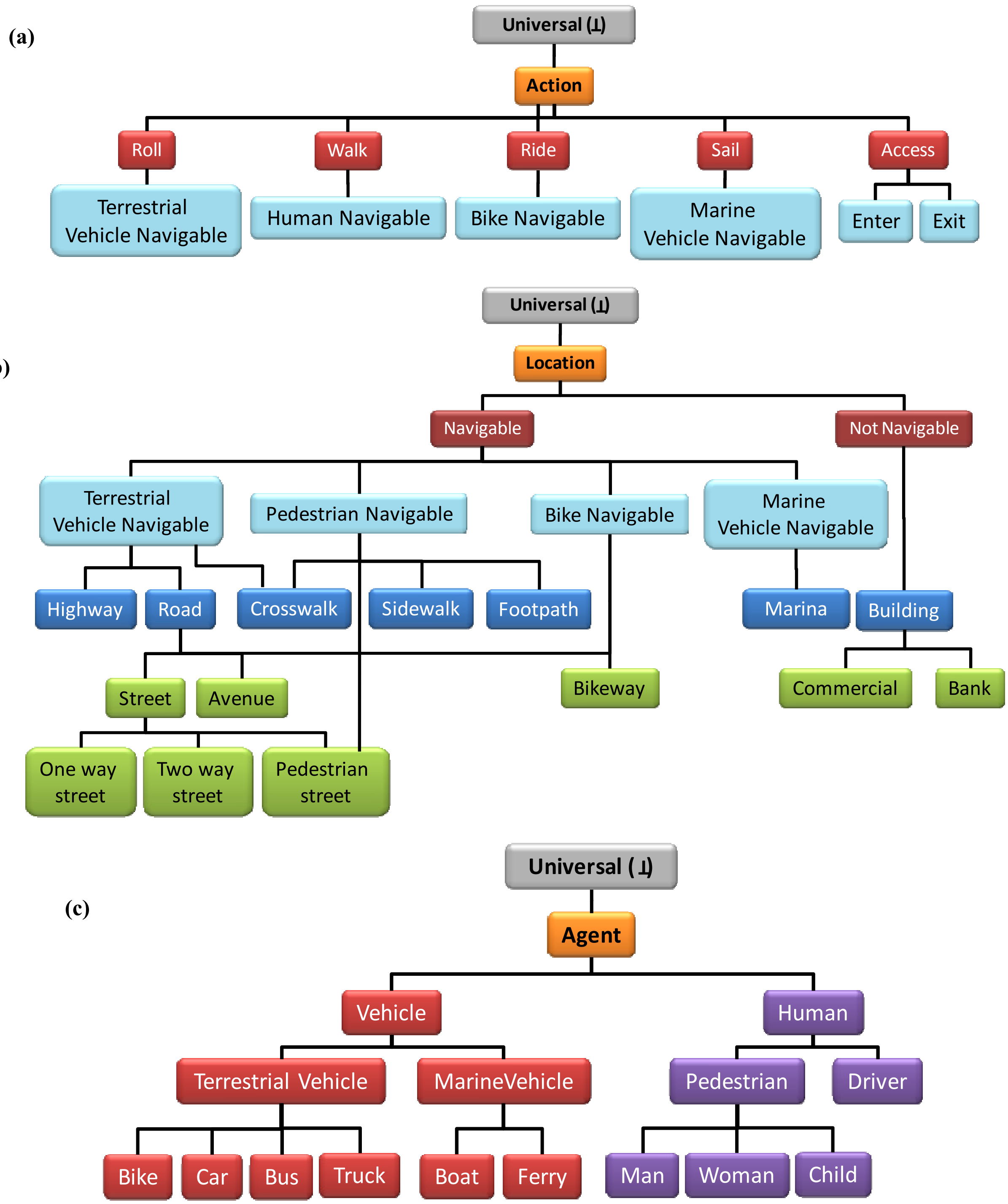


Figure 11: an example of a conceptual description of agent archetypes, actions performed, and locations situated in a geographic environment; (a) a specialization of the concept action; (b) a specialization of the concept location including the input semantics (dark blue); (c) a specialization of the concept agent.

Propagation of Input Semantics

The IVGE obtained as a result of the GIS importation technique emphasizes qualified areas by defining the semantics of their boundaries. But these informed boundaries are difficult to exploit when dealing with the semantics associated with a position, as for example if we want to check if a position is inside a building. This is why we propose to enhance the information provided by the IVGE by spreading the boundaries' semantics to the cells. Three related processes are necessary, and explained in the following subsections: *graph analysis*, *resolution of potential conflicts*, and *semantics assignment*.

The graph analysis is a traversal algorithm which explores the environmental graph while qualifying the cells towards a given semantic. This algorithm is applied to the entire graph for each semantic to be propagated. While exploring the graph, the algorithm collects three kinds of cells which are stored in three container structures for future use: *Inside* cells are within an area delimited by borders associated with the propagated semantic. *Outside* cells are outside any area defining the propagated semantic. *Conflict* cells are both qualified inside and outside by the algorithm.

Three parameters influence the traversal: 1) the semantic *sem* to propagate; 2) a set of *starting cells*, indicating where to start the exploration of the graph; a set is provided instead of a single cell in order to be able to manage disconnected graphs; and 3) a boolean value *startin* indicating whether the semantic must be assigned to the *starting cells* or not. The recursive algorithm is composed of four steps:

- **Step 1:** For each cell c in *starting cells*, proceed to *step 2*;
- **Step 2:** If *startin*=True let C be *Inside* and NC be *Outside*; else let C be *Outside* and NC be *Inside*. Proceed to *step 3*.
- **Step 3:** If c is in NC, transfer c to *Conflict*. If c is in *Conflict* or in C, Proceed to *step 1*. Put c in C, then for each neighboring cell n adjacent to c through a border b proceed to *step 4*.
- **Step 4:** If *sem* is defined for border b, inverse C and NC. Proceed to *step 3* and replace n by c.

After each graph traversal, we must deal with the cells that are potentially in conflict. Indeed, these cells must be assigned to either the Inside or the Outside container so that the system can continue with the next step. Cells are in con.ict when the shapes of two input features with the same semantic share a segment. Two alternative methods are proposed: 1) a fast assignment where the conflicting cells are arbitrarily transferred to one of the target containers, and 2) a deductive assignment where an algorithm selects the best option based on geometric considerations.

The arbitrary assignment is used when the internal details of a shape are not relevant to the target application. For example, Fig. **12a** presents two buildings. The building on the left-hand side has internal walls producing conflicting cells. These conflicts can be solved using the arbitrary assignment since they are not relevant when considering a building as an obstacle. The deductive assignment is used when the internal details of a shape are relevant. For example, Fig. **12b** presents four connected roads producing conflicting cells. They are resolved using the deductive method in order to determine which parts are effectively a road, and which parts are not. Both methods are carried out in two steps: 1) a local conflicting graph extraction which is the same for both methods, and 2) a decision step which is specific to each method.

The local conflicting graph extraction collects all the cells surrounding a conflicting cell, but only if they are reachable through a border which is not marked by the propagated semantic. Each orange zone in Fig. **12** shows an extracted local conflicting graph. Every time a cell is discovered, it is transferred from the global container Conflict to a local container. Then, the algorithm recursively explores the neighbors which are reachable through a border which is not marked by the propagated semantic, transferring them to the local container. At the end, the algorithm obtains a set of local conflict containers, corresponding to the amount of local graphs considered in conflict.

The decision part of the arbitrary assignment only consists of transferring the local conflicting cells to one of the Inside or the Outside containers. The decision part of the deduction algorithm is based on geometric considerations. If the local conflicting zone is mainly surrounded by Outside cells, then the con.ict is resolved as Inside, and vice-versa. In order to check the surrounding of a conflicting zone, three perimeter lengths are computed: the inside perimeter P_{in} which is the total length of the zone's boundaries connected to Inside cells; the outside perimeter P_{out} for Outside cells; and the unknown perimeter P_{ukn} when connected to Conflict cells. Then, three cases can occur:

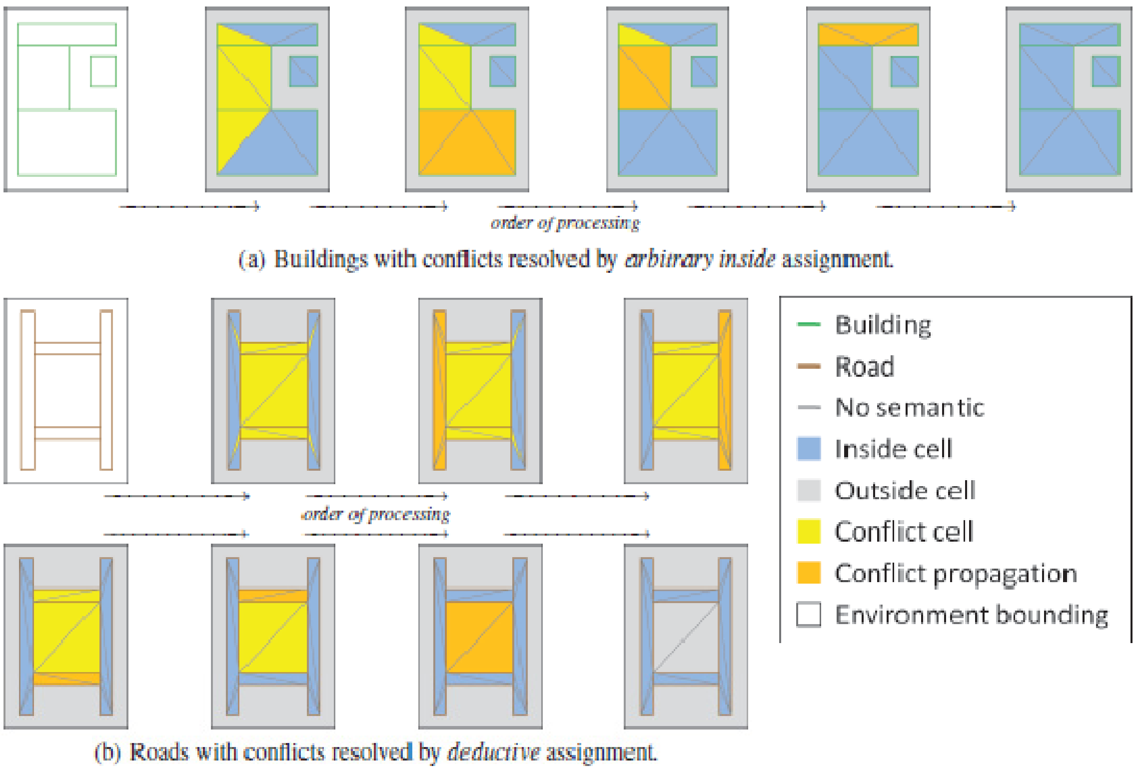


Figure 12: Two examples of the conflicting cells resolutions, where the left most figure is the original environment, the second figure is the graph analysis, the last figure is obtained after the resolution of all conflicts.

1. If $P_{in} > P_{out} + P_{ukn}$ the conflicting cells are assigned to Outside;
2. If $P_{out} > P_{in} + P_{ukn}$ the conflicting cells are assigned to Inside;
3. Otherwise, none of the inside and outside perimeters are significant and the decision will be re-evaluated during a subsequent step.

The third case requires a second decision step for all undecided conflicts, where the unknown perimeter is ignored in order to force a decision. The need for two decision steps is illustrated in Fig. **12b**: if the last conflict resolution had been made first, the decision would have been impossible, the zone being totally surrounded by unknown cells; the displayed resolution order avoids any decision failure, only requiring one step.

The last step of the semantic propagation consists of assigning the final semantics to the cells. The process is quite simple: each propagated semantic is assigned to all the corresponding *Inside* cells. One can notice that some cells may have multiple semantics when they are present in more than one *Inside* container. Additionally, it is possible to keep track of the Outside cells by assigning them a negative semantic, as for example in order to know what is not a road in the environment is. Finally, an optional process can be performed to remove the borders' semantics of some detected conflicting cells. Indeed, such borders may distort some spatial reasoning algorithms. For example, when considering road borders as obstacles to plan a path, a simulated vehicle would not be able to go through some passageways (*i.e.* left hand side of Fig. **13b**). After resolution, the semantic of the problematic borders is removed, making them crossable (*i.e.* right hand side of Fig. **13b**). These problematic borders are the ones that are marked with a propagated semantic and which connect two Inside cells. One can note that only the cells previously detected as conflicting need to be tested.

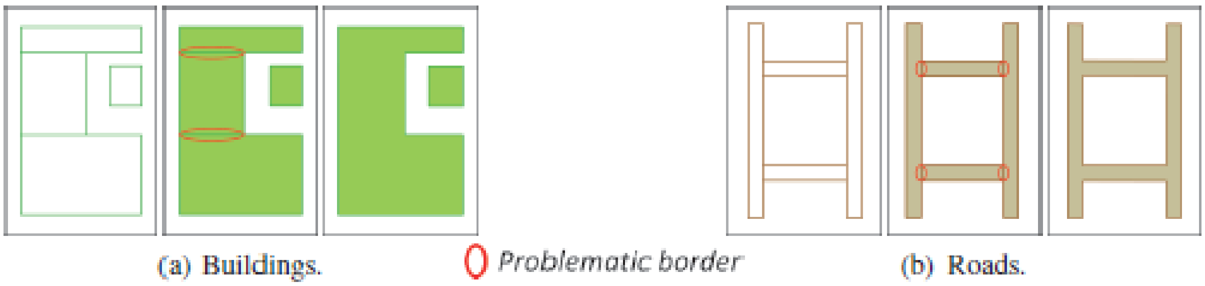
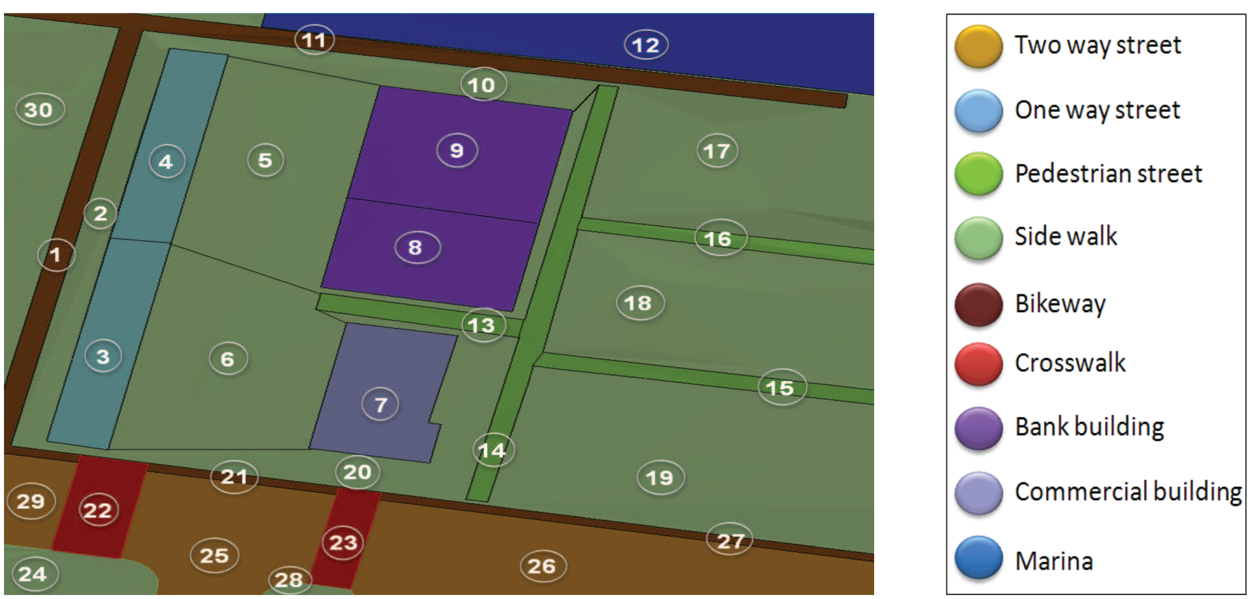


Figure 13: Two examples of the result of borders' semantics propagation. In both figures: left is the original environment, centre is obtained after con.ict resolution, right is obtained after assigning semantics and filtering conflict borders.

Semantic Abstraction

Depending on their characteristics and goals, agents interact with the environment in different ways. For example, the spatial semantics associated with each cell can be used by agents to navigate in favored areas (for example, a pedestrian might avoid navigation in areas such as roads while preferring walking on sidewalks, crossings, *etc.*). The aim of the semantic abstraction is to provide a semantic representation of the topological graph for each agent archetype (as a human and a car). We call such a representation an abstract view. While navigating, agents can refine this abstract view by querying local or global refinements on cell types or on graph structures. The semantic abstraction provides multi-level views of the IVGE in which agents are navigating. A specific view can be generated, using the semantic abstraction, for each agent archetype, depending on its goals and needs. In this way, an agent can retrieve abstract views of the environment containing relevant information for its decision making. For example, an agent simulating a virtual human navigating in a conventional VGE would not distinguish between a road and a sidewalk since there is no difference between them in the virtual environment. However, using our semantic abstraction, this agent will be able to adapt its behavior with respect to the environment's characteristics in order to "*speed up when crossing the road*" or to "*cross the street using crosswalk areas*"

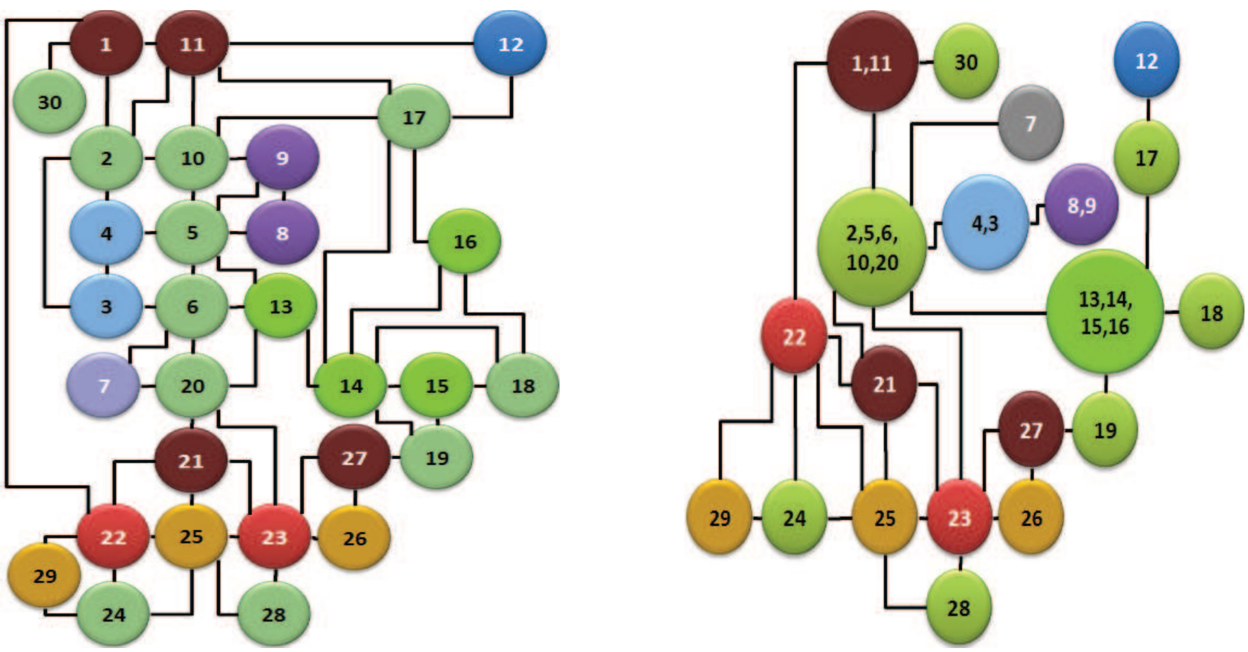


(a) A sample virtual environment (b) Legend

Figure 14: (a) a virtual environment of a part of Quebec City situated near the marina; (b) the legend.

In order to build such a semantic abstraction, we first define a semantic type hierarchy as well as a set of semantic expressions (rules) linking various semantic concepts. The semantic type hierarchy describes the super-type/sub-type links between semantic concepts. Fig. **11b** shows an example of a semantic type hierarchy for the concept location. This semantic type lattice includes various semantic concepts such as "road" which is a subtype of "*Terrestrial Vehicle Navigable*" and a super-type of "*street*". Fig. **15a** presents a simplified example of an informed graph involving semantics such as *sidewalk*, *crosswalk*, *one-way street*, *two-way street*, *pedestrian street*, *bikeway*, *etc.* Taking advantage of our semantic type lattice and of the graph structure of the IVGE, we are able to semantically abstract the informed topological graph and to build a new semantic representation of the environment. This new semantic graph overlays the topological graph and enables the creation of a hierarchical graph structure of the IVGE which is useful for situated reasoning such as path planning.

In order to explain the semantic abstraction process, let us consider the virtual environment depicted in Fig. **14a** which represents a part of Quebec City. The color codes of the spatial semantics are explained in Fig. **14b**. Besides, Fig. **15a** illustrates the informed topological graph associated with this geographic area. The first level of semantic abstraction consists of grouping cells sharing identical elementary semantics as illustrated in Fig. **15b**. This first abstraction level only simplifies the environment structure by reducing the number of cells. The semantic concept type diagram is then used to guide the computation of the second level of semantic abstraction. The second level of semantic abstraction takes into account the characteristics of a specific category of situated actions occurring in a specific location and involving a specific type of agents. For example, in Fig. **16a**, adjacent cells of type bikeway, crosswalk, one-way street, two-way street are grouped as navigable areas for bike riders. Indeed, we assume that bike riders are neither allowed to ride on sidewalks nor on pedestrian streets. In addition, in Fig. **16b**, adjacent cells of type crosswalk, one-way street, two-way street are grouped as navigable areas dedicated to cars. Here, we suppose that bikeways are prohibited areas for cars.

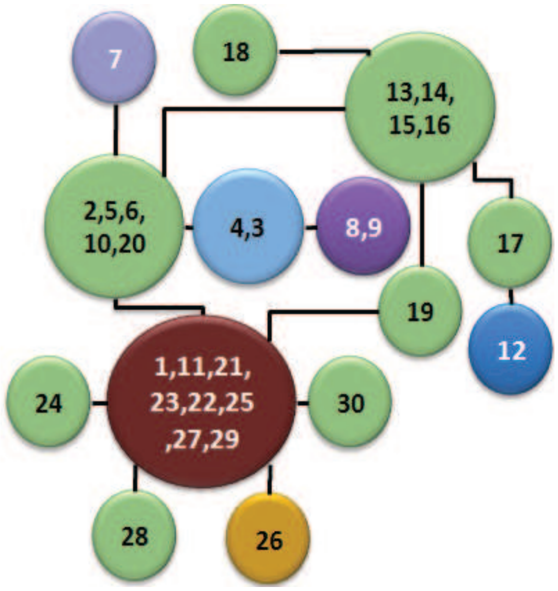


(a) Informed topological graph. (b) The first level of the semantic abstraction.

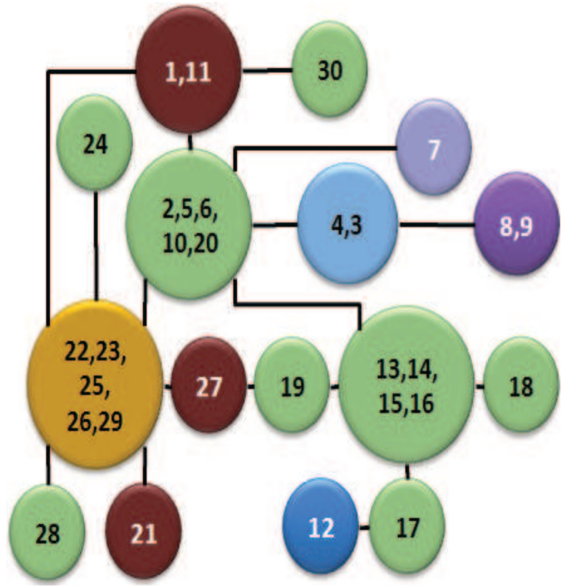
Figure 15: (b) is the first level of the semantic abstraction (adjacent cells sharing identical semantic information are grouped together) of the informed topological graph (a).

To conclude, once the geometric, topologic, and semantic abstraction levels are computed, an agent can plan a path inside a view of the IVGE which takes into account both the environment's and the agent's characteristics. Indeed, Figs. **14** and **15** illustrate the first and second levels of the semantic abstraction process while taking advantage of the semantic type lattices we proposed in Fig. **11**. The semantic abstraction allows the creation of the hierarchical graph structure of the IVGE which is adapted to hierarchical decision processes.

Such a hierarchical graph structure allows precise queries and local refinements and enables the description of complex reasoning processes based on the environment topology. It also supports high-level spatial reasoning processes such as hierarchical path planning that wediscuss in the following section.



(a) The second level of the semantic abstraction based on the bike navigation semantic lattice defined in Figure 10. Note how *bikeway* nodes are grouped with adjacent *one-way street* and *two-way street* nodes.



(b) The second level of the semantic abstraction based on the car navigation semantic lattice defined in Figure 10. Note how *crosswalk* nodes are grouped with adjacent *one-way street* and *two-way street* nodes.

Figure 16:. Second level of the semantic abstraction based on the *bike* and the *car* navigation semantic lattices; (a) note how *bikeway* nodes are grouped with adjacent *one-way street* and two-way street nodes; (b) note how *crosswalk* nodes are grouped with adjacent *one-way street* and with *two-way street* nodes.

HIERARCHICAL PATH PLANNING

The displacement of a human agent in a virtual environment relies on its capability to move from one starting position to reach a final destination position, both of them being situated in the IVGE. This is usually known as the path planning problem: how to find a path from a starting location up to a target position, while avoiding obstacles located in the environment? In order to generate a plausible path, we need to minimize certain criteria, such as the distance. However, other criteria related to the environment's characteristics such as slopes, may be considered in order to generate plausible paths. Indeed, the movement of a human in a geographic environment is often constrained by the terrain's characteristics. For example, a steep slope or a ravine may be obstacles which must be avoided. Taking advantage of our topological graph structure, the path planning problem can be reduced to a graph traversal algorithm which minimizes the sum of weights associated with the edges linking nodes. In addition, the geometric abstraction helps qualify the elevation of the terrain using elevation semantics. This semantic information is stored in the cells that correspond to nodes of the graph topology. Therefore, the path planning which takes into account the elevation of the land is easily calculated using the IVGE model that we propose.

Several graph algorithms can be used to compute the shortest path. Among these, the most often used algorithm is A* [31] whose approach is as follows: each explored node is assigned a value corresponding to the estimated total cost of the path traversing that node. This value is computed using the current distance from the initial position which is added to the estimated distance to the target position (provided by a heuristic h). The convergence efficiency of the A* algorithm mainly depends on the quality of the heuristic h. In the case of a shortest path search, this heuristic is often an estimate of the distance. However, this heuristic needs to be adapted in order to take into account the terrain elevation which is described by elevation semantics and associated with the visited nodes. In contrast to A*, Dijkstra's algorithm

[24] starts with the origin node (the current position) and successively explores all the successors, while favoring those which are closest to the origin, and so until the fulfillment of a stopping condition. The stopping condition is usually based on the final destination position, but needs to be adapted in order to take advantage of the spatial semantics provided by the IVGE. Indeed, the major issue faced by a path planner in a complex geographic environment is the large number of cells (including triangles, hexagons, or squares) produced by the space decomposition process (sub-section 4.2). The topological and semantical abstractions that we propose allow us to reduce the size of the topological graph. However, we still need efficient path planning approaches when dealing with large scale and complex geographic environments. An effective method for the computation of paths in such geographic environments is to search within a smaller abstract part of the space. Abstractions (topologic and semantic) can be used to factor a search problem into several smaller problems and thus, allow agents to reason about path planning strategies in terms of abstracted views of the environment. This is known as hierarchical path planning [17]. In the following sub-sections, we first provide an overview of current path planning techniques. Next, we present the algorithm that we propose to implement the hierarchical path planning while taking advantage of the IVGE enhanced with CGs.

An Overview of Current Path Planners

The majority of current path planners, including recent hierarchical planners [5, 9, 14, 38] do not take into account the environment's characteristics (topological and semantic) and the agents' types and capabilities. For example, they assume that all agents are equally capable of reaching most areas in a given map, and any terrain portion which is not traversable by one agent is considered to be not traversable by the other agents. Further assumptions are often made about the characteristics of each agent: a path computed for one agent is equally valid for all the agents because agents' categories are not distinguished. Such assumptions limit the applicability of these techniques that can only solve a narrow set of problems: path planning of homogeneous agents in a homogeneous environment.

Besides, two recent hierarchical triangulation-based path planning approaches, namely Triangulation A* and Triangulation Reduction A*, which are relevant to our work are described in [9]. TA* makes use of the Delaunay Triangulation (DT) technique to build a polygonal representation of the environment without considering the semantic information. This results in an undirected graph connected by constrained and un-constrained edges; the former being traversable and the latter not. TRA* is an extension of TA* and abstracts the triangle mesh into a structure resembling a roadmap. Like our method, both TA* and TRA* are able to accurately answer agents' path queries since they make use of the DT technique. The abstraction technique used by TA* and TRA* however are notably different from our work. Where we aim to topologically abstract the IVGE by merging triangles into convex polygons, they rather aim to maximize triangle size. We also handle semantically enriched environment descriptions including qualification of space and terrain's elevation while both TA* and TRA* assume a homogeneous flat environment.

In this section, we address the issue of path planning for agents having different capabilities and evolving in environments with various characteristics. We propose an algorithm that implements hierarchical path planning and takes advantage of the hierarchical graph structure of the IVGE and of the rich semantic information provided by CGs in order to compute paths in semantically-enriched IVGE.

Hierachical Path Planning Algorithm

Let us consider the topological graph extracted from the exact spatial decomposition before highlighting the usefulness of the topological and semantical abstractions. Since cells are convex, it is possible to build an obstacle free path by linearly connecting positions located at two different borders belonging to a given cell.
Thus, it is also possible to use borders, represented by edges in the graph, to compute obstacle free paths between different locations in the environment. Since the topological graph structure is hierarchical, each node at a given level i (except at level 0) represents a group of convex cells or abstract cells of a lower level i - 1. Hence, our approach can be used to compute a path linking two abstract nodes at any level.

Let us consider a hierarchical topological graph G composed of i levels. Nodes belonging to level 0 are called leaves and represent convex cells produced by the exact spatial decomposition. Nodes belonging to higher levels ($i > 0$) are called abstract nodes and are composed of groups. Given a starting position, a final destination, and a hierarchical topological graph G composed of i levels, the objective of our algorithm is to plan a path from the current position to the destination using G. The algorithm starts from the highest level of the hierarchy and proceeds as follows:

- **Step 1:** Identify to which abstract nodes belong the starting position and the final destination.

Two cases need to be considered:

  - Case 1: Both are in the same abstract node *k* at level *i*. Proceed to *step 1* with the groups (at level *i - 1*) belonging to node k.
  - Case 2: They are in different abstract nodes *k* and *j* at level *i*. Proceed to *step 2*.

- **Step 2:** Compute the path from the abstract node *k* to the abstract node *j*. For each pair of consecutive nodes (*s*, *t*) belonging to this path, two cases are possible:
  - Case 1: Both are leaves. Proceed to *step 4*.
  - Case 2: Both are abstract nodes. Proceed to *step 3*.
- Step 3:
  - If the starting position belongs to *s* then identify to which group *gs* of s it belongs and proceed to *step 2*, in order to compute the path from the abstract node *gs* to the closet common boundary with the abstract node *t*. Else proceed to *step 2* in order to compute the path from the center of the abstract node *s* to the closet common boundary with the abstract node *t*.
  - If the final destination position belongs to *t* then identify to which group *gd* of t it belongs and proceed to *step 2*, in order to compute the path from the closet common boundary with the abstract node *s* to *gd*. Else proceed to *step 2* in order to compute the path from the closet common boundary with the abstract node *s* to the centre of the abstract node *t*.
- **Step 4:** Once in a leaf, apply a path planner algorithm (we use the Djkstra and A* algorithms) from the starting position to the final goal using the convex cells which belong to the informed graph.

The strategy adopted in this algorithm is to refine the path planning when getting closer to the destination. The algorithm starts by planning a global path between the start and the destination abstract nodes (step 1). Then, for each pair of successive abstract nodes, it recursively plans paths between groups (of lower levels) until reaching leaves (steps 2 and 3). Once at leaves (convex cells at level 0), the algorithm proceeds by applying a path planning algorithm such as Dijkstra and A* (step 4). Hence, at level *i*, the path planner exploration is constrained by the nodes belonging to the path computed at level *i + 1*.

Moving agents can use this algorithm in order to plan paths within the IVGE. The path computed in step 2 is actually a coarse grain path whose direction is only indicative. Since the path is refined in a "*depth-first*" way, agents can perform a local and accurate navigation inside an abstract node without requiring a complete and .ne grained path computation towards the final destination. The lower levels' sub-paths (related to other abstract nodes) are computed only when needed, as the agent moves. Such a "'just in time"' path planning approach is particularly relevant when dealing with dynamic environments. Indeed classical path planning approaches use the entire set of cells representing the environment and compute the complete path between a starting position and a final position. These classical approaches suffer from two major drawbacks: 1) the computation time of a path is considerable since it involves all the cells composing the environment; 2) the planned path may become invalid as a consequence of environment changes. An interesting property of our hierarchical path planning approach is the optimization of calculation costs over time. Indeed, the entire path is only computed for the most abstracted graph, which contains a small number of abstract nodes compared to the informed graph (convex cells at level 0). In addition, our approach provides a just in time path planning process which allows to handle the environment's dynamics. Furthermore, this hierarchical path planning is adapted to any type of agents; whenever we are able to generate the abstracted graphs taking into account both the geographic environment and agents' characteristics.

RESULTS

In this section, we present the results of the implementation of our IVGE generation approach. First, we introduce IVGE-Viewer, a software that we developed in order to automatically generate semantically-enhanced virtual geographic environments and to inspect the environment partitioning in 2D and 3D. Next, we show how we take advantage of the multi-level graph structure and the semantically-enriched description of the IVGE in order to support agents' path planning, taking into account both the environment's and the situated agents' characteristics. We only present the standard path planning algorithm. The Hierarchical Path Planning algorithm introduced in Section 7 is detailed in [28].

Ivge-Viewer Tool

The proposed environment extraction method is used to create an accurate semantically-enhanced virtual geographic environment (IVGE) and provides the advantages of the semantic merging of grids along with the accuracy of vector data layers. Using the automatic extraction method that we propose, our system handles the IVGE construction directly from a specified set of input vector GIS files. The performances of the extraction process are very good, being able to process an area such as the centre part of Quebec-city, with one elevation map and five semantic layers, in less than five seconds on a standard computer (Intel Core 2 Duo processor 2.13Ghz, 1Go RAM). The obtained unified map approximately contains 122,000 triangles covering an area of 30km2. Besides, the necessary time to obtain the triangle corresponding to a given coordinate is negligible (less than 10-4 seconds). Moreover, the geometric abstraction produces approximately 73,000 groups of cells in 2.8 seconds. Besides, the custom and automated discretisation processes are performed respectively in 1.8 and 1.2 seconds using eight semantic elevation labels.

Regarding the topological abstraction, we tested our model using two abstraction levels. The first level fully respects the convexity (*i.e.* convex ratio = 1) and produces around 87, 000 convex groups of cells, which corresponds to a reduction of almost 28%. The second level performs groupings of adjacent convex groups if and only if the convex ratio is better than 0.9. The second level yields approximately 28,000 cells which represent 22% of the total number of convex cells produced by the spatial decomposition.

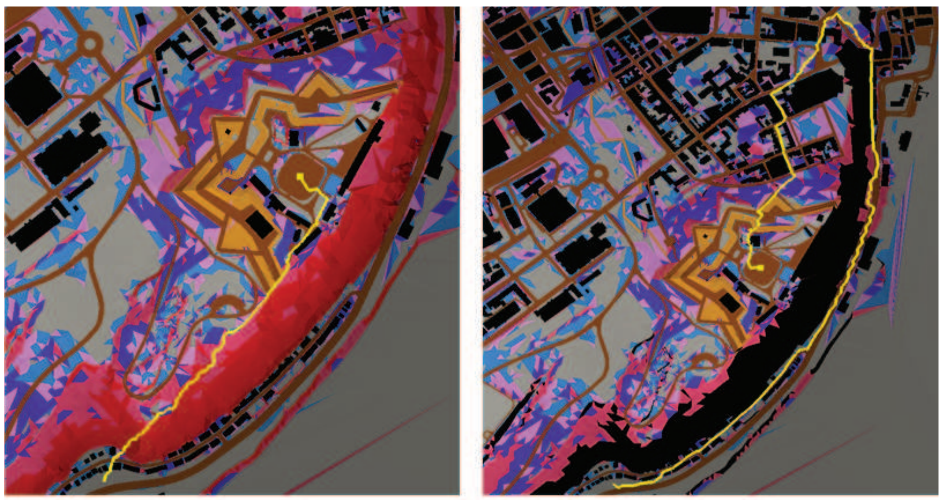
Regarding the semantic abstraction, we used the elementary semantic concepts illustrated in Fig. **5** together with the semantic concept type hierarchy detailed in Fig. **10**. The semantic abstraction reduced the number of abstract nodes belonging to the second level of the topological graph from 28,000 to approx-imately 17,000.

Figure 17: 2D map visualization of the GIS data; (1) unified map; (2) selected position's information including geometric and semantic data.

The IVGE-Viewer tool provides two visualization modes for the computed data. First, a 3D view (as shown in Fig. **6d**) allows the user to freely navigate in the virtual environment. We propose an optional mode for this view where the camera is constrained at a given height above the ground, allowing the user to follow the elevation variations when navigating. Second, we propose an upper view with orthogonal projection to represent the GIS data as a standard map. In this view the user can scroll and zoom the map (1 in Fig. **17**), and can accurately view any portion of the environment at any scale. Additionally, one can select a position in the environment in order to retrieve the corresponding data (2 in Fig. **17**), such as the underlying triangle geometry, the corresponding height and the associated semantics, including semantic elevation.

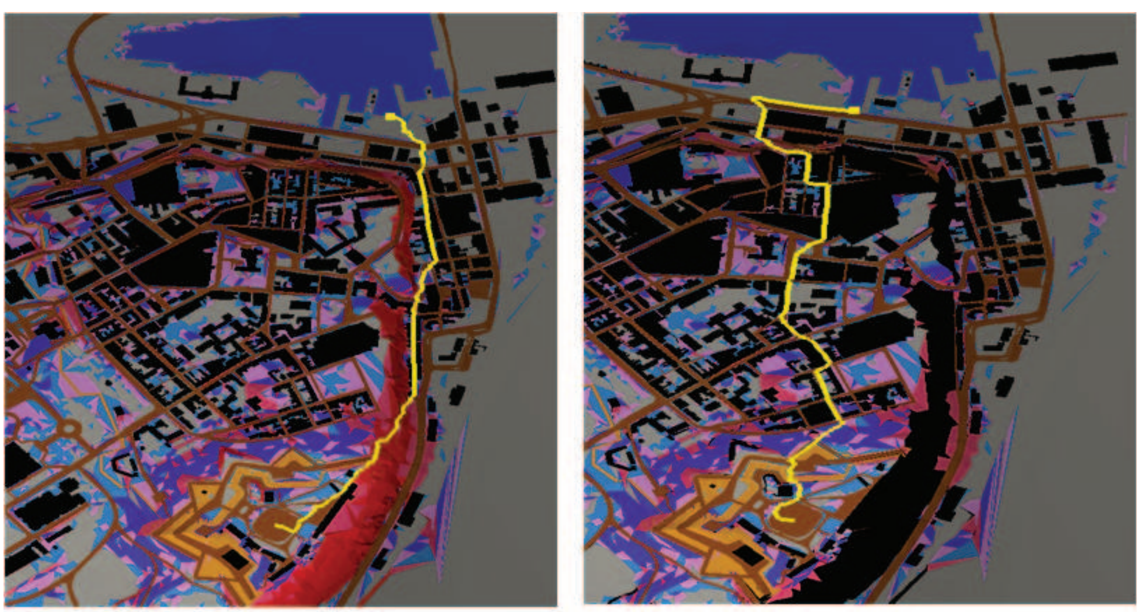
Standard Path Planning with Constraints

Fig. **18** illustrates two paths linking two locations situated in the IVGE. Fig. **18a** shows a path planning (colored in yellow) which does not take into account the terrain elevation. This path only avoids obstacles such as buildings, walls which are colored in black. Therefore, the path crosses an area colored in red which represents a steep slope. In Fig. **18b**, the algorithm has generated a path which respects both terrain's elevation and obstacles in the IVGE. Indeed, the steep slopes (initially colored in red) are avoided since they are now considered as obstacles (colored in black). This path is longer, but it fully respects the constraints of the environment and the elevation of the terrain.



(a) (b)

Figure 18: Path planning linking two locations in the IVGE (the computed path is colored in yellow). (a) path computed with no respect top terrain's elevation. (b) path computed with respect to terrain's elevation.



(a) (b)

Figure 19: Search path to get to a place in the IVGE (place described by semantics). The calculated path is colored in yellow; (a) without taking into account the elevation of the land; (b) taking into account the ground elevation.

The movement of a human agent in a virtual environment may not aim at reaching a particular position in the IVGE but rather a particular area. This type of path planning answers the question: how to find a path to reach a specific area while avoiding obstacles located in the environment? The semantic information integrated in the IVGE description helps answering such a question. Indeed, semantic information is used to describe spatial objects (building, house, marina, wall of the old city, *etc.*). Using the topological graph and using the Dijkstra algorithm the system computes the shortest path to reach a specific area located in the IVGE. In our case, the stopping condition of the algorithm is related to the semantics of the visited nodes. If the semantics associated with the visited node correspond to the target areas' characteristics, the algorithm stops and the path is generated. To illustrate the path planning for a target area qualified by one or several semantics (instead of a target position) in the IVGE, we propose the following example: a tourist who moves using a rolling chair is located inside the old city of Quebec. This tourist wants to visit an attraction spot called the marina. Hence, the marina is not identified by coordinates (x, y, z), but rather by a semantic information. Fig. **19** shows the computed path to reach the marina (the marina is colored in blue at the top of the figure). On the one hand, Fig. **19a** shows a path which does not take into account the terrain's elevation. This path avoids the obstacles of the environment (buildings colored in black) but crosses steep slopes areas (colored in red). Such a path is obviously not acceptable for our tourist. On the other hand, Fig. **19b** highlights a path planning which avoids steep slopes (colored in black) as well as obstacles situated in the IVGE (buildings colored in black).

DISCUSSION, CONCLUSION, AND FUTURE WORKS

In this chapter, we proposed an accurate and automated approach for the generation of semantically-enhanced and geometrically-accurate virtual geographic environments (IVGE) using GIS data. This novel approach offers several advantages. First, the description of the IVGE is realistic since it is based on standard GIS data. Besides, this description is also accurate because it is produced by an exact spatial decomposition technique which uses data in a vector format. Hence, this description preserves both the geometrical and the topological characteristics of the geographic environment and enables a graph-based description of the virtual environment. The topological approach goes beyond grid-based techniques by combining the semantic information merging and the vector-based representations accuracy. Table **1** provides a comparative table of our IVGE topological approach with vector-based and grid-based approaches for the representation of virtual environments. For each approach, the table highlights its advantages (in green) and drawbacks (in red). First, the grid-based approach is the most commonly used for the representation of virtual environments. This approach is simple, easy to build, and offers a fast access to data. However, it is very sensitive to the scale at which the data are represented. In addition, merging provided by various sources in different formats can cause problems of accuracy and consistency. Moreover, the complexity of the grid-based representation depends on both the geometry of the environment and the data resolution (scale). The second approach for the representation of virtual environments is the vector-based method. This method is known for its precision as it preserves the geometrical characteristics of the environment. Besides, the complexity of the vector-based representation only depends on the geometry of the environment. However, access, manipulation and merging of data is not as easy as it is using the grid-based approach. Finally, our topological approach combines these two methods (grid-based on vector-based) in order to compensate their respective drawbacks. The representation of the environment as a topological graph allows us: 1) to preserve the accuracy of the geometry of the environment generated by the exact spatial decomposition technique; 2) to benefit from graph theory to easily and rapidly access the representation of the virtual environment; 3) to take advantage of the extensible structure of graphs in order to enrich the description of the IVGE using semantic information.

Table 1: Advantages and drawbacks of the existing environment representation approaches com-pared to our topological model.

Grid	Vector layer	Topological
✔ Easy to fill ✔ Multiple semantics merging ✔ Easy to examine - - - - - - - - ✖ Constant scale ✖ Problem of accurate localisation ✖ Balance accuracy / memory complexity	✔ Adaptable scale ✔ Memory complexity depends on geometry ✔ Exact representation of input data (accuracy) - - - - - - - - ✖ Hard to examine ✖ Very hard to merge spatially overlapping semantic data	✔ Adaptable scale ✔ Memory complexity depends on geometry ✔ Exact representation of input data (accuracy) ✔ Multiple semantics merging ✔ Easy to examine ✔ Graphs algorithms and improvements

The description of the IVGE is also enriched with spatial semantics automatically extracted from input data and processed from the terrain's elevation. Regarding the representation, propagation, and exploitation of the semantic information, we use the CG formalism, which constitutes another original aspect of this work. First, we demonstrated how easily we can represent spatial semantics, how we take advantage of the concept type lattice to specify rules at the abstract concept level, and how they remain valid for more specialized concepts. Second, we outlined how the propagation of semantic information from the borders to the cells improves the description of the IVGE and provides a coherent qualification of space (such as roads and buildings). Third, we pointed out how we enhance the hierarchical topological graph using the concept type lattice in order to build different views of the IVGE which take into account both the environment and the agent type's characteristics.

The main outcome of such a semantically-enhanced and geometrically accurate virtual geographic envi-ronments concerns agents' situated reasoning capabilities such as path planning in large scale and complex geographic environments. We proposed a hierarchical path planning algorithm (using Dijkstra and A*) which takes advantage of our IVGE model to provide paths which take into account the agents' and environment's characteristics. Furthermore, we highlighted the applicability of our model with the IVGE-Viewer tool. Finally, our model not only supports path planning between two locations (positions) in the IVGE, but also path planning toward a semantically-qualified area situated in the IVGE. We are currently working on the implementation of the hierarchical path planning algorithm which is closer to human behavior by offering more plausible geo-simulations of virtual humans displacements.

When considering the coupling of agent-based models and GIS-based spatial data models, Browns and his colleagues [6] distinguish four key relationships affecting how spatial data model and agent-based process models interact: identity, causal, topological, and temporal. In our IVGE model, agents can interact with a semantically-enhanced description of the virtual environment. Thus, the identity relationship is supported using both the classification of situated agents' archetypes and the qualification of spatial features. Regarding the causality relationship, agents have the ability to perform actions that affect spatial features and/or their attributes while taking into account both their own characteristics and the environment's enriched description. The IVGE also provides situated agents with the geometric and topologic information required for navigation and path planning processes.

Nevertheless, the temporal dimension is essential to understand and model the dynamics of geographic phenomena [11]. Since time can only be perceived through its effects, capturing the temporal dimension requires detection and analysis of the changes that affect agents evolving within the virtual geo-graphic environment [25]. In order to support agents' interactions within a virtual geographic environment, an explicit and adequately linked representation of space and time needs to be developed [11]. Currently, our IVGE model does not support the temporal relationship. Indeed, such a relationship is conceptually part of the agent-based simulation engine. Nevertheless, we are currently elaborating a multi-agent geo-simulation platform that leverages our IVGE model for the support of micro-simulation of agents' spatial behaviors in virtual geographic environments. We can also take advantage of the work done by Haddad and Moulin [16] on the use of cognitive archetypes and conceptual graphs to model dynamic phenomena in spatial environments. Haddad and Moulin's approach might extend our IVGE model in order to support spatio-temporal models and thus the temporal relationship.

In the future, we expect to further improve the description of the IVGE by integrating enriched knowledge representations (called the environment knowledge) using CGs aiming at assisting situated agents' interactions with the IVGE depending on their archetypes, on the environment's characteristics, and on their goals. The above-mentioned contributions of our model open opportunities for several applications in a variety of application domains. Examples of such domains include the entertainment industry (games and movies), security planning and crowd management (planning events involving large crowds such as demonstrations, popular celebrations such as soccer games, and religious celebrations), and military operations in urban settings involving civilian crowds.

ACKNOWLEDGEMENT

This research has been supported by the Canadian Defence (DRDC Valcartier), and the Canadian Network of Centres of Excellence in Geomatics (GEOIDE). M. Mekni benefited from a PhD scholarship granted by FQRNT (Fonds Québécois de la Recherche sur la Nature et les Technologies).

REFERENCES

[1] W. Ali and B. Moulin, '2D-3D multi-agent geosimulation with knowledge-based agents of customers' shopping behavior in a shopping mall' in Spatial Information Theory, pp. 445–458, 2005.

[2] R. Andersen, J. L. Berrou, and A. Gerodimos, 'On some limitations of grid-based (ca) pedestrian simulation models', in First International Workshop on Crowd Simulation (V-Crowds'05), VRlab, EPFL, 2005.

[3] M. Batty, J. Barros, and S. J. Alves, 'Cities: Continuity, transformation and emergence', In CASA Working Paper no.72, Centre for Advanced Spatial Analysis, University College London, 2004.

[4] I. Benenson and P. Torrens, 'Geosimulation: Automata-based modeling of urban phenomena, Wiley, 2004.

[5] A. Botea, M. Müller, and J. Schaeffer, 'Near optimal hierarchical path-finding', Journal of Game Development, vol. 1, pp. 7–28, 2004.

[6] D. G. Brown, R. Riolo, D. T. Robinson, M. North, and W. Rand, 'Spatial process and data models: Toward integration of agent-based models and GIS', Journal of Geographical Systems, vol. 7, no. 1, pp. 25–47, 2005.

[7] S.-G. Chen and J.-Y. Wu, 'A geometric interpretation of weighted normal vectors and its improvements', Computer Graphics, Imaging and Vision: New Trends, pp. 422-425, 2005.

[8] S. Chu and B. Cesnik, 'Knowledge representation and retrieval using conceptual graphs and free text document self-organisation techniques', International Journal of Medical Informatics, vol. 62, no. 2-3, pp. 121 – 133, 2001.

[9] D. Demyen and M. Buro, 'Efficient triangulation-based pathfinding', in Proceedings of the Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference (AAAI'06), Boston, Massachusetts, USA, July 16-20, 2006.

[10] S. Donikian, 'Vuems: A virtual urban environment modeling system', in Proceedings of the Conference on Computer Graphics International (CGI'97), Washington, DC, USA, June 23-27, pp. 84–92, 1997.

[11] S. Dragicevic and D. J. Marceau, 'Spatio-temporal interpolation and fuzzy logic for GIS simulation of rural to urban transition', Cartography and Geographic Information Science, vol. 26, no. 14, pp. 125–138, 1999.

[12] N. Farenc, R. Boulic, and D. Thalmann, 'An informed environment dedicated to the simulation of virtual humans in urban context', in Computer Graphics Forum, P. Brunet and R. Scopigno, eds, vol. 18, no. 3, pp. 309–318, 1999.

[13] N. Farenc, S. Raupp Musse, E. Schweiss, M. Kallmann, O. Aune, R. Boulic, and D. Thalmann, 'One step towards virtual human management for urban environment simulation', in Workshop on intelligent user interfaces, 1998.

[14] R. Geraerts and M. H. Overmars, 'The corridor map method: a general framework for real-time high-quality path planning', Computer Animation Virtual Worlds, vol. 18, no. 2, pp. 107–119, 2007.

[15] J. Gong and L. Hui, 'Virtual geographical environments: Concept, design, and applications', in International symposium on Digital Earth, Beijing, China, November 29 -December 2, pp. 369–375, 1999.

[16] H. Haddad and B. Moulin, 'Using cognitive archetypes and conceptual graphs to model dynamic phenomena in spatial environments', in LNAI, ICCS 2007, U. Priss, S. Polovina, eds, vol. 4604, pp. 69–82, 2007.

[17] D. Harabor and A. Botea, 'Hierarchical path planning for multi-size agents in heterogeneous environments', in Proceedings of the International Conference on Automated Planning and Scheduling (ICAPS'08), Sydney, Australia, September 14-18, 2008.

[18] S. J. Harper, J. D. Westervelt, and A.-M. Trame, 'Integrating geographic information systems and agent-based modeling techniques for simulating social and ecological processes', Oxford University Press, pp. 105–124, 2002.

[19] S. Hensman and J. Dunnion, 'Using linguistic resources to construct conceptual graph representation of texts', in Text, speech and dialogue, Springer Berlin/Heidelberg, pp. 81–88, 2004.

[20] M. Kallmann, H. Bieri, and D. Thalmann, 'Fully dynamic constrained delaunay triangulations', Geometric Modelling for Scientic Visualization, 2003.

[21] A. Karalopoulos, M. Kokla, and M. Kavouras, 'Geographic knowledge representation using conceptual graphs', in Proceedings of the 7th Agile conference on Geographic Information Science, Heraklion, Greece, April 29-1st May, 2004.

[22] A. Karalopoulos, M. Kokla, and M. Kavouras, 'GeoSpatial Semantics', Lecture Notes in Computer Science, pp. 1–14, 2005.

[23] F. Lamarche and S. Donikian, 'Crowds of virtual humans: a new approach for real time navigation in complex and structured environments', Computer Graphics Forum, Eurographics'04, 2004.

[24] J. Lengyel, M. Reichert, B. R. Donald, and D. P. Greenberg, 'Real-time robot motion planning using rasterizing computer graphics hardware', in SIGGRAPH '90: Proceedings of the 17th annual conference on Computer graphics and interactive techniques, New York, NY, USA, pp. 327–335, 1990.

[25] D. J. Marceau, L. Guindon, M. Bruel, and C. Marois, 'Building temporal topology in a GIS database to study the land-use changes in a rural-urban environment', The Professional Geographer, vol. 53, no. 4, pp. 546–558, 2001.

[26] J.-E. Marvie, J. Perret, and K. Bouatouch, 'Remote interactive walkthrough of city models', in Proceedings of the 11th Pacific Conference on Computer Graphics and Applications, pp. 389–393, 2003.

[27] M. Mekni, N. Sahli, and B. Moulin, 'A geosimulation approach involving spatially-aware agents a case study on the identification of risky areas for trains', in Agent-Directed Simulation Symposium, pp. 39–44, 2008.

[28] M. Mekni, and B. Moulin, 'Hierarchical path planning for multi-agent systems situated in informed virtual geographic environments', in Second International Conference on Information, Process, and Knowledge Management Symposium, pp. 48–55, 2010.

[29] B. Moulin, W. Chaker, J. Perron, P. Pelletier, H. Jimmy, and E. Gbei, 'Mags project: Multi-agent geosimulation and crowd simulation', Lecture notes in Computer Science, vol. 2825, pp. 151–168, 2003.

[30] R. Najlis and M. J. North, 'Repast for GIS', in Proceedings of Agent 2004: Social Dynamics: Interaction, Reflexivity and Emergence, University of Chicago and Argonne National Laboratory, IL, USA, 2004.

[31] J. Nilsson, 'Principles of artificial intelligence', Springer-Verlag, 1982.

[32] 'Open Source Geospatial Foundation (OSGEO)', GDAL-OGR: Geospatial Data Abstraction Library/Simple Features Library Software, Aug. 2008. http://www.gdal.org.

[33] S. Paris, S. Donikian, and N. Bonvalet, 'Environmental abstraction and path planning techniques for realistic crowd simulation', Computer Animation and Virtual Worlds, vol. 17, pp. 325–335, 2006.

[34] D. Parker, 'GIS, Spatial Analysis and Modelling', Redlands, ESRI Press, pp. 2–20, 2004.

[35] W. Shao and D. Terzopoulos, 'Environmental modeling for autonomous virtual pedestrians', Digital Human Modeling for Design and Engineering Symposium, 2005.

[36] J. F. Sowa, 'Conceptual structures: information processing in mind and machine', Addison-Wesley Longman Publishing Co., 1984.

[37] J. F. Sowa, 'Knowledge Representation: Logical, philosophical, and computational foundations', Course Technology, 1999.

[38] N. Sturtevant and M. Buro, 'Improving collaborative pathfinding using map abstraction', AAAI, 2006.

[39] F. Tecchia, C. Loscos, and Y. Chrysanthou, 'Visualizing crowds in real-time', Computer Graphics Forum, vol. 21, no. 4, pp. 753–765, 2002.

[40] G. Thomas and S. Donikian, 'Virtual humans animation in informed urban environments', Proceedings of Computer Animation, pp. 112–119, 2000.

[41] P. Torrens and I. Benenson, 'Geographic automata systems', Journal of Geographic Information Systems, vol. 19, no. 4, pp. 385–412, 2005.

[42] J. Zhu, J. Gong, H. Lin, W. Li, J. Zhang, and X. Wu, 'Spatial analysis services in virtual geographic environment based on grid technologies', MIPPR 2005: Geospatial Information, Data Mining, and Applications, vol. 6045, no. 1, pp. 604–615, 2005.

CHAPTER 6

Empirical Calibration of Spatially Explicit Agent-Based Models

S. Heckbert[1,*] and I. Bishop[2]

[1]CSIRO Ecosystem Sciences, Tropical Ecosystems Research Centre, Darwin, Australia and [2]Department of Geomatics, University of Melbourne, Australia.

Abstract: Spatially explicit agent-based models integrate human and environmental systems and reveal patterns that arise from a multitude of individual decisions. Properly defining micro-scale behaviours therefore has large importance for macro outcomes, bringing model validation and calibration issues into focus as a challenge for the research community. This paper describes techniques to empirically calibrate representations of decision making processes in agent-based models. Examples are reviewed using survey data, participatory approaches with geovisualisation and experimental economics. Novel approaches are presented, including experimental economics directly integrated with a spatially explicit agent-based model to reveal trading behaviours in markets. The experimental economics calibration tool directly integrated with an agent-based model reduces the need for interpretation in subsequent use of participant data to re-calibrate artificial agents.

AGENT-BASED MODELLING OF COMPLEX SYSTEMS

Geosimulation broadly refers to computational modelling of spatial processes, where environmental and social-economic systems are connected as an integrated system. The system is simulated through time and over geographical space to explore the outcomes of model assumptions. Spatial decision making can be represented in geosimulation using agent-based modelling (ABM). These models allow for macro level spatial patterns to emerge from micro level behaviours of autonomous agents, in order to quantitatively explain the underlying processes that contribute to the system outcomes. However, defensibly defining and calibrating these micro level behaviours remains an outstanding challenge. This paper describes current techniques to empirically calibrate representations of decision making in agent-based models. A variety of approaches are reviewed as case studies. The first case study presents a method to estimate agent preference functions from discrete choice surveys. The second case study presents methods to elicit behaviours from participants through context-rich geovisualisation media. The final case study uses experimental economics to calibrate agent behaviour. The use of an experimental economics platform directly integrated into an ABM is presented as a novel technique to derive empirically calibrated behaviours of human decision makers. Integration of experiments and ABMs allows for a tight definition of the decision making situation and reduces subjective interpretation that arises from model and experimental design inconsistencies.

In spatially-explicit ABMs, emergence of higher-order patterns arises from the multitude of underlying interactions of agents. For a useful definition, ABM is the computational study of systems of interacting autonomous entities, each with dynamic behaviour and heterogeneous characteristics [1]. The ‘agents’ interact with each other and their spatial environment in the case of geosimulation, resulting in emergent outcomes. Interactions can be direct such as communication and physical interaction by mobile agents in space, or indirect *via* feedbacks from aggregate outcomes such as environmental feedbacks from the spatial landscape. The dynamic behaviour of heterogeneous agents is represented by decision making functions, and takes into account the different characteristics of each agent.

Increasingly, researchers are using multiple methods to calibrate ABMs. These include surveys, semi-structured interviews, existing data sources such as GIS and census data, direct participant observation, role playing games, and laboratory experiments. In addition to the normal survey and observational tools primary data collection can be supported by location-based survey techniques [2] and participatory modelling with geovisualization tools [3]. Existing data sets at increasingly finer resolutions are becoming available for use in geosimulation including satellite imagery, fine resolution GIS databases, census and marketing data increasingly at the individual level. From the data gained through these sources, statistical relationships can be derived, and / or decision making rules constructed from interpretations of the data. Primary data often

***Address correspondence to Scott Heckbert:** Institute for Sustainable Solutions, Portland State University, Portland, OR, USA. 97201: Email: scott.heckbert@gmail.com

Danielle J. Marceau and Itzhak Benenson (Eds)

collection is performed when there is a gap in available data from which to generate the population of agents. The following three case studies present methods of collecting data for agent calibration: estimating preference functions from discrete choice surveys, eliciting behaviours from participants through context-rich media, and using experimental economics to inform agent behaviour in tightly defined decision making situations.

AGENT BEHAVIOURS DRAWN FROM SURVEYS

This section reviews a case study using survey data in calibrating a spatially explicit ABM. Surveys and other data sets can gather information to derive individual or household behavioural models based on microeconomic theory and to generate statistical descriptions of the attributes of agents. Survey data can identify types of agents based on cluster analysis, and provide information on the distributions of characteristics, beliefs and preferences within a group. Surveys are good for sampling and extrapolating to the population level [4].

Survey data are frequently collected for use in stated preference discrete choice experiments. These surveys ask participants to make selections between sets of choices with the aim of isolating the preference for certain attributes of the choice. A well designed survey can produce data for analysis using regression to infer a causal link between an attribute of the landscape and the associated behaviour.

An agent population can be constructed from data collected by surveys. A typology of different representations of agents can be identified from survey data, as outlined in [5]. In re-assigning behaviours to the various types of agents, a straightforward way to parameterize models is presented in [6] using a consistent sampling frame in terms of scale to select observation units for both biophysical measurements and socioeconomic surveys. Parameters are then extrapolated based on estimated probability functions and assigned to the model agents and landscape units (such as cells). The resulting formulation of landscape units and agent population are statistically consistent with empirical data [6].

It is common to use surveys in combination with other approaches such as geovisualization choice experiments (discussed here in the second case study) and similar stated and revealed preference techniques [7], [8]. For example, in developing agent-based models for recreational behaviour and management of sensitive tourist environments, models are frequently calibrated using a combination of observed visitor behaviour and surveys of visitor beliefs, values and preferences [9 - 11]. In urban geosimulation, Brown and Robinson [12] use econometric estimates from survey data to design agent preference functions for residential selection and then predict land use change. These techniques use stated preferences and reported behaviours to derive behavioural functions for agents.

This paper's first case study presents a novel methodology for calibrating agent behaviours within models using preference measurement techniques from economics. The novelty of this approach is that it is based in rigorous econometric methods and can be replicable across many ABMs where agents are faced with making a selection from a choice set. Using this technique allows discrete choice surveys and subsequent econometric analysis to define the functional form of equations and parameter weightings.

The model described here is presented in full in Heckbert *et al.* [13], and represents the development of forestry roads and the indirect effect of roads on wildlife through hunting. We provide an example of using survey data for the empirical estimation of the preference functions and the calibration of the decision-making process of agents. Estimating preference functions is a common and well established methodology in econometric analysis. Behaviours can be elicited through stated preference survey techniques, where respondents are asked to make a series of tradeoffs in discrete choice surveys. Econometric analysis, specifically random utility modelling as summarised in [14] and [15] determines the influence of attributes of the choice set on decision making.

In this first case study, simulated agents represent hunters who select a hunting site. The decision maker selects the site with the highest expected utility based on three parameters: travel cost (with negative marginal utility), the abundance of game quarry (with positive marginal utility), and interactions with other hunter agents (negative marginal utility). The parameters of the utility function are estimated based on the variety of existing survey data from discrete choice studies [16 - 19].

The simulation model is implemented with Net Logo 4.0.4, and is available online [13]. The model includes mobile hunter agents who operate over a forest landscape represented by a grid of cells, populated by the game populations.

The model focuses on the evolution of the road network over the cellular landscape. At each time step, hunter agents $i = 1...n$ evaluate the expected utility U_{ij} that would be gained by attending a given cell j.

$$U_{ij} = \sum_{k=1}^{K} \beta_i^k X_j^k$$

Each cell contains a vector of attributes X_j^k for the attributes k of travel cost, game availability, and hunter congestion. The utility function is a parameterized linear equation which takes on the functional form estimated by random utility models, with preference parameters β_i^k representing the marginal utility derived by hunter i, from attribute k, at hunting site j. The coefficients β_i^k are calculated using random utility models using survey-based discrete choice studies [16 - 19] of Canadian hunters' preferences. Preference parameters β_i^k are initialized across the agent population as a random number with mean β_i^k and standard deviation σ_β. Parameters β^k and σ_β for the utility function are drawn from [16 - 19]. Agent heterogeneity is introduced through assignment of individual preference parameters similar to the method described in [6]. The resulting initialized population of hunting agents have heterogeneous preference structures.

The influence of interacting with other hunters is introduced though a variable tracking hunter congestion, which is simply the number of hunters in a cell's eight neighbouring cells. Game population N_{jt} [animals/km^2] in each cell is developing according to a discrete logistic model,

$$N_t = N_{(t-1)} + r\left(\frac{1 - N_{(t-1)}}{K}\right) N_{(t-1)} - H_t$$

where r is the intrinsic population growth rate [0.2/iteration], K is the population carrying capacity [4.4 a/km^2], and H_t is the number of animals harvested in time t. H is calculated by the number of hunters in the cell, multiplied by the hunting success rate [0.3 kills / hunter].

Overall Travel cost TC_j to attend the hunting site is a weighted average of the costs of travelling through forest, forestry roads, and highways, and is given by:

$$TC_j = \alpha^{DF} * DF_j + \alpha^{DR} * DR_j + \alpha^{DH} * DH_j$$

where α are unique cost parameters and relate to the distance through the forest to the nearest road DF_j, the distance along forestry roads to the highway DR_j, and the distance from the highway intersection to the city DH_j.

The agent selects the hunting site expected to yield the highest utility, and multiple hunters can attend the same site. Hunters return to the origin at the end of each time step. Fig. **1** shows outcomes for the simplified case where the hunter origin is located at the centre of the map and there are no roads. In this case, hunters select surrounding cells for hunting. The simulation is run for 250 iterations of 80 hunters, in an 80 x 80 grid at a resolution of 1 km^2. The darkest cells contain game populations at carrying capacity while the white cells have experienced a local extirpation with all animals removed. The different patterns of light and dark cells visible in Fig. **1** panels are created by parameterizing the hunter agent utility function from stated preference studies listed above.

Simulation runs depicted in Fig. **1** visualize the impact the agent population has over time using different combinations of preference parameter weightings. Each scenario is identical except for the parameter weightings applied to preferences for travel cost and game population, listed in Table **1**. Fig. **1**, panels 2-4 are parameterised from results reported in [18], [16] and [17] respectively. These papers report estimated preference weightings based on random utility modelling using discrete choice surveys. The reported preference parameters are inputted as the β_i^k parameters in agents' utility functions.

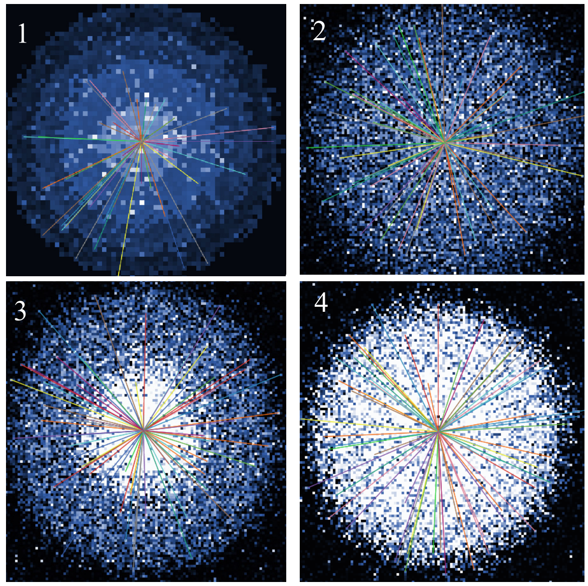


Figure 1: Spatial configurations resulting from parameterizing hunter agents from discrete choice surveys (2-4) and one combination revealing a sustainable resource use pattern (1). Darker cells have more game, white cells represent extirpation through hunting, and rays depict routes traversed by hunters disseminating from the origin.

Table 1: Preference weightings used in parameterising agents' utility functions, for three attributes of a hunting site; travel cost, game population, and hunter congestion, as measured from three stated preference studies.

Preference Study		Bottan 1999 [16]	Adamowicz *et al.* 2004 [18]	Haener *et al.* 2001 [19]
Travel cost / cell		-0.008	-0.004	-0.0125
Wildlife Pop'n [animal / km2]	0	-0.25	-0.49	-1.41
	1	-0.20	0.15	-1.41
	2	-0.15	0.15	0.22
	3	0.50	0.85	0.22
	4	0.50	0.85	0.22

In some cases, the agents apply an even spread of hunting pressure that does not push game populations within cells to extirpation. In other cases, with the same number of agents, game populations are uniformly extirpated simply due to a small difference in preferences. Fig. **2** presents simulated outcomes for the number of cells extirpated and the utility (indexed to initial value for comparability) derived by agents over time. The graphical depictions in Fig. **1** display the game population at time step 250, and in Fig. **2** the trajectory through time can be seen. The higher number of extirpations in the parameterisation based on [17] is due to a large (negative) travel cost parameter and weak preference for abundant game. These agents apply hunting pressure close to the origin and do not avoid cells whose game population is already diminished. As a result a greater number of cells are extirpated at a near constant rate, and utility derived by agents decreases as cells are extirpated. In comparison, the parameterisation based on [16] incurs approximately half as many extirpations, at a rate that begins to level out over time. This is caused by a smaller (negative) travel cost parameter and larger preference for cells with abundant game. As a result many cells near the origin are extirpated, but overall the agent population travels further afield to find cells with abundant game. The effect is to spread the overall hunting pressure across a larger area, enabling cells' game populations to regenerate. This also halts the diminishing levels of derived utility. The parameterisation based on [18], agents are willing to travel further and have a high preference for abundant game, effectively spreading out the overall hunting pressure, with few extirpations and nearly achieving non-declining utility over time.

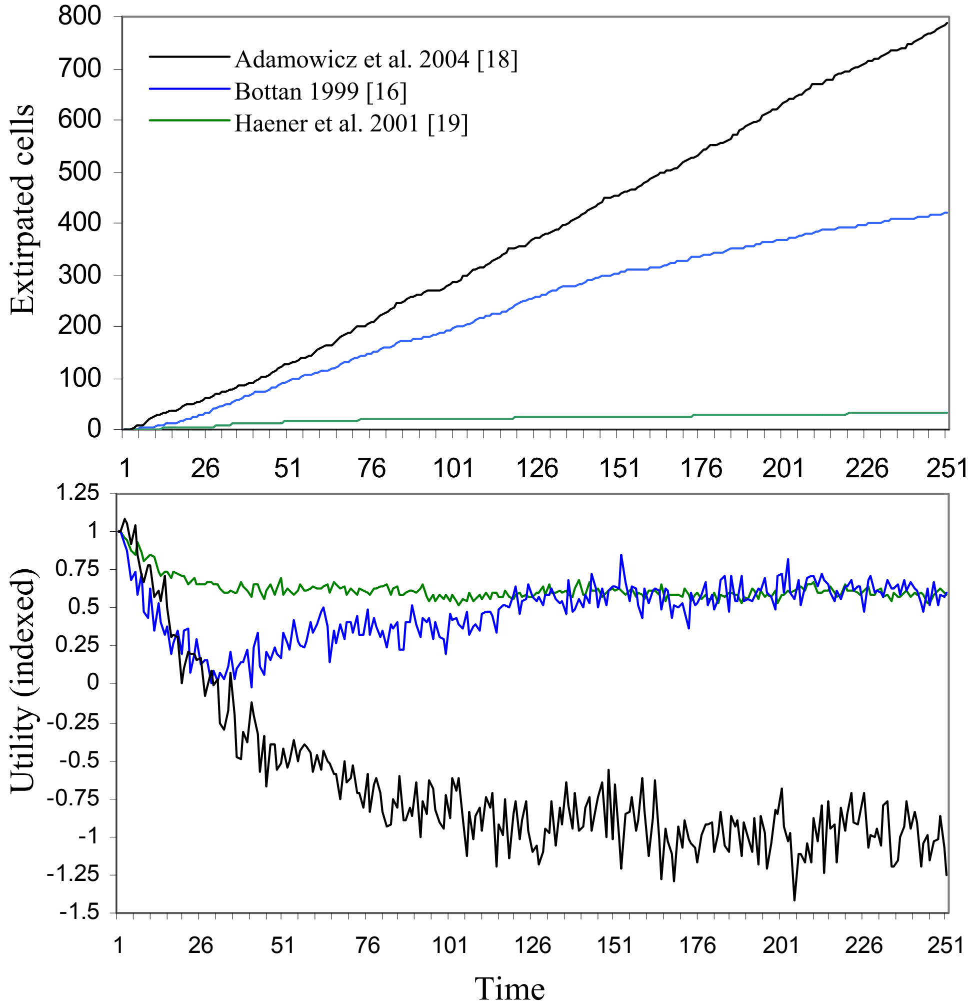


Figure 2: Number of extirpated cells and indexed utility derived by agents over time for three parameterisations of agent utility function.

One parameter configuration, depicted in Fig. **1** panel 1, reveals a resonating pattern which distributes hunting pressure across the landscape, and as a result extinctions do not occur. In this simulation, a series of concentric circles visually emanates from the hunter origin as the simulation progresses, dispersing as they move further away from the centre. This pattern is caused by the trade off between distance (negative utility) and game abundance (positive utility); when cells near to the origin have game numbers reduced, hunters travel further afield reducing game numbers there. Over time, populations recover and hunters return to closer sites. En mass, this has the effect to generate a spatial pattern of dispersing horizons of hunting pressure much like a pulsar, with a frequency determined by the relative weighting of preference parameters for travel cost and game abundance. This is an example of a spatial emergent property of a sustainable system of renewable resource harvest that self regulates and avoids crashing. This pattern yields an even flow of non-declining landscape utility, whereas all other calibration settings show declines in utility as sites are extirpated.

This example of agents operating with a simplified spatial landscape is now combined with a modelled road network to examine the relationship between hunting and road building/decommissioning. Interactions between forestry roads and hunting are well documented, with studies finding hunting pressure being related to accessibility of a hunting site. Industries which create access within forested areas may indirectly affect wildlife populations as hunters use roads to travel to hunting sites which were previously more difficult to reach.

A modelled road network serves to alter the relative accessibility of cells *via* the travel cost calculation, and disturbs the uniformity of space assumed in simulations shown in Fig. **1** The road network is represented as a series of linked nodes. Fig. **3** shows the case where a road network alters the travel cost landscape surface.

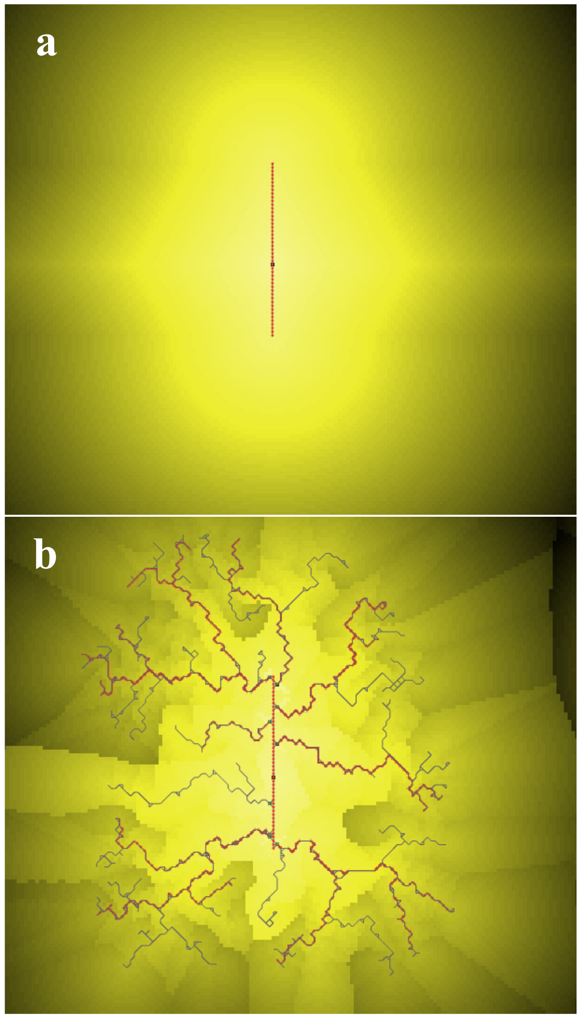


Figure 3: Uniform travel cost surface (a) with darker cells having higher travel cost, and altered surface (b) where a network of roads changes travel cost from the origin.

Roads are generated adjacent to existing road nodes and removed over time (decommissioned) using a process fully described in [13]. This process continues over the model time steps to generate an evolving network of roads that spreads across the landscape and serve to alter the travel cost calculation to cells. Hence when a new road is built, the surrounding cells return a higher expected utility calculation given their increased accessibility and hence lesser negative marginal utility for distance.

The outcomes of the combined model of road development, agent behaviour, and game population growth and decline are presented in Fig. **4** (c) depicts circular road nodes (red for highway and yellow for forestry roads) with hunter agents overtop of cells displaying game population levels in shades of blue, similarly to Fig. **1**.

Using the utility seeking agents and the evolving road networks, simulations were run to test the effect of different decommissioning ages of forestry roads. In the simulation results depicted in Fig. **4** (a, b), two decommissioning ages are compared (2 and 10 year decommissioning). Overall extirpations are 4448 and 3870, respectively. In the case where roads are decommissioned later, the decreased numbers of extirpations is caused by a more even spread of hunting pressure across the landscape.

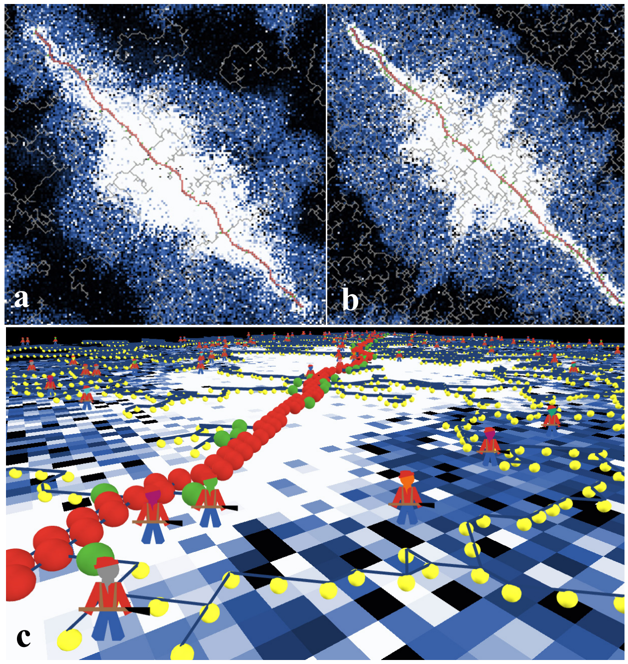


Figure 4: Spatial configurations in game populations (a, b) resulting from hunting under different forestry road decommissioning policies, and model view (c) of agents and road network on spatial landscape.

This case study highlights that attention to empirical calibration is important to be able to understand patterns that arise at a higher scale. Furthermore, information from surveys alone is not sufficient to understand adaptive behaviour of human decision makers. Data collected in surveys capture a snapshot in time and within a certain context of agent characteristics and behaviour, and surveys do not easily reveal motivations and strategies. To learn directly the motivations behind people's behaviour, semi-structured interviews can be used to explore drivers behind dynamic decision making. Interviews, focus groups, field interviews, and participant observation from anthropological techniques [20] have been used to guide definitions of agent decision making.

PARTICIPATORY TECHNIQUES FOR GATHERING CONTEXT-RICH BEHAVIOURS

Whereas the previous section described data collection techniques based in stated preferences and behaviours, in this section we describe techniques that begin to elicit revealed preferences from participants. This is important because what respondents report in surveys and interviews may differ from actual behaviours. Behaviours that depart from otherwise rational decision making are often couched in the context of a specific decision making situation. Here we describe methods for purposefully introducing context in data gathering activities with participants. The goal of these activities is to identify factors influencing decision making by providing research participants with a controlled setting that represents elements of the real-world situation.

Participatory modelling linked to ABMs has been done extensively using the companion modelling technique [21]. This allows for stakeholders to explore system interactions in role-playing games laden with context. Companion modelling develops collective learning in stakeholders and supports decision making processes by representing different perceptions of a complex situation, and by collectively exploring possible futures [21].

Real-world stakeholders are the participants and they play their roles while information is gathered to be used in developing the associated simulation models [22-24]. The information collected on stakeholder behaviour is evaluated by the stakeholders, including post-game interviews and cross-checks, then transformed into rule-based agents in models. This process can offer system-level awareness building and an opportunity to observe agent-agent interactions, but is limited by issues of objective knowledge of stakeholders, subjective interpretation of behaviours by researchers, and has been criticized for playing a role in re-enforcing power relationships in the stakeholder groups of case studies, often in developing countries [25].

Observation of behaviour in the landscape is referred to as the experiential or phenomenological approach to landscape assessment [26], [27]. When combined with consumer choices this behaviour can also be analysed using economic theory. Classic examples are use of hedonic price analysis to determine the value of views or quietness in housing [28], or trip cost analysis to estimate the value of a recreational site or a landscape [29]. However, these are based on accumulated, emergent, behaviour rather than individual behaviour.

At the individual level, global positioning systems (GPS) create the opportunity to track people's behaviour and also to prompt them with location specific questions when they arrive at particular places, such as a path intersection. The use of location sensitive questioning is not yet common. However there are developments based on similar technologies for delivering location specific information [30], path tracking [31], and mobile computer based questioning [32]. Loiterton and Bishop [3] designed a system based on a mobile computing device with GPS which tracked participants as they roam freely within the environment of interest. At specific time intervals, as the visitor comes to an intersection, a set of generic questions pop up, allowing the user to indicate the physical and perceived properties of the environment that are affecting their path decisions as they progress. Each time these questions are answered, the participant also gives an indication of their current feelings (*i.e.* levels of boredom, fatigue, hunger, etc) using a simple scroll-bar design. The answers, when combined with the tracked behaviour, provide data for geosimulation calibration.

A further option in the development of experimental platforms lies in the option of using visualization technology to provide subjects with a realistic representation of the environment in which decisions are to be made. This potential increases the level of association of the subject with the environment and hence the ability to observe responses to contextual information. It also potentially increases the degree to which options and hence future conditions, under particular choices, are effectively communicated to the stakeholder.

Simulated environments have been used in environmental psychology research, especially in landscape assessment [33], extended to willingness to pay [9], and choice experiments studies. Bishop *et al.* [34] have argued for linkage of the multi-user human-agent modelling environment with the realistic collaborative virtual environment (with linked environmental process models) to provide a complete virtual decision environment. In such a combined system of modelling and visualization of landscapes, information about the environment is derived from the senses (primarily sight) and also on request from neighbours, advisors, the internet and other sources.

Figure 5: Virtual intersection – original (top) and lake views (bottom) from [3].

Questions naturally arise about the validity of the sensory environment provided in this way. Since images were first used in landscape assessment studies, researchers have sought to establish their validity within the specific research context. A review of application and validity in the context of outdoor recreation is provided by [35]. Rohrmann and Bishop [36] and Bishop and Rohrmann [37] tested the validity of animation of urban walks for assessment of affective and cognitive responses. They found a good reliability of relative changes in responses but had more difficulty getting accurate absolute measures. Jallouli and Moreau [38] compared personal responses to wind energy turbines in a real-time interactive virtual reality environment with responses in the real world. Loiterton and Bishop [3] used animations to support experiments in path choice within a public urban garden environment, as shown in Fig. **5**. The animations were modified to test the specific influence of variables such as slope, enclosure and views to water on visitor behaviour.

Using computer generated visualizations, the visible features of a landscape can be controlled while making calculated alterations and then monitoring the responses of human subjects. In this way, we can gain a greater understanding of human perception and behaviour in both the natural and built environment. The value of virtual environments in understanding human perception and response has been established in several contexts including navigation and psychotherapy studies [39], [40].

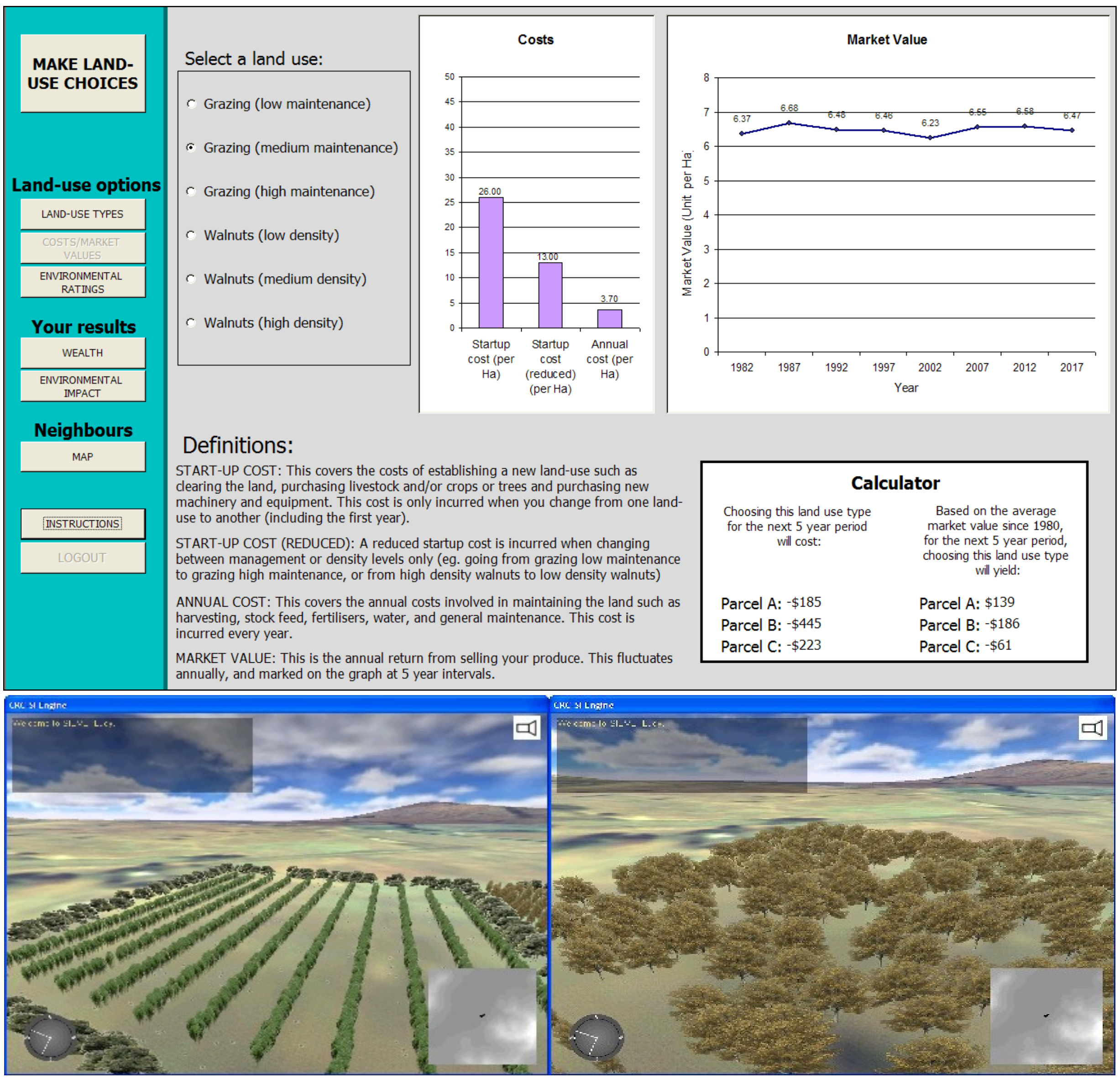


Figure 6: The experimental land use selection interface was used in conjunction with an immersive display from [41].

In the land-use change context, a preliminary study using realistic visualization is described by Kennedy and Bishop [41]. They tested land-use decision making in response to the presence or absence of neighbours (who are also making decisions) and the presence or absence of visual representation of the landscape emerging from the personal and neighbours decision making, as shown in Fig. **6**. They were however unable to establish any clear influence of the visualization on the decision-making or its validity although subjects reported that the visualization assisted their decision making.

EXPERIMENTAL ECONOMICS FOR CALIBRATING AGENT DECISION MAKING

Whereas landscape visualization and companion modelling provide a context-rich medium, laboratory experiments can be used to strip away contextual influences and test human behaviour in very controlled decision making settings. Laboratory experiments and agent-based modelling can be used together to derive model functions that represent human behaviours with parameters calibrated from experimental data. Experiments can also be conducted to support or negate a proposed agent behavioural algorithm when modellers are choosing between several possible representations.

Experimental economics uses participants in laboratory settings to study economic behaviour and test theories of decision making. Participants are paid according to outcomes, attempting to re-create real-world incentives to understand how markets operate. Reeson and Nolles [42] describe the application of experimental economics to natural resource management, outlining how experiments can reveal the intricacies of human behaviour in markets, showing how apparently minor details of market design can lead to significant differences in outcomes. Experiments test how people respond in non-market situations, such as when faces with a trade-off between self interest and the collective good (as is common in many environmental dilemmas) [42].

It is common to build models of human decision making with the results of experiments in mind, but the use of agent-based models and experiments in combination is very limited. First examples are [43], and in the non-spatial field of agent-based computational economics [44], which presents agent modelling and experiments informing one another (but not integrated) as a means of isolating the sources of aggregate phenomena in trading markets. This concept is applied to land-use change in Evans *et al.* [43] who model reforestation and agricultural land-use decisions, and present an example of how experiments can be used to test very specific but important assumptions. In [43] the authors used a mixed methods approach to calibration. First remote sensing data were used to estimate agent preference parameters by regression; then surveys were used to collect social and demographic information. However, in simulation results only a weak relationship between income and reforestation was found using this approach, and other factors such as learning, information, knowledge, risk aversion, and influence of social networks were hypothesized to play a role, but not able to be captured in surveys [45]. Computer-based laboratory experiments were designed to further test hypotheses about decision making and resulting behaviours. The experiments in Evans *et al.* [43] assessed how people make allocation decisions between agricultural use and reforestation. Subjects were allocated spatial areas to one of two land uses, receiving revenue according to an increasing price for one, and a decreasing price for the other. Experiments found considerable variance in behaviours of allocating land to each of the two uses, and where a ‘rational’ decision maker would have changed land use, the majority of experiment participants took many rounds to complete the reallocation, and some persisted in allocating to the disadvantaged option [45]. The overall outcome of participant decision making was to generate overall reforestation patterns with more ‘edge’ compared to a model populated with agents who make fully rational choices.

By calibrating against experimental results, fully informed and rational decision makers can be replaced with alternative representations grounded in empirical data. This technique in presented in the third case study of this paper, an application calibrating agent behaviour using an experimental economics platform integrated within the running ABM. Heckbert [46] [47] use ABM to address issues of water quality and coastal land use change in the Great Barrier Reef region of Australia. The model is constructed to test policy options for introducing a cap-and-trade system for allocating and trading permits for agricultural fertilisers. Within the simulation model, depicted in Fig. **7**, agricultural agents participate in a cap-and-trade emissions trading system for ‘water quality’ permits.

The model is constructed in Netlogo 4.0.3 and uses GIS inputs of agricultural landholdings for the Douglas Shire, Queensland to generate an agent population. Agents are the decision making unit associated with 164 properties as identified through cadastral data inputted into the model. Agents perform a series of scheduled operations for each time step, representing one year. The main operations pertain to production decisions involving choosing application rates for fertiliser, and trading behaviour in a market.

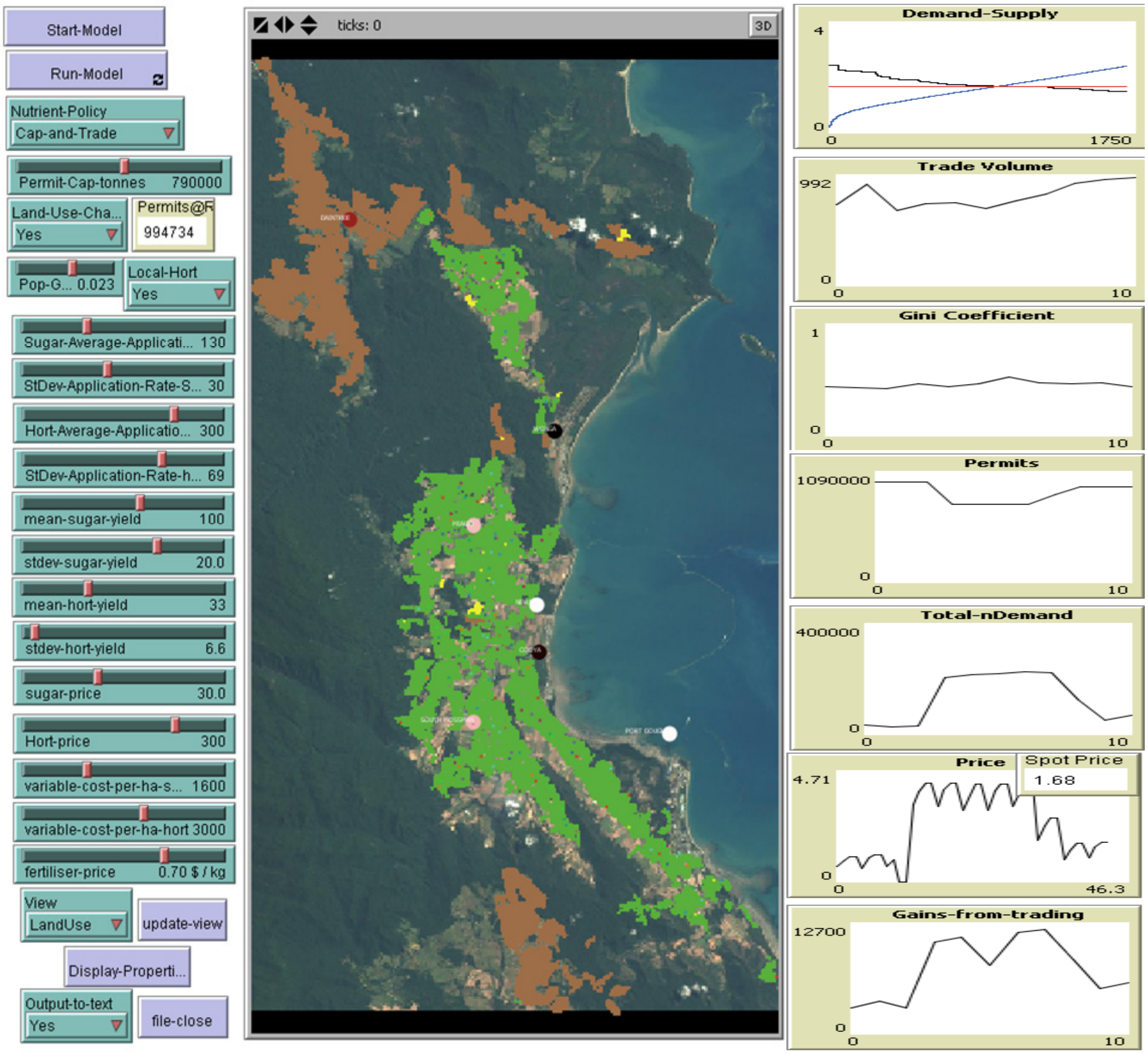


Figure 7: Interface for geosimulation of water quality cap-and-trade model testing the effect of land use conversion on market dynamics.

Market designs have been tested in the laboratory and in ABMs. An appropriate market design is described in Hailu and Thoyer [48] using a multiple-unit market design in an agent-based trading model. Unlike single-unit auctions, multi-unit auctions allow traders to submit offers with quantity and price schedules rather than single quantity-price bundles [49]. The population of bidders has private values reflecting different production (demand) and cost (supply) structures, indicating the amount they would be willing to supply at different prices [50].

In order to calculate variables for production and trading decisions, agents perform a series of functions calculating tradeoffs between input use and the effect on agricultural output. For specification of parameters for the following equations refer to [46].

Fertiliser application rates $N_{j,t}$ [kg/ha] are initialised at model setup for each agent $j = 1...164$ according to a normal distribution,

$$N_{j,t} \sim N(RR, SD)$$

where RR is the mean recommended application rate [kg/ha], and SD is the standard deviation which determines the level of heterogeneity in fertiliser application rates across the agent population. Crop yield is calculated as:

$$O_{j,t} = \kappa_j * \left(1 - \delta * e^{-\beta * N_{j,t}}\right) - \gamma * N_{j,t}$$

where O_j is the crop yield [t/ha] realised using fertiliser application rate from above, and γ, δ, β are static yield parameters. The parameter κ_j is the property-specific yield parameter, varying among properties as

$$\kappa_j \sim N(PP, PSD)$$

where PP is mean productivity and PSD is the productivity standard deviation, normally distributed across the agent population.

Agents derive revenues and incur costs from growing crops and from buying and/or selling fertiliser permits [kg] from one another, with total profits expressed as:

$$\pi_{j,t} = FR_{j,t} - FC_{j,t} + \left(O_{j,t} * A_{j,t} * CP\right)$$

$$-\left(\varepsilon * V + V * A_{j,t} + N_{j,t} * A_{j,t} * FP\right)$$

where $FR_{j,t}$ and $FC_{j,t}$ are revenues and costs [$] respectively, incurred from selling and purchasing fertiliser permits. Production revenues include the area $A_{j,t}$ [ha] of the property and CP is the commodity price [$/t] paid for yield. Variable costs for each land use are V [$/ha], and ε [ha] is the minimum farm size observed in our data set, used to infer a measure of fixed costs. The commercial cost of fertiliser is FP [$/kg].

In order to trade permits within the market, agents make pricing decisions and choose the amount of fertilizer they would like to buy or sell given their calculated profitability at different application rates. Permits are limited based on the stringency of the cap, and agents have heterogeneous production functions resulting in a population-level distribution of willingness to pay (WTP) and / or accept (WTA) for permits.

Permits $P_{i,t}$ [kg] are assigned to agents based on recommended rates for each land use,

$$P_{j,t} = RR * A_{j,t}$$

From here, demand for additional fertiliser permits $D_{j,t}$ [kg] is the calculated shortfall between the agent's fertiliser application rate and permits available.

$$D_{j,t} = \left(N_{j,t} - P_{j,t}\right) * A_{j,t}$$

At this point agents calculate two price-quantity schedules; one of WTP for demanded permits and one of WTA values for permits the agent might supply to the market. The artificial agent demand schedule is an array list of WTP values populated with the marginal value $MVD_{j,t}$ [$/kg] for additional fertiliser units,

$$MVD_{j,t} = \left(O_{j,t}^{n+1} - O_{j,t}^{n}\right) * CP - FP$$

where $O_{j,t}^{n}$ is yield [t/ha] at fertiliser input level n. $MVD_{j,t}$ is calculated for all values of n within the range of $n = P_{j,t}$ to $n = D_{j,t} + P_{j,t}$, and FP again is the price of fertiliser [$/kg].

In a similar fashion, sellers also calculate a WTA schedule which is again stored as an array list. The WTA is calculated as the marginal cost [$/kg] of supplying additional permits to the market,

$$MCS_{j,t} = \left(O_{j,t}^{n} - O_{j,t}^{n-1}\right) * CP + FP$$

where $O_{j,t}^{n}$ is yield across the range $n = P_{j,t}$ to $n = 0$. The WTA and WTP values are then added to a list, sorted ascendingly for offers to sell (WTA), and sorted descendingly for bids to buy (WTP). The overall market price SP_t is defined at the pair-wise combination where the demand and supply curves intersect. The heterogeneity in WTP and WTA is sufficient to generate market supply and demand curves and thus, determine expected trading prices, trade volume and market efficiency.

Simulations were run for a number of scenarios which explore market and production outcomes at different cap levels for aggregate fertiliser application. Indicators tracked within the model inform the effectiveness of the cap-and-trade scheme under different cap levels. The trading price [$/kg] and trading volume [t] inform the costs faced by traders and the market throughput.

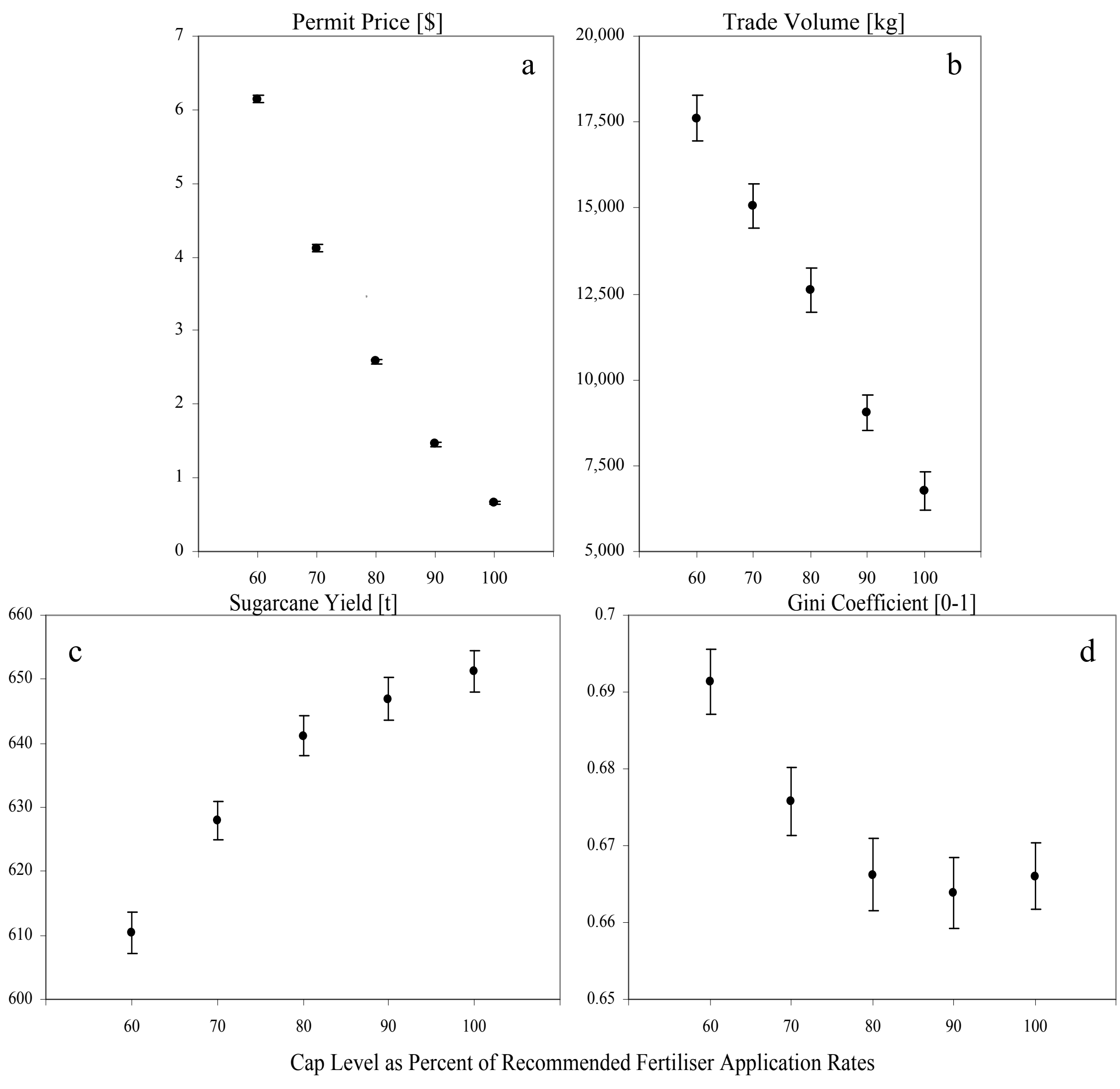


Figure 8: Simulation outcomes for a) trading price, b) market trade volume, c) total sugarcane yield and d) the gini coefficient for farm profits, describing the distribution of profits across the agent population. Values are mean estimates for 100 trading periods and confidence intervals.

Simulations were run testing five cap levels which constrain aggregate fertiliser application, the cap being the total number of limited fertiliser permits available. Results are presented in Fig. **8**, comparing the outcomes for trading price for permits, trade volume, agricultural yield, and the gini coefficient of income distribution. The trading price increases super-linearly with lower (more stringent) cap levels, as might be expected. Trade volume increases linearly as less permits are available, and agents price them higher yet demand more under stricter cap levels. As would be expected, the total aggregate sugarcane yield decreases as the cap levels tighten, with significantly different results between 70% and 80% of agronomic recommended rates. An indicator of income distribution is measured using a gini coefficient. The gini coefficient is a statistic that measures distribution of profit within the population, where 0 represents a completely homogenous population (exactly equal, with all agents receiving the same profits), and values approaching 1 where the distribution of profit is skewed towards only a few individuals (less equal). The gini coefficient is higher with lower cap levels, indicating that profit distribution becomes less equitable with lower caps. Cap levels between 70% and 80% are again significantly different for the gini indicator, revealing that the distribution of profits within the agent population begins to change past this point due to concentration of profits within a smaller number of agents. These results suggest that the cap-and-trade system market would function well up to a given level of cap stringency, but would start to result in significantly different production losses and uneven distribution of income at a cap level around 70% of agronomic recommended rates.

Re-Calibrating Trading Behaviour using Experimental Economics

The results presented in Fig. **8** are based on model assumptions and data, particularly the fashion by which agents make production and trading decisions. However, this representation of agents assumes they are fully informed and optimise their production and trading behaviour, whereas agents programmed with behaviours more akin to those of real humans might result in a different set of WTP and WTA values and different production and trading decisions, which in turn serves to alter price and market efficiency. Depending on market design, experience of traders, access to information and other conditions required of markets to operate efficiently, behaviour in real markets is expected to deviate from optimal behaviours. We understand from experimental evidence that traders often make decisions with all manner of behaviours which deviate from assumptions of fully informed rational agents [42] [49] [50].

Calculating WTP values for price-quantity bundles while trading is a very precise and defined decision, well suited to study within experiments. Experiments were designed and conducted using the same base ABM described in [46] in order to examine economic decisions requiring tradeoffs, in this case the WTP for permits of fertiliser and the price-quantity trading bundles that experiment participants choose. An experimental interface was built and integrated directly with the ABM, and participants login as agents, making agent WTP decisions and choosing price-quantity bundles to trade. Experiments include trading with artificial agents, other human participants, or some combination of both. Treatments test the provision of information and the productivity heterogeneity of the population.

The experimental interface is depicted in Fig. **9**. Participants see information about their landholding including current bank balance, 'productivity' compared to the population average, trading prices, and a summary table that tracks outcomes over time. In some treatments participants are provided with a marginal value table (right side of interface) which informs them of optimal WTP values for different size trading bundles. Lastly, sliders are moved by the participant to select the price and quantity of trade. Data from respondents are collected and stored for analysis. The participants interact with the running ABM. Information is provided to the experiment participant, and they select parameter levels for equations within the artificial agents< trading operations.

Participants were given written instructions and answered a simple multiple choice test before the experiments start. Participants were paid with cash based on their performance, and filled out a post-experiment survey which elicited descriptions of the strategies they used. From a pool of 20 participants, more than 100 replications were conducted over three days, an average replication having five participants, and as many as 20 participants for treatments where only human traders were included. Wording of the instructions and interface is intentionally neutral, such that participants apply 'Inputs' for 'Production' rather than 'fertiliser' for 'sugarcane'.

The experiment consisted of three parts meant to train the participant for performing trades. Firstly participants select input application rates over multiple time steps, and are presented with a history of outcomes as well as the results of their best outcomes so far. When all participants have decided on application rates, the second part of the experiment begins, where a random amount of inputs is 'lost', akin to losing an amount of fertiliser due to an unknown frequency and intensity of rainfall events. Participants are asked to adjust application rates based on this stochastic event, again repeated over several times steps. Lastly, participants are presented with a limited cap on the number of inputs which can be used, and asked to select prices and quantities to trade, first without random loss of inputs which is introduced at the end of the replication. All information and parameters from the interface are recorded in spreadsheets along with the participants' selections for inputs and trade price and quantity. From the output data several statistics are calculated, including the difference between 'optimal' and chosen input levels, the learning rate of participants in finding selected values, the adjustment of inputs in response to risk, and the effect of risk in trading price and quantity.

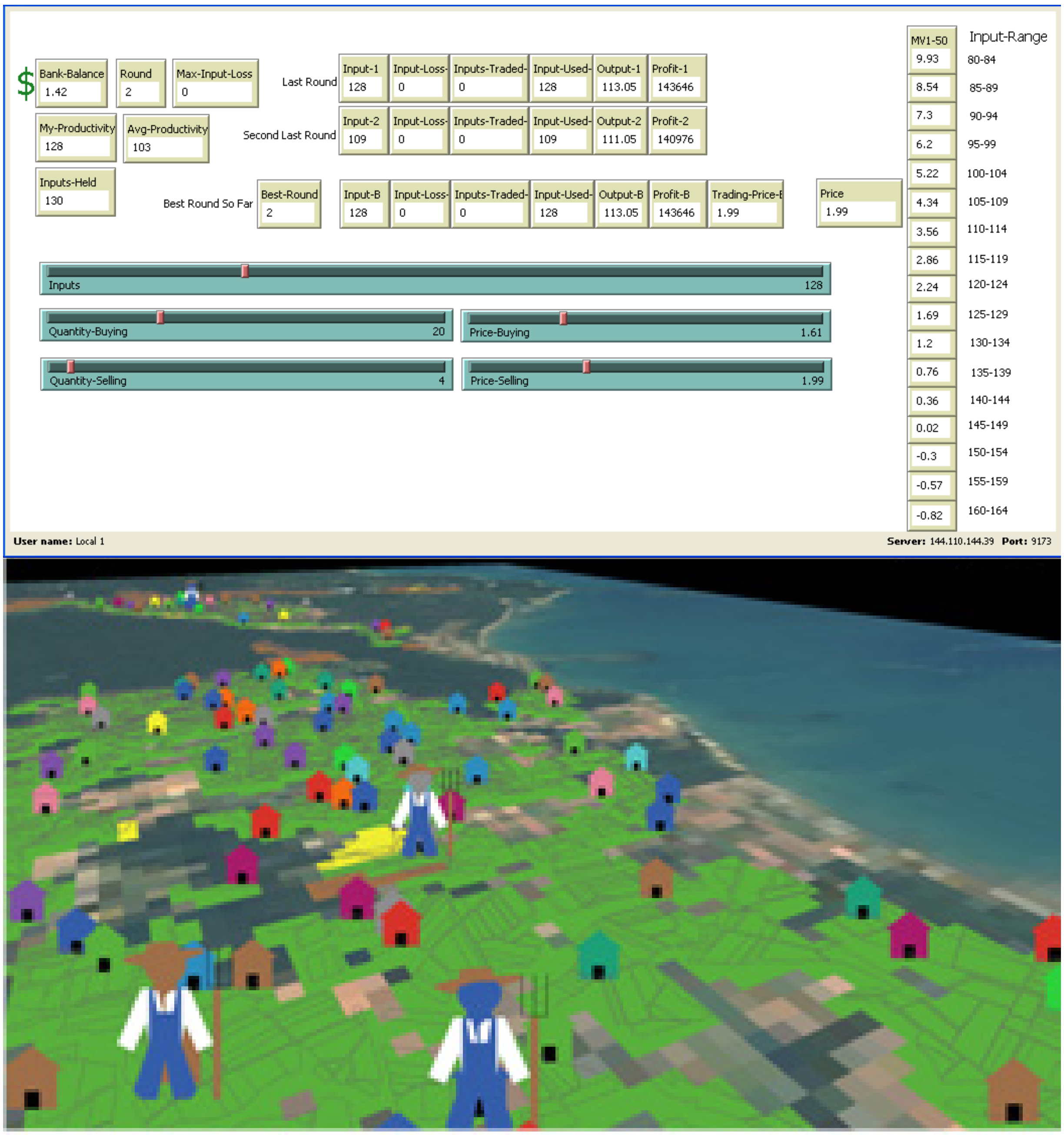


Figure 9: Experiment and spatial interface depicting experiment participants within a running agent–based model. Participants use sliders for selecting price-quantity bundles in a cap-and-trade water quality market.

As it is found in real-world markets, participants included a 'markup' on their prices, as an extra margin above (below) the optimal selling (buying) price. The overall effect on supply and demand curves is to raise the equilibrium price and reduce the market efficiency. As expected, market efficiency is increased with groups of experienced traders. An interesting finding consistent with anecdotal evidence from the real-world is that participants often over-applied inputs in response to risk. When faced with losing inputs determined by a uniformly distributed random number with a known mean, participants often increased application rates by more than the mean, which would be the expected value over repeated time steps. Agriculturalists have shown this same behaviour in selecting fertiliser application rates as a risk buffer, preferring to over-apply at a net cost than risk losing significant production.

Statistics are calculated from participant output, and then re-inserted into the production and trading algorithms to parameterise artificial agents. The re-calibration with markup values serves to raise price and decrease market efficiency. Over-application in response to risk serves to increase overall demand, again raising price. The distribution of revenues as measured by a gini coefficient is shown to be less equitable across the agent population as a result of this increased price and decreased efficiency. The outcomes resulting from re-assigning artificial agents with experimental findings serve to highlight the importance of calibrating models beyond assumptions of fully informed and rational decision making. The increase in price and decrease in market efficacy revealed by the re-calibration bring into question the balance of benefits and costs associated with introducing the market mechanism. The altered price signal will have flow-on effects for selection of inputs to production and success in trading permits. The reduced market efficiency brings into question the benefits of introducing a market-based instrument given the need for public funding to initiate and maintain the trading system. The cost savings associated with using a market mechanism to manage environmental goods may be less than other policy options such as direct regulation of application rates. This finding has been observed in real-world examples of initiating market-based instruments for environmental regulation, with [51] and [52] finding that such markets often have not attained the cost savings that were initially projected.

This example presents a novel use of integrated ABM and experimental economics to ground agent decision making in behaviours expressed by real humans. Previously studies have used experiments and ABM separately to guide the design of agent behaviours, but are still reliant on subjective interpretation by the researcher in identifying and re-assigning behaviours. The example described here directly integrates the experiment with the ABM, reducing risks of misinterpretation provided the experimental design and model are consistent. Because the experimental interface is coupled directly with the same functions used by agents, this consistency is maintained through directly linking parameters and equations programmed into the ABM to the participant interface. This allows for a very tight definition of the decision making problem, allows control over what information is available, and helps to address the issue of whether there are inconsistencies between the experimental design and interpreted behaviours. This example of an integrated experiment and agent-based model highlights ways that the outstanding challenge of calibrating models can be approached through gathering data on human behaviour. Alongside data from surveys, interviews, companion modelling and geovisualization, experimental economics can provide empirical data for designing and parameterizing human behaviours.

DISCUSSION

Empirical calibration of ABM is important to generate confidence in model outcomes. This paper describes three case studies which highlight issues of sensitivity in micro calibration and macro outcomes, the use of context in eliciting revealed behaviours of human participants, and the importance of grounding simulated decision making in behavioural economics. Models that represent individual decision making can draw on a variety of techniques from social and economic sciences to ground assumptions in empirical data. The use of survey data can inform selection and parameterization of decision making functions, but survey data remains a snapshot in time and does not necessarily reveal motivations, strategies and adaptive behaviour. Interviews are commonly used to inform adaptive decision making and formulate decision making algorithms. Participatory approaches such as companion modelling formalize the research process of using agent models and role playing games. Geovisualization allows a context-rich visual environment to couch decisions in a realistic setting and presents stakeholders with a wealth of information. Finally, experiments with human participants offer a controlled laboratory setting where defined decision situations can be explored. Experiments can explore elements of human decision making that move beyond assumptions of fully informed rational behaviour, and re-calibration of ABMs with human-like decision making can reveal differences in model outcomes which may not be apparent when basing assumptions solely on neo-classical economic theory.

The first case study presents a method to elicit empirically-based preference functions from survey data. Survey respondents are asked to make tradeoffs between different options, eliciting preferences for attributes of the choice set. Survey data are analysed using random utility models to econometrically estimate preference functions that can be assigned to agents. This method of determining decision making functions for agents is comparable and replicable across many ABMs involving selection of a preferred option from a choice set, and is grounded in established economic theory.

The second case study presents methods to introduce context into participatory data gathering techniques. Whereas surveys and interviews elicit stated preferences and behaviours, these techniques move towards gathering revealed preference and behavioural information. In this case the specific context of the decision making situation can be used to tease out important factors influencing decision making. Geovisualization is presented as a way to communicate large amounts of contextual information to participants.

The third case study presents a novel application of integrated ABM and experimental economics. The application immerses participants in a running ABM using an experimental economics interface. The integration of the two tools assists in ensuring consistency between the experimental design and ABM equations, and also consistency between the participant data collected and the interpretation and reassignment of behaviours to artificial agents. Data from participant responses are used to re-parameterise artificial agents, comparing the difference for macro-level outcomes between optimally performing agents and agents with behaviours informed by experimental results. In the example shown here, participant trading behaviours served to increase market price and reduce efficiency.

Of the variety of calibration techniques reviewed, a multi-methods approach is suggested, through gathering information using different methods based on their merits. Surveys gather information on demographics and broad preferences, interviews begin to reveal motivations, participatory approaches introduce relationships and context with stakeholders, geovisualisation presents complex information accessibly for learning and feedback, and lastly experiments reveal behaviours for controlled decision making situations beyond neo-classical assumptions of fully informed and rational decision making. Empirically-based techniques of gathering information to develop ABMs improve the confidence in results and the applicability to real-world situations.

REFERENCES

[1] S. Heckbert, T. Baynes, and A. Reeson, 'Agent-based modeling in ecological economics', Annals of the New York Academy of Sciences, Ecological Economics Reviews, vol. 1185, pp. 39-53, 2010.

[2] D. Loiterton, and I.D. Bishop, 'Virtual environments and location-based questioning for understanding visitor movement in urban parks and gardens,' in Proceedings Real-time Visualization and Participation, 2005.

[3] D. Loiterton and I.D. Bishop, 'Simulation, calibration and validation of recreational agents in an urban park environment,' in Monitoring, Simulation and management of visitor landscapes, R. Gimblett, H. Skov-Petersen, and A. Muhar, eds, University of Arizona Press, pp. 107-122, 2008.

[4] D.T. Robinson, D.G. Brown, D.C. Parker, P. Schreinemachers, M.A. Janssen, M. Huigen, *et al*., 'Comparison of empirical methods for building agent-based models in land use science' Journal of Land Use Science, vol. 2, no. 1, pp. 31-55, 2007.

[5] D. Valbuena, P.H. Verburg, and A.K. Bregt, 'A method to define a typology for agent-based analysis in regional land-use research', Agriculture Ecosystems and Environment, vol. 128, pp. 27-36, 2008.

[6] T. Berger and P. Schreinemachers, 'Creating agents and landscapes for multi-agent systems from random samples', Ecology and Society, vol. 11, no. 2, 2006.

[7] R. Laing, A. Davies, and S. Scott, 'Combining visualisation with choice experimentation in the built environment,' in Visualisation for landscape and environmental planning: technology and applications, I. Bishop and E. Lange, eds, Taylor and Francis, pp. 212-291, 2005.

[8] A. Grêt-Regamey, I.D. Bishop, and P. Bebi, 'Predicting the scenic beauty value of mapped landscape changes in a mountainous region using GIS', Environment and Planning B, vol. 34, pp. 50-67, 2007.

[9] H.R. Gimblett and H. Skov-Petersen, 'Monitoring, simulation and management of visitor landscapes', University of Arizona Press, 2008.

[10] R. Jochem, R. van Marwijk, R. Pouwel, and D. Pitt, 'MASOOR: Modeling the transaction of people and environment on dense trail networks in natural resource settings,' in Monitoring, Simulation and Management of Visitor Landscapes, H.R. Gimblett and H. Skov-Petersen, eds, University of Arizona Press, pp. 269-294, 2007.

[11] R. Nelson, C. Sharp, C. Prew, and R. Gimblett, 'Balancing increased visitor use and wilderness character using simulation modelling in Misty Fjords National Monument', in Monitoring, Simulation and Management of Visitor Landscapes, R. Gimblett and H. Skov-Petersen, eds, University of Arizona Press, pp. 315-330, 2007.

[12] D.G. Brown, D.T. Robinson, L. An, J.I. Nassauer, M. Zellner, and W. Rand, 'Exurbia from the bottom-up: Confronting empirical challenges to characterizing a complex system', Geoforum, vol. 39, pp. 805-818, 2008.

[13] S. Heckbert, W. Adamowicz, P. Boxall, and D. Hanneman, 'Cumulative effects and emergent properties of multiple-use natural resources' Lecture Notes in Computer Science, vol. 5683, pp. 1-13, 2010.

[14] P. Boxall, W. Adamowicz, J. Swait, M. Williams, and J. Louviere, 'A comparison of stated preference methods for environmental valuation', Ecological Economics, vol. 18, no. 3, pp. 243-253, 1996.

[15] N. Hanley, R. Write, and W. Adamowicz, 'Using choice experiments to value the environment: Design issues, current experience and future prospects', Environmental and Resource Economics, vol. 11, no. 3, pp. 413-428, 1998.

[16] B. Bottan, 'Exploring the human dimension of Thunder Bay moose hunters with focus of choice behavior and environmental preferences', Unpublished thesis, Lakehead University, 1999.

[17] P. Boxall, and B. MacNab, 'Exploring the preferences of wildlife recreationists for features of boreal forest management: a choice experiment approach', Canadian Journal of Forest Research, vol. 30, pp. 1931-1941, 2000.

[18] W. Adamowicz, P. Boxall, M. Haener, Y. Zhang, D. Dosman, and J. Marois, 'An assessment of the impacts of forest management on aboriginal hunters: Evidence from stated and revealed preference data', Forest Science, vol. 50, no. 2, pp. 139-152, 2004.

[19] M. Haener, D. Dosman, W. Adamowicz, and P. Boxall, 'Can stated preference methods be used to value attributes of subsistence hunting by Aboriginal Peoples? A case study in Northern Saskatchewan', American Journal of Agricultural Economics, vol. 83, pp. 1334-1340, 2001.

[20] M. Huigen, K.P. Overmars, and W.T. de Groot, 'Multiactor modeling of settling decisions and behavior in the San Mariano watershed, the Philippines: a first application with the MameLuke framework', Ecology and Society, vol. 11, no. 2, 2006.

[21] N. Gilbert, 'Sciences sociales computationnelles: Simulation sociale multi-agents,' in Modélisation et simulation multi-agents: Applications pour les sciences de l'homme et de la société, F. Amblard and D. Phan, eds, Lavoisier, pp. 141-157, 2006.

[22] J.C. Castella, S.P. Kam, D.D. Quang, P.H. Verburg, and C.T. Hoanh, 'Combining top-down and bottom-up modelling approaches of land use/cover change to support public policies: Application to sustainable management of natural resources in northern Vietnam', Land Use Policy, vol. 24, no. 3, pp. 531-545, 2007.

[23] N. Becu, A. Neef, P. Schreinemachers, and C. Sangkapitux, 'Participatory computer simulation to support collective decision making: Potential and limits of stakeholder involvement', Land Use Policy, vol. 25, no. 4, pp. 498-509, 2008.

[24] P. Guyot and S. Honiden, 'Agent-based participatory simulations: Merging multi-agent systems and role-playing games', Journal of Artificial Societies and Social Simulations, vol. 9, no. 4, 2006.

[25] Y. Maru, K. Alexandridis, and P. Perez, 'Taking 'participatory' in participatory modeling seriously,' in Proceedings of the 18th World IMACS / MODSIM Congress, Cairns, Australia, 13-17 July, 2009.

[26] T.C. Daniel and J. Vining, 'Methodological Issues in the Assessment of Landscape Quality,' in Behavior and the Natural Environment, F. Altman and J.F. Wohlwill, eds, Plenum, pp. 39-84, 1983.

[27] E.H. Zube, J.L. Sell, and J.G. Taylor, 'Landscape perception: Research, application and theory', Landscape Planning, vol. 9, pp. 1-33, 1982.

[28] R.W. Paterson and K.J. Boyle, 'Out of sight, out of mind? Using GIS to incorporate visibility in hedonic property value models', Land Economics, vol. 78, pp. 417-425, 2002.

[29] J. Bergin and C. Price, 'The travel cost method and landscape quality', Landscape Research, vol. 19, pp. 21-23, 1994.

[30] K. Virrantaus, J. Markkula, A. Garmash, V. Terziyan, J. Veijalainen, A. Katanosov *et al*., 'Developing GIS-supported location based services', in Proceedings of the First International Workshop on Web Geographical Information Systems, Kyoto, Japan, 2001.

[31] P. Forer, H. F. Chen, and J.F. Zhao, 'Building, unpacking and visualising human flows with GIS', in Proceedings of GISR-UK, Norwich, pp. 334-336, 2004.

[32] S.M. Nusser, D.M. Thompson, and G.S. De Lozier, 'Using personal digital assistants to collect survey data', in Proceedings of the Section on Survey Research Methods, pp. 780-785, 1996.

[33] I.D. Bishop and R.B. Hull, 'Integrating technologies for visual resource management', Journal of Environmental Management, vol. 32, pp. 295-312, 1991.

[34] I.D. Bishop, C. Stock, and K. Williams, 'Using virtual environments and agent models in multi-criteria decision making', Land Use Policy, vol. 26, pp. 87-94, 2009.

[35] R. Manning and W. Freimund, 'Use of visual research methods to measure standards of quality for parks and outdoor recreation', Journal of Leisure Research, vol. 36, pp. 557-579, 2004.

[36] B. Rohrmann and I.D. Bishop, 'Subjective responses to computer simulations of urban environments', Journal of Environmental Psychology, vol. 22, pp. 319-331, 2002.

[37] I.D. Bishop and B. Rohrmann, 'Subjective responses to simulated and real environments: a comparison', Landscape and Urban Planning, vol. 65, pp. 261-277, 2003.

[38] J. Jallouli and G. Moreau, 'An immersive, multisensory and dynamic approach for landscape study', in Innovations for building and construction, 12th European Conference Proceedings, E. Dado, R. Beheshti, and K. Zreik, eds, Europia Productions, pp. 121-140, 2009.

[39] H.A. van Veen, H.K, Distler, S.J. Braun, and H.H. Bulthoff, 'Navigating through a virtual city: Using virtual reality technology to study human action and perception' Future Generation Computer Systems, vol. 14, pp. 231-242, 1998.

[40] L. Hodges, B. Rothbaum, R. Kooper, D. Opdyke, T. Meyer, M. North *et al*., 'Virtual environments for treating fear of heights', IEEE Computer, vol. 28, no. 7, pp. 27-34, 1995.

[41] L. Kennedy and I.D. Bishop, 'Investigating land-use decision making in a virtual environment', in Landscape Analysis and Visualisation, Spatial Models for Natural Resource Management and Planning, C. Pettit, W. Cartwright, I. Bishop, K. Lowell, D. Puller, and D. Duncan, eds, Springer-Verlag, pp. 591-608, 2008.

[42] A. Reeson and K. Nolles, 'Experimental economics: Applications to environmental policy', Socio-Economics and the Environment in Discussion, CSIRO Working Paper Series, vol. 3, pp. 1834-5638, 2009.

[43] T. Evans, W. Sun, and H. Kelley, 'Spatially explicit experiments for the exploration of land-use decision-making dynamics', International Journal of Geographical Information Science, vol. 20, no. 9, pp. 1013-1037, 2006.

[44] J. Duffy and M. Unver, 'Asset price bubbles and crashes with near-zero-intelligence traders', Economic Theory, vol. 27, pp. 537-563, 2006.

[45] S. Manson and T. Evans, 'Agent-based modeling of deforestation in southern Yucata´ in, Mexico, and reforestation in the Midwest United States', Proceedings of the National Academy of Sciences, vol. 104, no. 52, pp. 20678-20683, 2007.

[46] S. Heckbert, 'Agent-based modeling of emissions trading for coastal landscapes in transition', Journal of Land Use Science, in press.

[47] S. Heckbert, 'Experimental economics and agent-based models', in Proceedings of the 18th World IMACS / MODSIM Congress, Cairns, Australia, 13-17 July, 2009.

[48] A. Hailu and S. Thoyer, 'Multi-unit auction format design', Journal of Economic Interaction and Coordination, vol. 1, pp. 129-146, 2006.

[49] T. Cason and D. Friedman, 'Price formation in double auction markets', Journal of Economic Dynamics and Control, vol. 20, pp. 1307-1337, 1996.

[50] T. Cason and D. Friedman, 'Price formation in single call markets' Econometrica, vol. 65, pp. 311-345, 1997.

[51] N. Netusil and J. Braden, 'Transaction costs and sequential bargaining in transferable discharge permit markets', Journal of Environmental Management, vol. 61, pp. 253-262, 2001.

[52] S. Atkinson and T. Titenberg, 'Market failure in incentive-based regulation: The case of emissions trading', Journal of Environmental Economics and Management, vol. 21, pp. 17-31, 1991.

CHAPTER 7

Geosimulation of Income-Based Urban Residential Patterns

Erez Hatna[1,*] and Itzhak Benenson[2]

[1]Institute for Management Research, Radboud University, the Netherlands and [2]Department of Geography and Human Environment, Tel Aviv University, Israel

Abstract: Maps of high-resolution residential patterns of Israeli cities reveal essential spatial heterogeneity in relation to family income, a situation that was found in eight of the nine cities investigated. Modern urban theory provides several explanations of how homogeneous urban patterns self-organize and persist; however the mechanisms that allow the persistence of heterogeneous patterns remain unexplained. We argue that the observed residential heterogeneity can be explained on the basis of householders' residential preferences and behavior. Namely, the householders differ in respect to their willingness to reside near poorer neighbors. Residential heterogeneity is the consequence of this variability.

In order to investigate this hypothesis, we developed an agent-based model of residential dynamics in the city, which extends the Schelling model of segregation. Model residential agents represent families, which differ in their economic abilities and their tolerance of poorer neighbors. Using the model, we demonstrate that quite a low fraction of tolerant agents is sufficient for the emergence and persistence of heterogeneous residential areas in the city. This result is robust to the uncertainty of our knowledge about the fraction of highly tolerant agents – the variation in this fraction only weakly influences the resulting urban heterogeneity. We thus conclude that the presence of residential agents who are tolerant of their poorer neighbors is sufficient to explain long-term urban heterogeneity and we discuss the possible consequences of this result for the theory of urban gentrification.

INTRODUCTION

The population of modern cities consists of various ethnic, cultural and socioeconomic groups. This heterogeneity has significant implications regarding the city structure and dynamics: the interaction between group members can increase social cohesion or, alternatively, result in social avoidance and segregation. However, until recently, the standard view of urban residential patterns is that of a pattern of relatively homogeneous areas, each characterized by a population of a given ethnicity or socio-economic status. This representation of cities dates back to social ecology models [1], [2]. For example, the concentric [2] and sector [3] models included homogeneous areas, characterized by low, medium and high status populations. Social ecology models were mostly theoretical; empirical investigations of the urban residential structure started in the 1970s with the introduction of Factorial Ecology. The Factorial Ecology methodology made use of census data and statistical methods, such as factor and cluster analysis, in order to indentify the main socio-economic characteristics of urban units and to recognize units with similar characteristics [4]. However, the urban data were always available at the level of census tracks or statistical areas, while urban patterns at the level of separate buildings and apartments remained unknown.

The importance of the high-resolution view of the urban residential pattern was recognized in the early 1970s, when Thomas Schelling's model of segregation was published [5]. The Schelling model is an illustration of how individual incentives and perceptions of the differences between social groups can lead to the collective phenomenon of segregation [6]. Although the model is abstract and thus can represent many phenomena, its interpretation in relation to residential dynamics of social groups is most evident. Thomas Schelling illustrates his idea by placing equal numbers of dimes and pennies (representing agents of two social groups) on a chessboard-size sheet of paper, one coin in a cell, and allowing the agents to change their residence (cell) by assuming that every agent wants more than a certain fraction **F** of its neighbors within the 3x3 neighborhood to be like himself. An agent located within an unsatisfactory neighborhood can migrate to the nearest cell that satisfies his demand [6]. Studying the model, Schelling [6] demonstrated that no matter what the initial pattern and the order of moves, the demand of **F = 1/3**, *i.e.*, for only one-third of the "friendly" neighbors within the

***Address correspondence to Erez Hatna:** Centre for Advanced Spatial Analysis (CASA), University College London (UCL), 90 Tottenham Court Road, London W1T 4TJ, UK; Email erezh51@gmail.com

Danielle J. Marceau and Itzhak Benenson (Eds)

neighborhood, results in rapid segregation of the pattern into large homogeneous patches of dimes and pennies. The Schelling model comprises residential agents of two mutually avoiding types of "Blacks" and "Whites". The reality is evidently more complex, with many factors that determine residential choice in cities. However, the theoretical result of Schelling remains valid – a relatively weak residential preference towards residing in a neighborhood that contains a sufficiently high fraction of agents, who are similar to the resident of the central cell, is enough to lead to self-organization into a segregated residential pattern that consists of extended homogeneous Black and White areas. However, real cities are essentially less homogenous, in comparison to the level produced by the Schelling model and several attempts to relate between the real-world residential distribution and Schelling model were performed during last decade by Mark Fossett [7], [8] and discussed in [9], [10]. The papers of Fossett and Clark consider complex situation of a modern American city, populated by three ethnic groups – White, Black and Hispanic, which members have different ethnic residential preferences and, in addition, differ in their economic abilities. However, the complexity of the model prevents the authors from direct account of the interaction between the resident and the habitat. The latter can be ignored when the ethnic relations are considered, but should be considered in detail when the economic segregation is considered. This paper thus focuses on explaining the level of income heterogeneity of the residential distribution in real cities, while using the concept of the Schelling model. Following Schelling's view of neighborhood preferences as essentially determining urban residential dynamics, we explore the homo/heterogeneity level of residential income patterns in Israeli cities. We base our analysis on the high-resolution data of the 1995 population census on family income. Based on experimental results, we propose a mechanism that can explain the patterns' genesis and persistence. The use of high-resolution GIS datasets, combined with Geosimulation methodology [11], enable the modeling of urban spatial patterns as an outcome of individual behavior. We employ this approach in order to simulate the level of residential heterogeneity observed in nine Israeli cities.

The paper is organized as follows: we start with the study of high resolution patterns of family income in nine Israeli cities. The high-resolution maps reveal high levels of spatial heterogeneity in eight cities. This finding suggests that individuals of different economic status tend to live in close proximity. We propose two explanations of this phenomenon: the first assumes that buildings of different price range and quality are located in close proximity. The second is related to tolerant householders, who are willing to settle or remain near neighbors of lower economic status. We use regression analysis to test both explanations and conclude that both explanations may contribute to the high spatial heterogeneity of income. The rest of the paper focuses on investigating the feasibility of the second explanation, using an extension of the Schelling model of segregation. Using the model, we demonstrate that a low fraction of tolerant residential agents can generate heterogeneous residential patterns. We further demonstrate that the level of heterogeneity generated is relatively insensitive to the amount of tolerant agents.

THE EMPIRICAL STUDY

The Database

The research is based on the database on families and individuals obtained by the Israeli Central Bureau of Statistics (ICBS) during the population census of 1995 [12]. The family data were related by the ICBS personal identifier to data on personal income in 1995, supplied by the Israeli National Insurance Institute. To follow Israel's privacy law, the database was coded by the ICBS personnel, in order to avoid the possible identification of the characteristics of specific individuals and then it became available for academic research [12]. The dataset included a GIS layer of building foundations and a non-spatial table that contained family and individual characteristics collected during the census, such as country of origin, age, level of education, number of children and age. Geo-referencing of the dataset in relation to the building foundation layer enabled the construction of maps at the resolution of individual buildings.

Our analysis was performed on nine Israeli cities located in the center of the country. Tel-Aviv, with 350,000 residents, is the largest and the populations of the others vary from 150,000 (Netanya) to less than 40,000 (Rosh-HaAyin) Table **1**. Three family characteristics – income, parents' origin and number of children, which are especially important for our study, were available for at least 90% of the families in each city. Data on educational level were available for 15% of the households, while householders' estimates of the year of building construction were available for 5% of the buildings.

THE CONSTRUCTION OF INCOME MAPS

In order to make the family income distribution closer to normal, the income values were transformed into a $\log_2$ scale. The $[x, x + 1]$ interval, in terms of a logarithmic scale, thus corresponds to the $[2^x, 2^{x+1}]$ interval of actual income. To make residential patterns visually comprehensible, the high-resolution maps we present are not based on the layer of building foundations, but on the layer of Voronoi polygons, built on the basis of building centroids. The Voronoi coverage enables a spatially continuous view of the urban pattern and also serves for defining neighborhood relations between the buildings. Namely, we defined the first order neighborhood of a building **H** as **H** itself, plus the set of buildings, which have Voronoi polygons that share a common edge or node with the Voronoi polygon of **H** (Fig. **1**), the green polygons). We define, recursively, the neighborhood of order **k,** as consisting of the buildings of the neighborhood of order **k-1** plus their neighbors. As presented in Fig. **1**, the red building's neighborhood of the second order consists of the buildings within the green and yellow Voronoi polygons, and that of the third order consists of the buildings within the green, yellow and aqua polygons. To obey legal limitations in regards to the privacy, in what follows, we investigate and present the data on the basis of the neighborhoods of the third order.

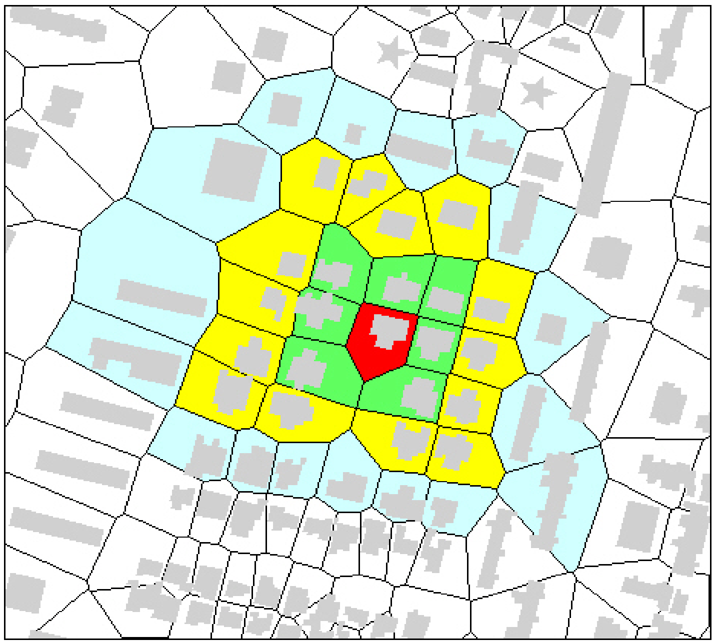


Figure 1: A layer of buildings and the corresponding layer of the buildings' Voronoi polygons. The polygons marked in green, yellow and blue represent (respectively) the first, second and third order neighbors of the central (red) building

To comply with privacy restrictions, we use the mean value of **Log_2(Income)** of the households in the building and present the maps of the non-weighted average of these values over all households in the building's neighborhood. In addition to maps of average values, we also present maps of standard deviation (STD) of the average values of **log_2(Income)** over the building's neighborhood. Although the map of the averages essentially smoothes the income pattern, it still reflects poor/rich urban areas; the STD-map reflects the local level of the urban income heterogeneity.

RESULTS OF THE EMPIRICAL STUDY

The Residential Income Maps

Fig. **2** presents the residential income pattern for Bat-Yam, the most homogeneous of the nine investigated cities. The average STD and the STD's 95^{th} percentile over Bat-Yam's neighborhoods are the lowest of the nine cities Table **1** thus the Bat-Yam maps of the Log_2(Income) local mean and STD coincide most perfectly with the Schelling model for the case of agents characterized by a continuous "economic status" – as demonstrated by [13]; if residential agents prefer neighbors with incomes similar to their own, then the residential pattern is characterized by the clear poor-rich gradient. Fig. **3** presents the opposite case – the residential pattern for Kfar-Saba, which is the most heterogeneous of the nine investigated cities, with rich and poor residents living side by side. The intermediate case of the city of Ramla is presented in Fig. **4**. Note that according to Table **1**, the low residential heterogeneity of Bat-Yam is an exception and the rest the cities are closer to Ramla and Kfar-Saba in their residential heterogeneity.

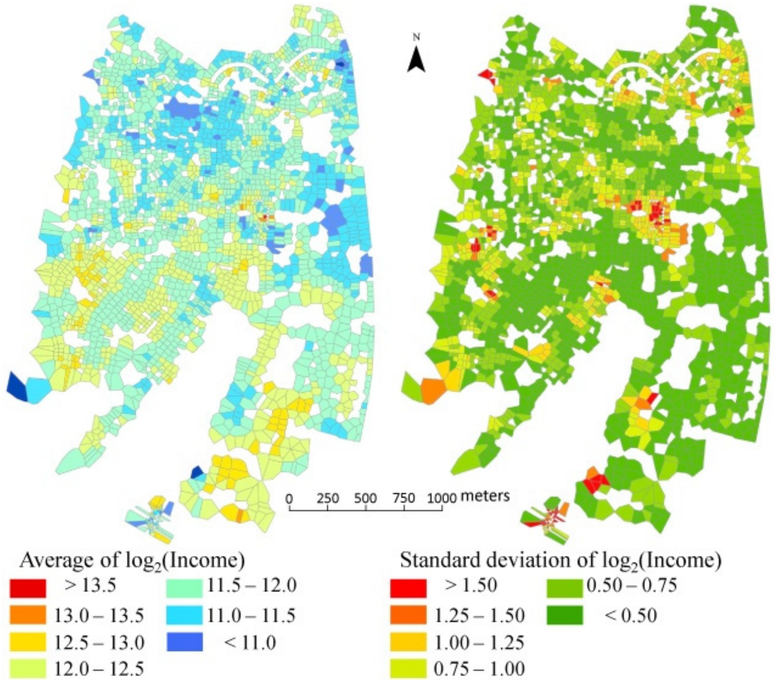


Figure 2: Income average and standard deviation over the residential buildings' neighborhood for the city of Bat-Yam

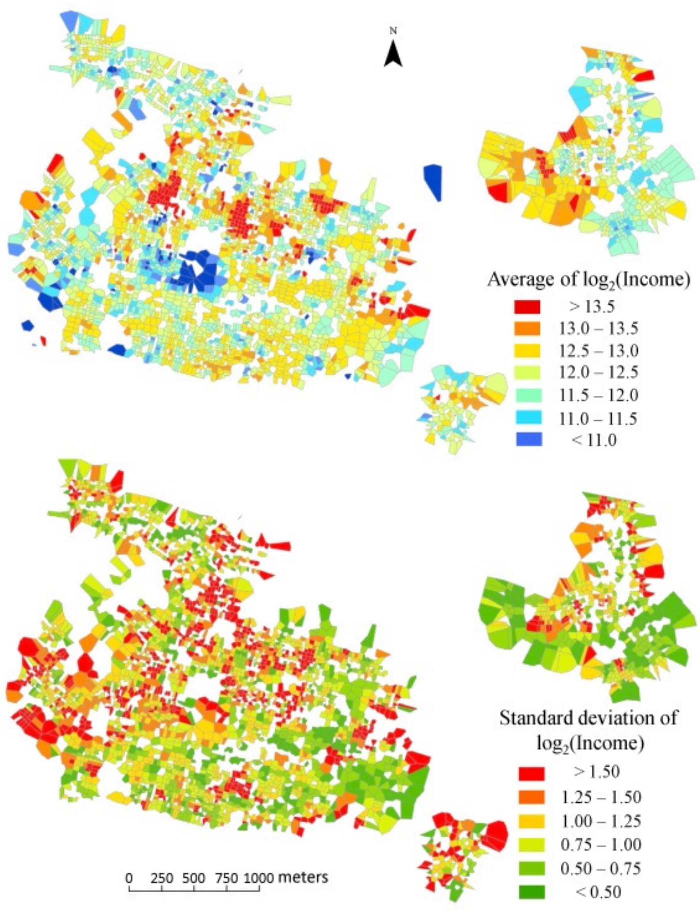


Figure 3: Income averages and standard deviations over the residential buildings' neighborhood for the city of Kfar-Saba

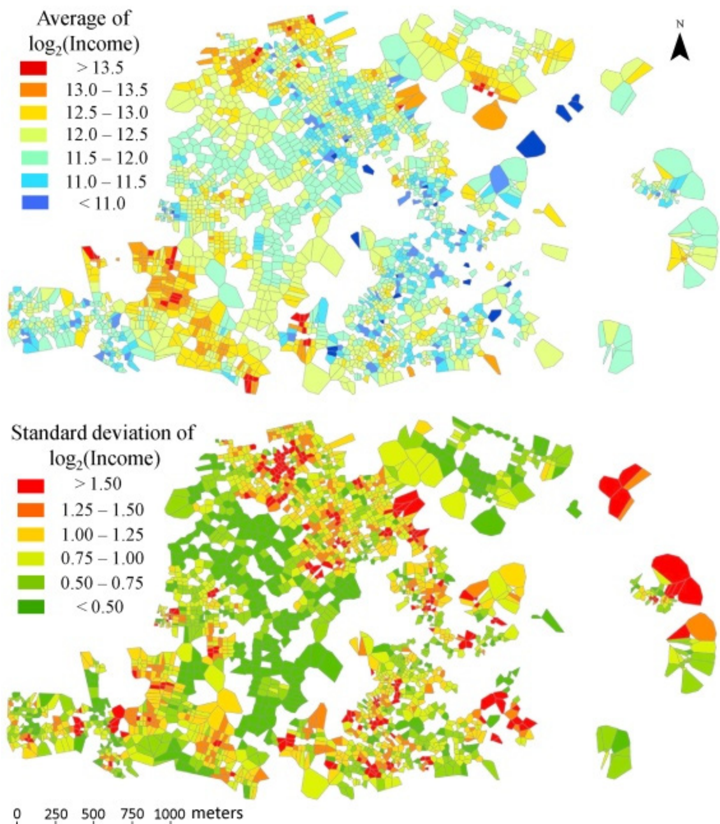


Figure 4: Income average and standard deviation over the residential buildings' neighborhood for the city of of Ramla

Table 1: Basic characteristics and spatial heterogeneity of the nine cities.

City	Bat-Yam	Ashdod	Lod	Tel-Aviv	Ramla	Netanya	Rosh HaAyin	Kfar Saba	Ramat Hasharon
Population (in 1000s)	140	130	52	350	40	150	40	70	40
Populated buildings	2,485	2,869	1,814	17,208	2404	5,287	3,017	3,234	3,140
Mean values of STD over the neighborhoods	0.593	0.804	0.876	0.888	0.923	0.941	1.047	1.075	1.325
95^{th} percentile of STD over the neighborhood	1.255	2.065	1.972	1.983	1.701	1.935	2.206	2.042	2.443

Several hypotheses that explain why richer householder stay close to poor ones can be proposed:

1. Lack of residential opportunities: householders may not have enough residential alternatives to choose from and are thus forced to stay close to the poorer neighbors at their current or potential location.
2. Wealthier householders are attracted to local features, such as a single new and expensive residential buildings constructed in poor areas and just ignore the proximity to the poor neighbors in this case.
3. Educated householders, who are expected to be more tolerant, accept staying within poorer neighborhoods, taking advantage of the cheaper living conditions there, while poorer families with children prefer to stay in wealthy neighborhoods, to take advantage of better schools and other advantages of the "educated neighborhood".

We did not investigate every explanation in depth and limit ourselves to the most general arguments when comparing the plausibility of the above hypotheses: We reject the first one, as the typical annual inter-city migration rate in Israel is very high, about 5% and we consider such a high migration rate as hardly possible for cities with an essentially limited set of residential opportunities. The second and third explanations can be verified in some depth, by estimating the correlation between families' income and other characteristics. As we are mostly interested in areas of high income heterogeneity, we calculate the correlations only over the buildings which are sited inside the heterogeneous areas.

Correlation Analysis

The census data include the year of building construction, number of years of education of the head of family and the number and age of children. We thus estimated correlations between the mean family income and the education level of the head of family, the number of children of school age per family, and the building's age. We defined an area as "heterogeneous" on the basis of third order neighborhoods (Fig. **1**), which are considered "heterogeneous" if the STD of the neighborhood buildings' **Log2(Income)** is above 1.25; these areas comprise 20-40% of the cities' areas, with the exceptions of Bat-Yam (3.5%) and Kfar Saba (69%) Table **2**.

Table 2: Correlation between mean household income in a building and neighborhood characteristics over the areas satisfying heterogeneity conditions

City	Percentage of city areas with STD > 1.25 over third-order neighborhoods	Fraction of children in a building	N	Fraction of householders who graduated from high school	N	Year of building's construction	N
Bat Yam	3.5	0.035	88	0.226	27	0.233	5
Ashdod	22.8	-0.269**	656	0.187**	224	0.224	29
Lod	20.0	-0.124*	365	0.354**	121	0.055	26
Tel Aviv	20.0	0.158**	3417	0.221**	1440	0.581**	117
Ramla	18.0	0.016	437	0.202**	164	0.390	15
Netanya	23.8	0.165**	1289	0.262**	517	0.140	30
Rosh HaAyin	34.9	0.058	1057	0.291**	283	0.702	6
Kfar Saba	44.2	0.177**	1435	0.391**	650	0.443**	98
Ramat Hasharon	68.8	0.126**	2166	0.255**	831	0.041	51

* $p<0.05$, ** $p<0.01$

The correlations between family income and the fraction of children remains relatively low and both positive and negative correlations exist, while the correlations between income and education, as well as income and building age are always positive – that is, richer householders in heterogeneous areas live in newer buildings and are more educated. The analysis is, thus, in favor of the second explanation: dwellings in more expensive and newer buildings within the heterogeneous neighborhoods are more often populated by more educated families, usually with higher incomes, while the proximity of buildings of different price ranges and age contributes to the income heterogeneity of cities.

Survey Of Householders' Attitude To Neighbors

However, the data do not allow us to quantify the relation between families' level of income or education and their tolerance to poorer neighbor. To acquire an additional dimension of understanding, we conducted a small survey. Namely, based on the available *spatial* information of the families' income, we identified a number of wealthier householders residing among poor neighbors. We visited them and asked why they stayed in the neighborhoods. We were not able to conduct a deep study and our questionnaire contained three simple and straightforward questions of the following kind: "How important is it for you that – X – among your neighbors is the same as yours?" where "X" referred to "economic status," "culture" or "education." The respondents were asked these questions twice, one regarding neighbors living in their residential buildings and once regarding neighbor living in their proximity (but in different buildings). The answers were coded using a Likert scale ranging from 1 (very unimportant) to 5 (very important). The survey was conducted in Tel-Aviv and included 20 wealthy families residing within heterogeneous neighborhoods and a control group of 20 wealthy families residing within homogeneous areas (and thus overall wealthy). The neighborhoods were chosen on the basis of high-resolution maps. Specific apartments of "wealthy" appearance were selected during visits to heterogeneous neighborhoods, whereas apartments in homogeneous wealthy neighborhoods were randomly selected. The results of the survey are presented in Table **3Error! Reference source not found.Table 3:** Mean of answers to following questions: How important is it for you that your neighbors in [your house/neighboring houses] are similar to you in [characteristic X]?

Characteristic X	**In your house**			**In neighboring houses**		
	Rich among poor (n = 18)	**Rich among rich (n = 13)**	**p (t-test)**	**Rich among poor (n = 20)**	**Rich among rich (n = 20)**	**p (t-test)**
Economic status	2.56	3.31	~0.85	2.20	3.10	~0.10
Cultural level	2.72	4.00	~0.01	2.35	3.75	<0.01
Level of education	2.21	3.38	<0.01	1.80	3.10	<0.01

The results suggest that the householders living in heterogeneous areas appear to be more tolerant of other families compared to those living in homogeneous areas. The differences in the mean response values of householders residing in the heterogeneous and homogeneous neighborhoods were found to be significant (based on two-side t-test) for two characteristics – "cultural level" and "level of education," both for neighbors in the same building and neighbors within the building's neighborhood. The difference in mean response in relation to economic status was found to be insignificant.

Although the survey is limited and preliminary, it supports the hypothesis that the residential heterogeneity is a result of tolerant wealthy householders willing to live near neighbors of lower economic status. In order to test the feasibility of this hypothesis, we developed a Schelling-like model of income-based residential dynamics in an abstract city, where residential agents are characterized not only by their level of income, but, also, by a *level of tolerance*. We use the model in order to investigate whether the presence of the tolerant householder agents can explain the residential income heterogeneity, as observed in the Israeli cities. In particular, we used the model to test whether a low fraction of tolerant agents in the city is sufficient for the emergence of a persistent heterogeneous residential pattern and how sensitive are the results to the fraction of tolerant agents.

THE MODEL OF INCOME-BASED RESIDENTIAL DYNAMICS

To investigate the residential income patterns, we consider the Schelling-like model, where householder agents occupy and migrate between one-family "buildings", represented by cells of a square grid – "city." The grid cells can be considered as polygons of the Voronoi coverage constructed for buildings located in the cells' centers and, just as above, two cells are neighbors if they have a common edge or node.

Householder agents represent families that differ in their economic abilities and tolerance to poorer neighbors and locate and relocate in the city. In addition to internal migration, we also account for immigration into the city and emigration from it.

Formally, the model householder agent **g** is characterized by monthly "income" **E(g)** that is represented using a base 2 logarithm scale. The value of **E(g)** is assigned according to a truncated log-normal distribution, based on the estimates obtained in the nine investigated cities. We consider $\mathbf{log_2(E(g))}$ as normally distributed, with the average of 12 (2^{12} = 4096 NIS per month), and STD of 2.2. We truncate this distribution by assuming that **E(g)** must be above $\mathbf{E_{min}}$ = 9 (that is, above 2^9 = 512 NIS, which is taken as the minimal possible monthly income in 1995, the census year).

Householder agent **g** is also characterized by tolerance level **TL(g),** which represents the willingness of an agent to reside near other agents of lower economic status. An agent **g** of zero tolerance, **TL(g)** = 0, is maximally intolerant of the neighbors, while that of the unit tolerance, **TL(g)** = 1 does not react to the differences between itself and the neighbors.

A dwelling **c** is characterized by a "price" **P(c)**, which is also represented in terms of log transformed monthly income. The price **P(c)** is set according to the income of the agents, which occupy **c**'s neighborhood (additional description is given below). The cells of the 3x3 neighborhood (Moore neighborhood) of **c** are considered to be **c**'s neighborhood and denoted **N(c)**. The householder agents located within **N(c)** are considered as **g**'s neighbors. The model operates in discrete time. At each time step, residents make decisions as to whether to stay at their current location or move to another dwelling. The decisions are based on the disutility of the current location and on the utilities of the vacant, *i.e.*, non-occupied, cells.

The Disutility of the Currently Occupied Dwelling

If agent **g** occupies cell **c**, then the disutility **DU(c, g)** of **c** for **g** is determined by the average income $\mathbf{E_{N(c)}(g)}$ of **g**'s neighbors; **g** reacts to its neighbors if they were, on average, poorer than **g**; the strength of the reaction depends on the agent's tolerance **TL(g)** as follows:

$$\mathbf{DU(c, g) = [E(g) - E_{N(c)}(g)] * [1 - TL(g)] \text{ if } E(g) > E_{N(c)}(g)} \quad (1)$$

$$\mathbf{DU(c, g) = 0 \text{ if } E(g) \leq E_{N(c)}(g)}$$

The disutility **DU(c, g)** is piecewise linearly converted into probability **PL(c, g)** of entering a search pool of internal migrants (Fig. **5a**):

$$\mathbf{P_L(c, g) = P_{Leave}*DU(c, g) / (0.5*P(c)) \text{ if } 0 \leq DU(c, g) \leq 0.5*P(c)} \quad (2)$$

$$\mathbf{P_L(c, g) = P_{Leave} \text{ if } 0.5*P(c) < DU(c, g),}$$

where $\mathbf{P_{Leave}}$ = 0.004 (which is a rough estimation of the, monthly migration rate in Tel-Aviv), *i.e.*, the probability that **g** starts to search for another residence reaches 1 when the disutility of the current location **c** for **g** equals half of its price.

The Utility of the a Vacant Dwelling

The utility of an agent in relation to a dwelling

The utility **U(v, g)** of a vacant location **v** for **g** is defined as the sum of two components : (a) an *Economic component* $\mathbf{U_E(v, g)}$ with respect to its price **P(v)** and (b) a *Social component* – $\mathbf{U_N(v,g)}$ with respect to the average income $\mathbf{E_{N(v)}(g)}$ of **g**'s potential neighbors at **v:**

$$\mathbf{U(v, g) = U_E(v, g) + U_N(v, g)} \quad (3)$$

The Economic Component of Utility

To estimate the economic utility $\mathbf{U_E(v, g)}$ of a vacant place **v,** we assume that a residential agent **g** is willing to spend a certain fraction **k** of its income **E(g)** for a dwelling. This "housing budget" **k(E(g))** is assumed to be the entire income **E(g)** of poor agents, although only a portion of income for wealthier agents. In the model, we assume that $\mathbf{k(E(g)) = k_0*E(g)}$, where:

$$\mathbf{k_0 = 1 \text{ if } E(g) \leq E_{avg}\ k_0 = 2/3 \text{ if } E(g) \geq E_{avg} + 2*E_{STD}} \quad (4)$$

where $\mathbf{E_{avg}}$ is the global average of agents' income in the city. We also assume that $\mathbf{k_0}$ linearly decreases from the value of 1 to 2/3 with increasing **E(g)** within the interval ($\mathbf{E_{avg}, E_{avg} + 2*E_{STD}}$) (Fig. **5b**). The "housing budget" is fully spent, even if agent **g** locates in a "cheap" house **v**, for which **P(v) < k(E(g))**. If so, we assume that the residual of the **g**'s budget **ΔE = k(E(g)) - P(v)** is nonetheless spent on upgrading **v** "to **g**'s level"; moreover, in respect to the situation in Israel, we assume that re-locating to the cheaper house and spending the rest of the budget for upgrading is worth more than direct investment of the entire dwelling budget into the residence. That is, in case of **P(v) ≤ k(E(g)),** when the agent's housing budget is higher than the house price, the residual **ΔE** of the dwelling budget is invested in upgrading and the resulting "economic utility" of **v** is higher than that of the direct purchasing a dwelling

v of a price **k(E(g))**. Alternatively, over-spending, in the case of a "too expensive" dwelling **v**, *i.e.* **k(E(g)) < P(c)**, results in an additional penalty (say, mortgage interest) on **v**'s utility.

To express the above assumptions, the economic component of utility $\mathbf{U_E(v, g)}$ of the vacant cell **v** for occupation by agent **g** is thus defined as follows:

$$\mathbf{U_E(v, g) = E(g) - k(E(g)) + m*(k(E(g)) - P(v)) \text{ if } P(v) \le k(E(g))} \quad \mathbf{(5)}$$

$$\mathbf{U_E(v, g) = E(g) - P(v) - m*(P(v) - k(E(g))) \text{ if } k(E(g)) < P(v)}$$

where for simplicity, the same constant **m** > 0 defines the benefit of upgrading, or the penalty for taking a mortgage.

The Social Component of Utility

The social utility $\mathbf{U_N(v, g)}$ of a vacancy **v** for occupation by **g** is defined as follows:

$$\mathbf{U_N(v, g) = E_{N(v)}(g) - E(g) \text{ if } E_{N(v)}(g) > E(g)} \quad \mathbf{(6)}$$

$$\mathbf{U_N(v, g) = (E_{N(v)}(g) - E(g)) * (1 - TL(g)) \text{ if } E_{N(v)}(g) \le E(g)}$$

That is, the social utility is positive where the neighbors are richer than the householder and negative if the neighbors are poorer. It is important to note that negative social utility might cause the total utility of the vacancy to be negative; these vacancies are ignored by agents when searching for a location to occupy.

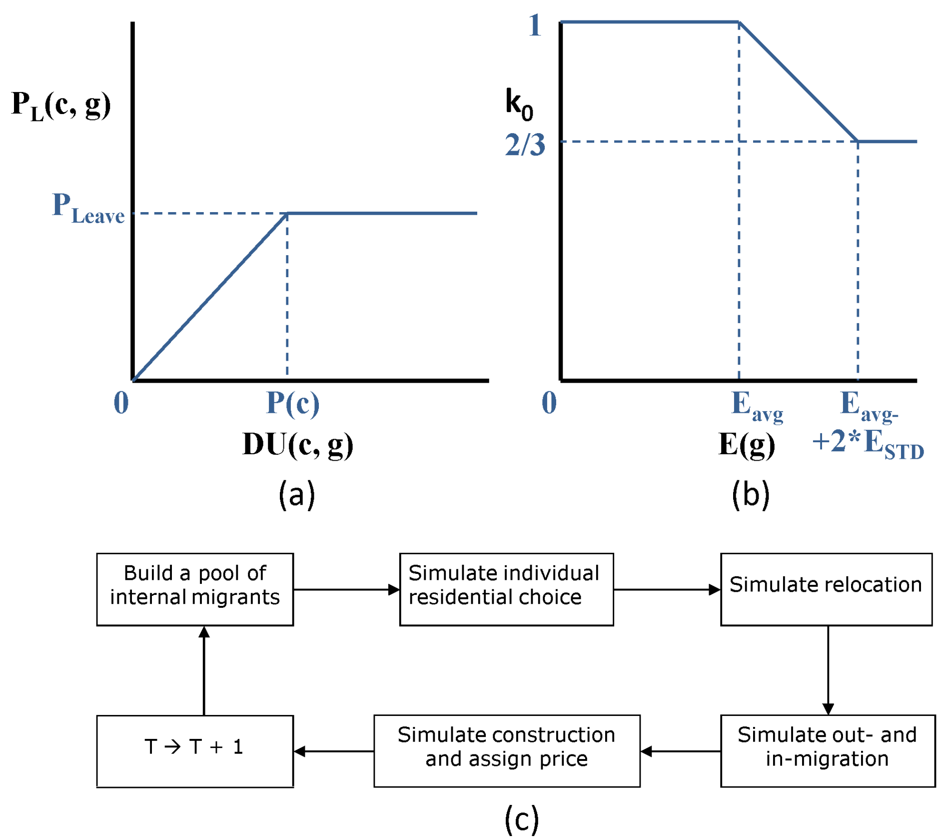


Figure 5: (a) The probability **PL(c, g)** that agent **g** occupying cell **c** will try to relocate as dependent on **DU(c, g)**, **c**'s disutility for **g**; (b) The income fraction $\mathbf{k_0}$ that an agent **g** is willing to spend on a dwelling as a function of agent's income **E(g),** and the average ($\mathbf{E_{avg}}$) and standard deviation ($\mathbf{E_{STD}}$) of the income over the entire population of the city. (c) Flow-chart of model events

Individual Residential Search and Relocation

The agents included in the search pool are considered in random order and when considered, each agent **g** in the pool constructs a list of **q** vacancies $\{v\}_g$. To construct $\{v\}_g$, **g** randomly selects **q** locations from the set of all vacant dwellings at a current time-step **t.** If **g** is an internal migrant, then a vacancy v_0 is included in the list $\{v\}_g$, if the utility $U(v_0, g)$ is higher than the utility of **g**'s current dwelling **c** and **g**'s income **E(g)** is higher than the price $P(v_0)$ of v_0 (We shall consider external immigrants in the next section). Agent **g** then sorts the dwellings in $\{v\}_g$ according to utility **U(v, g)** from high to low utility values.

After constructing a list of vacancies $\{v\}_g$, **g** attempts to occupy the best dwelling in it (the dwelling with the highest utility). If this dwelling is still vacant, then **g** relocates to the new dwelling. If the best vacancy in $\{v\}_g$ is already occupied by one of the agents who considered it before **g**, **g** remains in the pool. During this first pass on the pool of agents who aim at relocating, some agents relocate to the "best-for-them" dwelling. Those who fail to do so are randomly reordered and try to occupy the second-best vacancy in their $\{v\}_g$. The process continues until all agents find another location, or all **q** vacancies are tested but not taken. In the latter case, **g** stays at the current location. The flow of model events is presented in Fig. **5c**.

In-Migration and Out- Migration

The city is considered as an open system, where new agents (immigrants) enter at each time-step. The number of immigrants entering the city at a time-step **t** is calculated as **n(t)/100 + 10**, where **n(t)** is the number of occupied dwellings in the city at time-step **t**. Based on the experimental data, the income **E(g)** of immigrant **g** is assigned according to a truncated log-normal distribution. The tolerance **TL(g)** of immigrant agent **g** is assigned independently of **E(g)**, and the distribution of **TL(g)** depends on the scenario. The agents can leave the city for random reasons; we assume that the probability of this equals P_{Leave} = 0.004 per time step, the latter being a rough estimate of the monthly emigration rate for Tel-Aviv.

Initial Conditions

Initially, the "city" starts as a 3x3 square of occupied buildings, located in the center of the 100x100-cell grid. The rest of the cells do not contain buildings. Respectively, nine first urban residents are located in these buildings and assigned income **E(g)** and tolerance **TL(g)** according to the chosen scenario. These agents are randomly located in cells of the **3x3** neighborhood in the center of a grid; the price of dwelling **c** occupied by agent **g** is set equal to **k(E(g)).**

Dwelling Construction and Pricing

The maximum number of dwellings that can be constructed (*i.e.*, the cells that are activated as vacant dwellings), at each time-step **t,** is equal to the number of immigrants entering the city. A dwelling cell is activated if at least three of the neighboring cells within its 3x3 Moore neighborhood are occupied. The price of a new dwelling **c,** activated at time-step **t**, is assigned according to the weighted average of the neighbors' willingness to pay, averaged over the 5x5 Moore neighborhood **N(c)** of **c:**

$$P_t(c) = \Sigma_{h\in N(c)} k(E_t(h_d)) * w_d(h) / \Sigma_{h\in N(c)} w_d(h) \quad (7)$$

where $w_d(h)$ denotes the influence of the neighboring building **h** on the price of **c**. In what follows, we employ the values $w_1 = 1$ for the buildings at a block distance 1 and $w_2 = 0.5$ for those at a block distance 2 from **c**. The price **P(c)** of the dwelling **c** remains unchanged throughout the simulation.

Scenarios

Our model is nothing but an abstract representation of the possible mechanism of residential agents' behavior, where the agents differ in their tolerance to the poorer neighbors. Thus, our model scenarios aim at investigating qualitative correspondence between residential patterns in Israeli cities and model patterns. For every scenario, we examine the maps of average income in terms of the average and STD of the $Log_2(E(g))$ over the 3x3 neighborhood, in the same way as we did for the real-world data (Figs. **2**-**4**). As for the real data, we present a histogram of the STDs over the city neighborhoods and 95th percentile STD_{95} of this distribution.

The investigated scenarios differ in respect to the distribution of agents' tolerance **TL(g)**. In the first series of scenarios, all agents have the same tolerance, *i.e.*, **TL(g) = $\mathbf{TL_0}$** and we compare the scenarios with $\mathbf{TL_0}$ = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9.

In the second series, we introduce a two-value distribution of tolerance. Some of the agents are characterized by relatively low tolerance **TL(g)** = $\mathbf{TL_{Min}}$ to the poor neighbors, while the rest are characterized by relatively high tolerance **TL(g)** = 0.9. We consider five series with $\mathbf{TL_{Min}}$ **=** 0.1, 0.2, 0.3, 0.4 and 0.5, respectively, and for each value of $\mathbf{TL_{Min}}$, we define eight sub-series that differ in the percentage of the highly tolerant agents with **TL(g)** = 0.9 in the population. The percentages $\mathbf{p_{0.9}}$ of the highly tolerant agents are as follows: $\mathbf{p_{0.9}}$ **=** 2.5%, 5%, 10%, 15%, 20%, 30%, 40% and 50%.

In the third series of scenarios, agent tolerance is assumed to be uniformly distributed on the interval **[$\mathbf{TL_{Min}}$,0.9]**; the values of $\mathbf{TL_{Min}}$ = 0.1, 0.2, 0.3, 0.4 and 0.5 are considered**.**

Depending on the scenario, the city sprawls over the 100x100 grid during 400-500 iterations and then the city pattern stabilizes.

The list of numeric model parameters that are kept constant between scenarios:

1. Model time-step: **one month.**
2. Weights employed for calculating neighborhood status: $\mathbf{w_1}$ = 1**,** $\mathbf{w_2}$ = 0.5.
3. Maximum probability of leaving location $\mathbf{P_{Leave}}$ (per month): $\mathbf{P_{Leave}}$ = 0.004.
4. Agent's minimal willingness to pay: $\mathbf{k_{min}}$**(E(g))** = **2/3*E(g)**.
5. Profit of upgrading/loss from mortgage coefficient: **m** = 0.2.
6. Mean and STD of the Log_2(Income) for immigrants: 12 and **2.2**, respectively**.**
7. Number of dwellings an agent evaluates during residential search: **q** = 30.

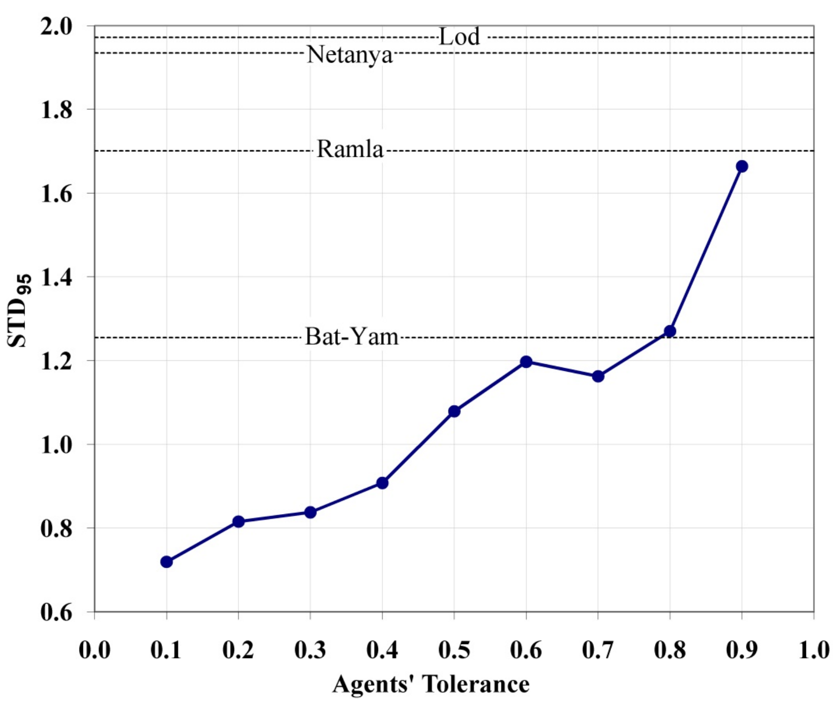


Figure 6: The 95th percentile (STD_{95}) of the distribution of standard deviations of Log_2(Income) over the 3x3 Moore neighborhood for cases in which all agents have the same tolerance.

MODEL INVESTIGATION

A City of Agents with Identical Tolerance

We employ the 95th percentile $\mathbf{STD_{95}}$ of the distribution of the $\mathbf{Log_2(Income)}$ over the Bat-Yam neighborhoods as a criterion for distinguishing between the homogeneous and non-homogeneous patterns. The $\mathbf{STD_{95}}$ for Bat-Yam equals 1.25 Table **1** and in what follows, we consider residential pattern as heterogeneous if $\mathbf{STD_{95} > 1.25}$. The values of $\mathbf{STD_{95}}$ for the patterns obtained in the first series of scenarios, in which all agents have the same value of intolerance of **TL(g)**, are presented in Fig. **6**. As it can be seen, the model residential pattern remains highly segregated until **TL(g)** exceeds the (very high) value of **TL** = 0.8 The corresponding homogeneous pattern obtained for **TL** = 0.5 is presented in Fig. **9a**.

Let us investigate the residential pattern in the city, the population of which contains both intolerant and tolerant agents.

A City Where Tolerant and Intolerant Agents Coexist

Let us recall that our hypothesis is that a low fraction of tolerant agents in the city is sufficient for the emergence of a persistently heterogeneous residential pattern. That is why we are interested in two aspects of the model's residential income distribution. First, we aim at estimating the fraction of tolerant agents that is necessary to generate a heterogeneous pattern; second, we aim at understanding the sensitivity of the pattern's heterogeneity to this fraction. We investigate the hypothesis with the second and third series of scenarios. The second series investigates residential patterns in a city with two-value distributions of tolerance: for some of the agents $\mathbf{TL(g) = TL_{Min}}$, while for the rest **TL(g)** = 0.9. The results, presented in Fig. 7, support our hypothesis that a low fraction of tolerant agents in the city essentially increases the heterogeneity of the persistent model pattern. Even in the scenario with $\mathbf{p_{0.9}}$ = 2.5%, *i.e.*, with only 2.5% of highly tolerant agents, the value of $\mathbf{STD_{95}}$ is 1.25, just as in Bat-Yam while $\mathbf{p_{0.9}}$ = 5% results in $\mathbf{STD_{95}} \approx 1.4$. Moreover, this result does not depend on the value of $\mathbf{TL_{min}}$, which is characteristic of the rest of the agents. The sensitivity of $\mathbf{STD_{95}}$ in regard to the fraction of the highly tolerant agents halts with the growth of $\mathbf{p_{0.9}}$, while for $\mathbf{p_{0.9}}$ > 20%, the value of $\mathbf{STD_{95}}$ remains at a level of 1.5 and hardly varies; even in a scenario with an unrealistically high $\mathbf{p_{0.9}}$ = 50%, $\mathbf{STD_{95}}$ reaches ~1.6 only.

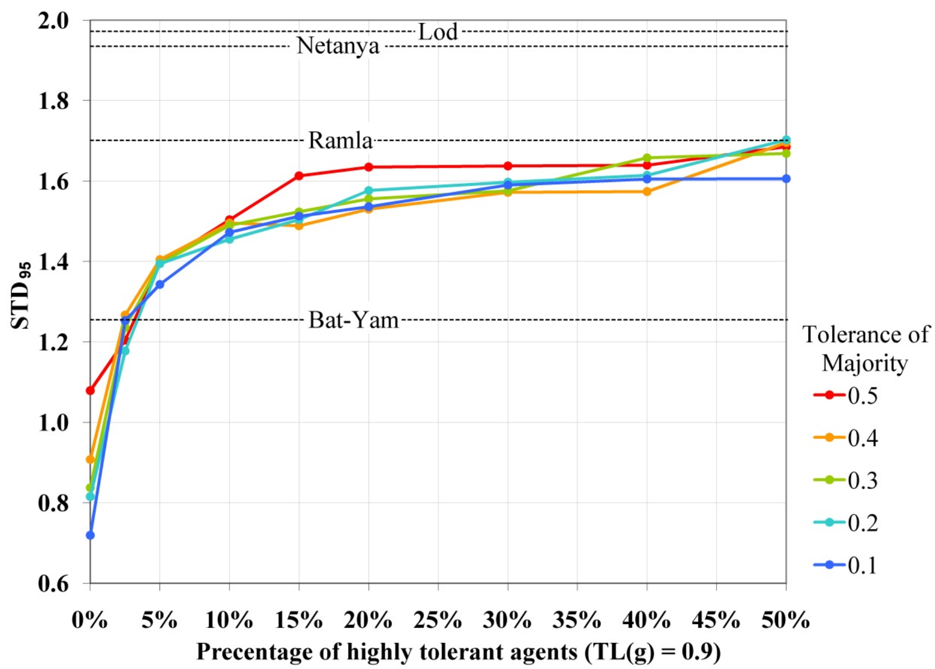


Figure 7: The 95th percentile (STD_{95}) of the distribution of standard deviations of Log_2(Income) over the 3x3 Moore neighborhood for the two-value distribution of tolerance. The scenarios differ according to the percentage of highly tolerant agents (x-axis) and the tolerance of the rest of the agents (different curves).

In the third series of scenarios, the agents' tolerance is distributed uniformly on **[TL_{Min}, 0.9]**, with **TL_{Min}** = 0.5, 0.6, 0.7, 0.8, and 0.9. The values of **STD_{95}** vary between 1.4 – 1.6 (Fig. **8**) and we consider all the patterns obtained in this case as heterogeneous. Note that, just as in the second series, the level of pattern heterogeneity is insensitive to the value of **TL_{Min}**, *i.e.*, the heterogeneity of the persistent urban income pattern, in case of the uniformly distributed tolerance, is similar to that obtained in the case of the two value tolerance in the second series of scenarios. The pattern of the model run with a tolerance distribution of **U[0.1, 0.5]** is presented in Fig. **9d**.

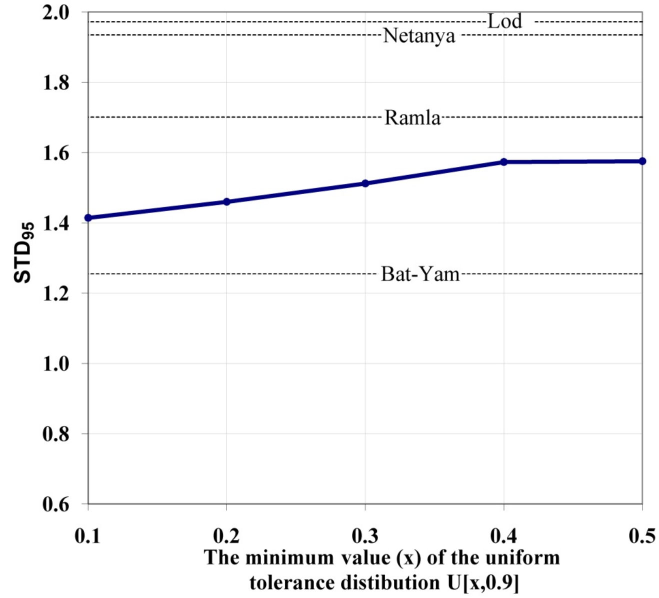


Figure 8: The 95th percentile (STD_{95}) of the distribution of standard deviations of Log_2(Income) over the 3x3 Moore neighborhood for cases of uniform distributions of tolerance.

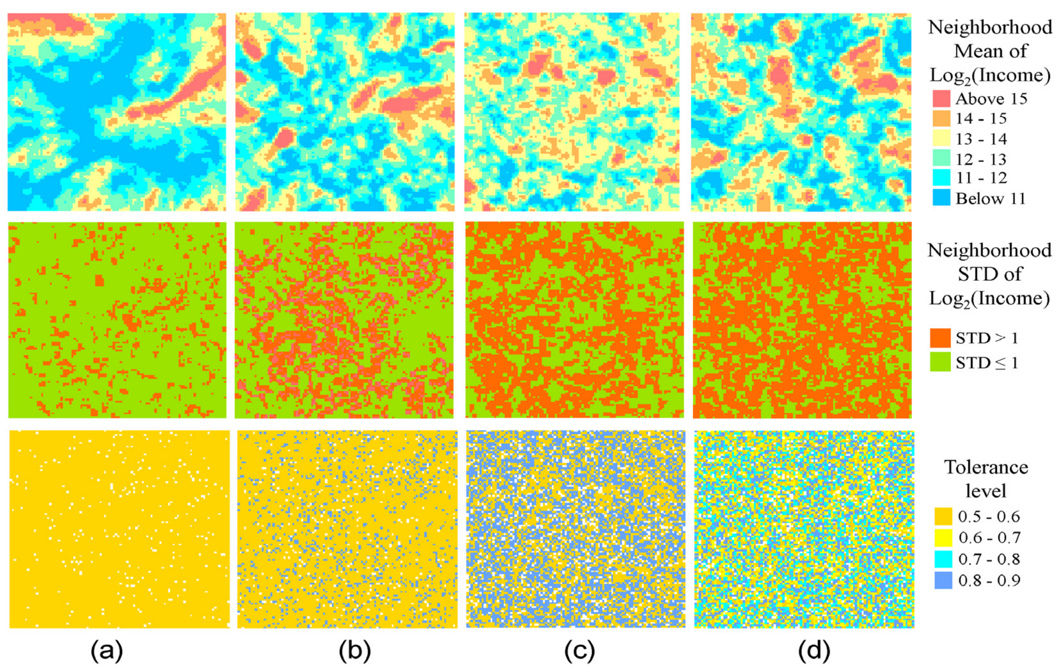


Figure 9: Maps of **Log_2(Income)** average over the 3x3 Moore neighborhood (top), the standard deviations of **Log_2(Income)** over the 3x3 Moore neighborhood (middle) and the tolerance level of agents (bottom) for four distributions of immigrants' tolerance **INT(g:** (a) **TL(g)** = 0.5 for 100% of immigrants; (b) **TL(g)** = 0.5 for 90%,**TL(g)** = 0.9 for 10% of immigrants; (c) **TL(g)** = 0.5 for 50%, **TL(g)** = 0.9 for 50% of immigrants; (d) **TL(g)** of the immigrants is distributed uniformly on **[0.5, 0.9].**

The Location of the Tolerant Agents

The second and third series of scenarios confirm our assumption that a small fraction of the tolerant agents is sufficient for the emergence of the heterogeneous urban income pattern. An interesting question is thus, "where are these tolerant agents located?" The map of agents' tolerance (Fig. **9**, third row) reveals that agents tend to segregate according to tolerance level. This tendency is especially clear when the fraction of tolerant agents is high (Fig. **9c**). The comparison between the average agents' tolerance (Fig. **9**, third row) and the STD of the agents' income (Fig. **9**, second row) reveals that a higher level of income heterogeneity is observed in the urban areas, where the average tolerance of the agents is also higher (Fig. **10**).

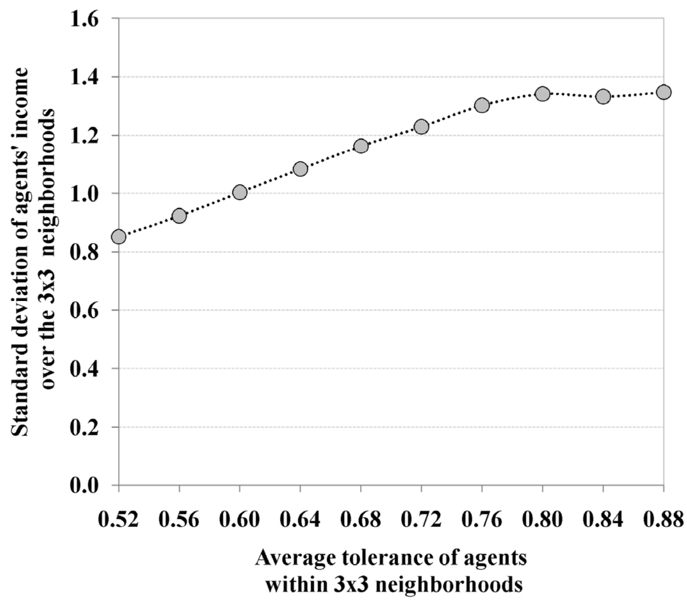


Figure 10: The dependence of the heterogeneity of agents' income in a neighborhood and the average tolerance of the agents for the scenario of 50% tolerant agents (TL(g)=0.9) and 50% intolerant agents (TL(g)=0.5) .

CONCLUSIONS

We started this study by constructing detailed maps of household incomes in nine Israeli cities and demonstrated the essential heterogeneity of the obtained pattern in eight of them. That is, in eight of the nine cities under review, some of the residents are willing to reside close to neighbors of lower economic status. We proposed two possible explanations of this finding. The first is based on the heterogeneity of the urban infrastructure – some wealthy householders locate to new and expensive residential buildings that are constructed in poor areas. The second explanation is related to the personal attitude of the residents to their neighbors: some of the householders, perhaps, the educated ones, are tolerant enough to reside within the poorer neighborhoods. Experimental data support both explanations.

We focused on testing the feasibility of the second explanation, using a model of urban residential dynamics that extends the Schelling model of ethnic segregation towards income-based residential relationships and choice. The model analysis reveals that the presence of tolerant agents essentially increases the spatial heterogeneity of a city. Furthermore, this effect is found to be insensitive to the population fraction of tolerant agents: As long as this fraction is higher than 10%-15%, the city is "heterogeneous". Taking these results literally, one can say that just 10%-15% householders tolerant of poorer neighbors is sufficient to preserve urban residential heterogeneity at a level comparable to that observed in Israeli cities.

Despite the real-world estimates of the distribution of householders' income and migration rates, the level of heterogeneity obtained by the model was lower than the observed level of heterogeneity in most of the investigated cities. This discrepancy can be explained by the fact that the model completely ignores the urban infrastructure, say the experimentally supported heterogeneity of the buildings patterns in respect to their age. Our oversimplified pricing mechanism does not account for the age of the building and, thus, part of the heterogeneity remains beyond the model frame. We also ignore potential variety of residents' behavior: The tolerance of the prospective migrants moving to a city can be different from the tolerance of already established resident moving from one part of the city to another; residents can essentially differ in their estimate of the relocation disutility; the distance to schools and public services can be important for the decision to relocate, *etc.* The Agent-Based modeling approach enables including all potentially important factors, but our knowledge about correlation between the factors and their influence is always limited. Thus we prefer the "Occam razor" view that aims at minimizing the number of factors and studying "major effects" only [14].

Going beyond the quantitative results, tolerant wealthy householders may affect urban pattern beyond the mere increase of heterogeneity. Their movement into lower-class areas may trigger a process of gentrification, as the prices of the properties grow and lower class householders leave or are priced out. This way, spatial heterogeneity of the income pattern may manifest an ongoing process of gentrification. We were not able to test this chain of arguments, because our data are limited to a single year (1995); to verify our result on the temporal persistence of urban heterogeneity, we need high-resolution data on the development of the population patterns in Israeli cities after 1995. Fortunately, the availability of high resolution datasets is steadily on the increase. In this respect, we hope to test the hypothesis raised in this paper on the basis of the data of the on-going Israeli population census, which started in 2008.

REFERENCES

[1] R. E. Park, ‘Secondary Contacts and City Life’, in Introduction to the Science of Sociology, R. E. Park, and E. W. Burgess, Chicago, University of Chicago Press, pp. 311-315, 1921.

[2] E. W. Burgess, ‘The growth of the city: an introduction to a research project’, in Publications of the American Sociology Society, vol. 18, pp. 85-97, 1924.

[3] H. Hoyt, ‘The structure and growth of residential neighborhoods in American cities’, Washington, Government Printing Office, 1939.

[4] B. T. Robson, ‘Urban analysis: A study of city structure’, Cambridge University Press, 1969.

[5] T. C. Schelling, ‘Dynamic models of segregation’, Journal of Mathematical Sociology, vol. 1, pp. 143-186, 1971.

[6] T. C. Schelling, ‘Micromotives and Macrobehavior’, WW Norton, 1978.

[7] M. A. Fossett, ‘Ethnic preferences, social distance dynamics, and residential segregation: Theoretical explorations using simulation analysis’, Mathematical Sociology, vol. 30, no. 3-4, pp. 185 - 273, 2006.

[8] M. A. Fossett, ‘Including preference and social distance dynamics in multi-factor theories of segregation’, Mathematical Sociology, vol. 30, no. 3-4, pp. 289-298, 2006.

[9] W. A. V. Clark, ‘Geography, space, and science: Perspectives from studies of migration and geographical sorting’, Geographical Analysis, vol. 40, pp. 258–275, 2008.

[10] W. A. V. Clark, and M. A. Fossett, ‘Understanding the social context of the Schelling segregation model’, PNAS, vol. 105, no. 11, pp. 4109–4114, 2008.

[11] I. Benenson, and P. M. Torrens, ‘Geosimulation: Automata-based modeling of urban phenomena’, Wiley, 2004.

[12] Israel, Central Bureau of Statistics, ‘Socio-economic characteristics of population and households in localities and statistical areas’, 2000.

[13] I. Benenson, ‘Modeling population dynamics in the city: from a regional to a multi-agent approach’ in Discrete dynamics in nature and society, vol. 3, no 2-3, pp. 149-170, 1999.

[14] I. Benenson, I. Omer, and E. Hatna, ‘Agent-based modeling of householders’ migration behavior and its consequences’ in Agent-based computational demography: Using simulation to improve our understanding of demographic behavior, F.C. Bilary, Prskawetz, A., eds, Springer, pp. 97-116, 2003.

CHAPTER 8

Open-Ended Agent-Based Economic Evolution

Bas Straatman[1,*] and Danielle J. Marceau[2]

[1]Energy and Environmental Systems Group, ISEEE, University of Calgary, Calgary, Alberta, Canada and [2]Department of Geomatics Engineering, University of Calgary, Calgary, Alberta, Canada

Abstract: Agent-based computational economics (ACE) is a recent development and adds to existing economic research by focusing on heterogeneous agents with bounded rationality who act upon incomplete information and who interact locally, instead of the uniform a-spatial fully rational *homo economicus*. The two approaches also differ in their aim. Traditional economics has its origins in equilibrium theory and aims for the stable state, while ACE focuses on the dynamics of the system and the processes that shape it. Instead of modelling to predict the outcome of the system, the focus is on the realism of the paths followed, and it is accepted that outcomes can vary. The ACE model presented in this chapter carries the idea of open-endedness a little further: not only are there multiple possible outcomes of the model, there is also an increase in complexity as the economy continues to produce novel forms (conform Standish's definition of open-endedness [1]). This exemplifies a general trend of complex systems: instead of returning to equilibrium, complex systems, such as the economy, increase their complexity, while novel elements are constantly introduced into the system. This novelty allows an increasingly complex organization of elements, processes and subsystems. To show that open-endedness can be modelled, this paper presents a model that incorporates open-endedness in the form of an agent-based artificial evolving economy. The model results show continuous expansion of the economy by novel products and technologies. This might very well be the first (economic) model to display this kind of evolutionary dynamics.]

INTRODUCTION

Evolutionary Economics (EE) is a novel paradigm that aims at understanding an economic system, defined as individuals, institutions, and activities organized to facilitate the production and exchange of goods and services, as an evolving process in space and time [2]. More specifically, EE focuses on the process of changing economic structures and the strategies of economic actors to adapt to these structures to survive [3]. Among the key concepts and assumptions of EE are the bounded rationality of the economic actors, considered as heterogeneous agents who interact and learn through social and economic networks in a complex and unstable environment, continuously fed by a rich variety of technologies, activities and structures. Consequently, understanding the dynamic interactions between economic actors and economic structures as a feedback mechanism and how novelty emerges and propagates from these interactions is a major area of investigation of EE [3].

To fulfill its objectives, EE adopted interdisciplinary approaches and modelling techniques among which computer simulation became increasingly popular over the last fifteen years. More specifically, the field of agent-based computational economics (ACE) has emerged at the interface of EE, Cognitive Science, and Computer Science [4] and is making headway in economic research. ACE is defined as the computational study of economies modelled as evolving systems of autonomous interacting agents [5]. It considers economic systems as composed of heterogeneous agents with bounded rationality who act upon incomplete information and interact locally in a spatial or social network. While traditional economics has its origins in equilibrium theory and aims for the stable state, ACE originates from the complex adaptive systems paradigm and focuses on the dynamics of the system and its self-organizing property. ACE relies on a bottom-up modelling approach to understand how macroeconomic structures emerge from the local interactions of adaptive agents in decentralized market economies [5].

Agent-based modelling involves a community of heterogeneous agents that interact with each other according to a set of rules. Agents can represent diverse individuals or institutions who play a specific role, can communicate among themselves, are aware of their environment, can learn and adapt to new situations, and can design strategies to achieve

***Address correspondence to Bas Straatman:** Energy and Environmental Systems Group, ISEEE, University of Calgary, 2500 University Dr NW, Calgary, AB, T2N 1N4, Canada; Tel: (403) 210-7621; Fax: (403) 210-3894; E-mail: b.straatman@mun.ca

Danielle J. Marceau and Itzhak Benenson (Eds)

their goals [6]. The aim of agent-based modelling is to explore the possible paths a complex, adaptive system can take through the dynamic simulation of agents' behavior and interactions [7], and as such as be proved to be a powerful tool for research in Economics (see the comprehensive reviews in [4, 8, 9]). ABM allow for the creation of virtual economic environments in which diverse scenarios can be tested. Compared to conventional equilibrium models, these models can take into account nonlinear behavior associated to human adaptation and learning. They incorporate rule-based decision heuristics and go beyond the perspective of fully rational and informed economic actors.

The independent studies of Schelling [10] and Sakoda [11] are often referred to as the first real applications of ABM in social sciences and economics. Their models aim at understanding households' migration behavior and cultural segregation using a population of interacting human agents. This topic is still a fertile field of investigation as illustrated by the recent study of Hatna and Benenson on income-based urban residential patterns [12, in this book]. Other active areas of ACE research include the formation of economic networks and organizations, the evolution of behavioral norms, and the bottom-up modelling of market processes to understand the self-organizing capabilities of specific types of market, such as financial, labour, retail, natural resource, and Internet exchange systems [5, 13]. Doyne Farmer and Foley [14] and Buchanan [15] report that this modelling approach has been successfully applied to explore how leverage affects fluctuations in stock prices, to optimize the flow of goods of a large consumer-products company through its network of suppliers, warehouses and stores, and by the NASDAQ stock exchange in New York to reproduce price fluctuations.

Among the variety of complex issues investigated by ACE is the open-ended evolution of individual behaviors and economic institutions [13]. In other words: given the appropriate conditions, to what extent an economic system might evolve and generate more advanced structures? An open-ended system is a system in which new adaptive traits continue to evolve over prolonged periods, without encountering a stopping point, and where there is a demonstrable sustained increased in the maximum complexity of this system. It is the continuing generation by evolutionary systems of novel characteristics [16]. Formulated within the context of the arrow-of-complexity hypothesis, this means that the functional complexity of an open-ended evolutionary system has a general tendency to increase with time [17].

The main idea advocated in this paper is when studying the evolution of economic systems not only are there multiple possible outcomes, but instead of returning to equilibrium the complexity of these systems seems to increase over the long run. Economies have a tendency to expand and become more complex, with the diversity of technology and products continuously increasing [18]. This statement is in accordance with Prigogine's concept of dissipative emergent structures [19]. According to this concept, a system if left alone will return to a trivial state; however when the system is open to inputs from outside, energy and materials are available and the surplus dissipates through increasingly complex self-emergent structures or organizations. Kauffman [18] further suggests that somehow complex systems move into the adjacent possible path fast enough not to become static, and not too fast in order to be able to avoid total chaos. This implies that complex systems can be truly open-ended as novel elements are constantly introduced into the system. The novelty allows for an increasingly complex organization of elements, processes and subsystems.

This study was undertaken to investigate if the economic market acts as an open-ended system using an agent-based approach. Two novel aspects are introduced. First, an agent-based economic model, called constructive is designed. Constructive dynamical systems arise from the field of Artificial Chemistry (AC). In attempts to decipher the origin and evolution of life, represented as a complex system, AC deals with elements that combine together, change or maintain themselves, and with systems that are able to construct new components [20]. Models that enable the construction of new component and new relationships among them, which may change the dynamics of the system, are referred to as constructive dynamic models. They differ from a conventional dynamic model where all the components and their interactions are given at the outset of the process [20]. Standard ABMs belong to that category as they do not allow for true evolutionary behavior. In this paper, a constructive agent-based economic model is designed to simulate an integrated production and market model in which successful and unsuccessful innovations are introduced, which serve as a selection mechanism in an evolutionary process.

The second novel aspect that is introduced is the use of a measure of evolutionary activity proposed by Bedau, Snyder and Packard [21] who classified long-term evolutionary dynamics into three categories: absent, bounded, and unbounded. This measure is used to evaluate the impact of the addition of new technology in the economic system being simulated, not only in terms of increasing its diversity, but in terms of generating a higher output.

The remaining of this paper is organized as follow. First a description of the Bedau's measure of evolutionary dynamics is presented. It is followed by a presentation of the constructive agent-based model implemented to investigate if the economic market acts as an open-ended system. A presentation of the results and a conclusion complete the paper.

A MEASURE FOR EVOLUTIONARY DYNAMICS

Bedau, McCaskill, Packard *et al.* [22] gives a list of open problems in artificial life research, and one of the issues stated is: ``What is inevitable in open-ended evolution of life?". In researching this question, many attempts have been made to simulate open-ended evolution, but so far even the most promising artificial life models such as Tierra and Avida have not succeeded. Therefore, Standish [1] reformulates the above: ``The issue of open-ended evolution can be summed up by asking under what conditions will an evolutionary system continue to produce novel forms". This leads to two questions, for both of which the answer is given in this paper. The first question is what does it mean to continue to produce new forms? That question is answered in this section by looking at Bedau's measure for evolutionary dynamics. The second question is how to build a model that can actually achieve the continuous generation of novelty. That question is answered in the next section.

The only absolutely certain case of evolution and increasing complexity is our own biosphere. In order to be able to compare dynamics generated by models with the evolutionary dynamics of our biosphere, Bedau, Snyder and Packard developed a measure to classify evolutionary dynamics, in other words to identify innovations that make a difference, where they consider an innovation to ``make a difference" if it persists and continues to be used [21, p.229]. Evolutionary activity is measured by diversity of activity, new activity and average aggregate activity. Based on these indicators, [21] distinguishes three classes of evolutionary dynamics: non-evolutionary, bounded, and unbounded Table **1**.

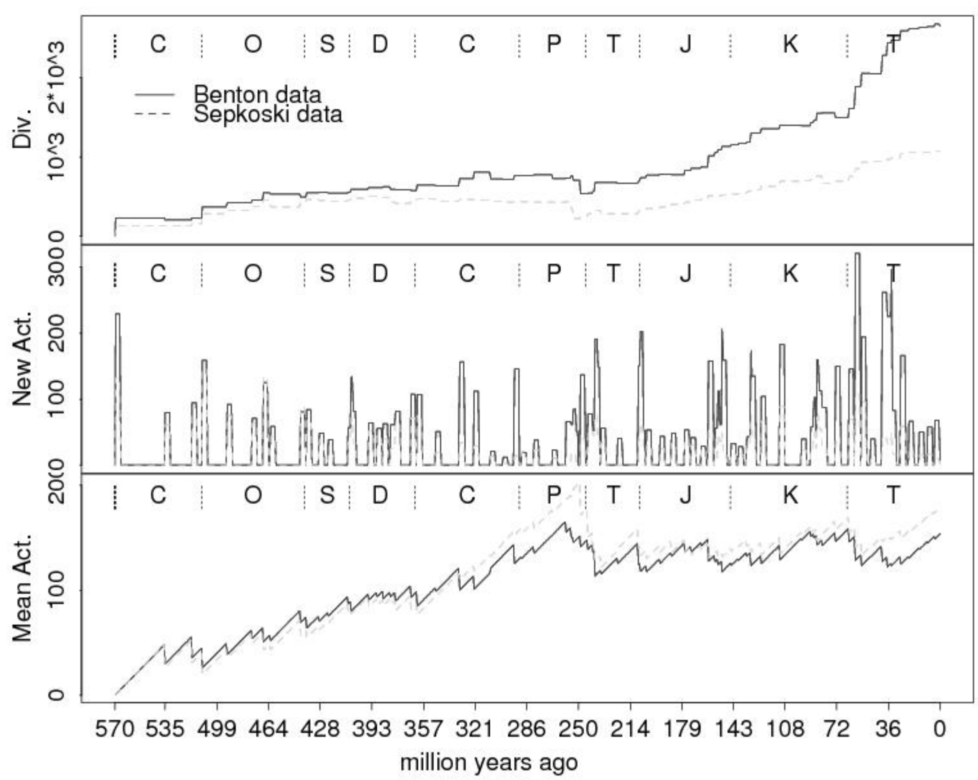


Figure 1: Analysis of the evolutionary dynamics of life as recorded in the fossil record data. Bedau, Snyder and Packard computed the diversity, the new activity and the mean activity in the Benton and Sepkoski fossil data sets on the existence of taxonomic families. The labels at the top of each graph show the boundaries between the standard geological periods, thus: Cambrian, Ordovician, Silurian, Devonian, Carboniferous, Permian, Triassic, Jurassic, Cretaceous, Tertiary. Image copied from [21].

Table 1: Classification of the evolutionary dynamics of artificial and natural evolving systems. Some models of artificial life show no evolutionary activity (class one), while some display bounded (class two) and others unbounded (class three) evolutionary activity [21]. In the headings D stands for Diversity, A_{new} stands for new activity, and $\overline{A}_{cum}$ stands for average aggregate activity.

class	evol. activity	D	A_{new}	$\overline{A}_{cum}$
1	none	bounded	zero	zero
2	bounded	bounded	positive	bounded
3	unbounded	unbounded	positive	bounded

The indicators D, A_{new}, and $\overline{A}_{cum}$ of Table **1** were determined for two data sets of fossil records, the Benton data set and the Sepkoski data set. These data sets record the taxonomic families. The Benton data set contains the fossil record of all families in all kingdoms, the Sepkoski data set contains the families of marine animals. This resulted in the graphs of Fig. **1**. The top graph shows the increase in diversity over time, where diversity is the number of taxonomic families. New activity (middle graph) indicates the appearance of new families, and mean activity (bottom graph) is aggregate activity divided by the number of families present at the time. Since diversity is increasing and mean activity is bounded, the aggregate activity is increasing unboundedly, and thus the evolutionary dynamics of the Phanerozoic biosphere as recorded in the fossil data sets belongs to class three: unbounded evolutionary activity.

AN EVOLVING AGENT-BASED MARKET MODEL

In the approach presented here, an agent-based model is developed to simulate an integrated production and market model. The market's role is to rule on successful and unsuccessful innovations *i.e.* it serves as a selection mechanism in an evolutionary process.

The technology of the model economy is represented by von Neumann input and output matrices [23]. These matrices were introduced in the context of a model developed to determine the conditions under which a general economic equilibrium exists. Here we are not concerned with von Neumann's conditions and the resulting equilibrium, but we use his matrix representation of the structure of the economy because it provides a compact description of the relation between technology and products. Each column in the pair of matrices represents a production process, or technology; more specifically, the columns in the input matrix show the inputs required to make each product, while the columns of the output matrix show the output generated by each of the technologies.

Let a and b be the input and output matrices respectively, with dimensions of both matrices $n \times m$, implying that there are m production processes and n products, and let $z(t)$ be the intensity or activity vector at time t. Von Neumann shows that under certain conditions there exist a growth rate α and an interest rate β, such that the output at time t, $b \cdot z(t)$ is sufficient to cover the input required at time $t+1$, $a \cdot z(t+1) = a \cdot \alpha z(t)$, and the profitability of each of the production processes is smaller than or equal to the interest rate.

$$a \cdot z(t+1) = a \cdot \alpha z(t) \leq b \cdot z(t)$$

$$\beta \cdot p \cdot a \geq p \cdot b$$

If we denote the total product inventory at time t as $S(t)$, then

$$S(t+1) = S(t) + (b-a) \cdot z(t).$$

AGENT-BASED MODELLING OF ECONOMIC ACTIVITY

In this approach we generate activity bottom up through the use of agents representing producers. Since prices and quantities are not determined algebraically but by agents interacting with each other repeatedly, the agent-based model is inherently dynamic.

Again we define a and b as the input and output matrices with m processes and n products, but now we add r agents to the model. Some of these agents will be producers, others will be consumers, depending on the technological capabilities they possess. In particular, consumers use consumables as inputs, and as output they generate labour.

Each of the agents A_i, regardless of type, is defined by

$$A_i(t) = A_i(f_i, I_{ij}, T_i, S_i(t)) \;\; for \;\; i = 1, \dots r$$

where f_i is the activity of agent A_i, I_{ij} is the interaction matrix specifying the agents with whom A_i interacts and T_i specifies which technologies agent A_i possesses; *i.e.* T_i specifies certain columns of the input and output matrices a and b. Interactions with other agents to acquire inputs and sell outputs, and production activity, which depends on technology T_i, operate on $S_i(t)$, which is the product set in agent A_i's possession at time t. In this section we assume agents can interact with all other agents, thus $I_{ij} = 1 \forall i, j$. During one iteration of the model, agents act sequentially to execute their activity, and the order in which the agents act is randomized at each iteration.

We now expand the matrices a and b by joining them with matrices M and I:

$$(\;\; manufacturing \;\; | \;\; buying \;\; | \;\; selling \;\;)$$

$$i = (\;\; a \;\; | \;\; M \;\; | \;\; I \;\;)$$

$$o = (\;\; b \;\; | \;\; I \;\; | \;\; M \;\;)$$

where I is the identity matrix of size n, and M is the square matrix of size n with all elements equal to 0 except for the second row which is equal to the price vector $p = \{p_1, p_2, \dots p_n\}$ -- in other words, the good corresponding to the second row of a and b is the numeraire, or simply money, and so the second row of M represents the prices of all goods. These expanded matrices link the input and output operations of manufacturing with the associated buying and selling of the goods.

The activity $f_i(t)$ of the agent A_i is given by

$$\max_{z_i \in \mathbb{Q}^{m+2n}} (p(t) \cdot (o - i) - (\beta, \dots, \beta, 0, \dots, 0)) \cdot z_i$$

subject to the conditions $\mathbf{B}, \mathbf{M}, \mathbf{T}, \mathbf{O}, \mathbf{F},$ *and* $\mathbf{E}$ listed below, and where $p(t)$ is the vector consisting of the prices $p_i(t)$ for each of the products $i = 1, \dots, n$ at time t and z_i is a vector of length $m + 2n$ indicating the actions (manufacturing, buying, selling) of agent i at time t. Prices are determined anew for every agent as it takes its turn in order to include the previous agent's effect on the balance of supply and demand. The term $p(t) \cdot (o - i)$ is a vector of profit rates for each of the actions (manufacturing, buying, selling) an agent can execute, and this vector is compared with the standard rate of return for each of these actions: β for manufacturing, and 0 for buying or selling at the current price.

If we write $z = \{m_1, \dots m_m, b_1, \dots b_n, s_1, \dots s_n\}$ to break up the activity vector into the different parts of manufacturing (m), buying (b) and selling (s), then agent A_i maximizes the value of the outcome of its actions z_i under the following constraints:

The state of agent i at time t minus expenditures plus output has to remain non-negative at all times:

$$S_i(t) + (o - i) \cdot z_i \geq 0$$

An agent is limited in how much it can do in one turn:

$$\sum_{i=1}^{m} z_i \leq maxproductivity$$

An agent cannot perform actions for which it lacks the technology, and at the same time it cannot perform each individual action more than $maxproductivity$ times:

$$\{m_1, \ldots, m_m\} \leq maxproductivity \cdot \{t_1, \ldots t_m\}$$

where $T_i = \{t_1, \ldots t_m\}$ is the technology list of agent i, which is a Boolean vector with 1s and 0s for technology it can and cannot perform.

How much an agent can buy depends on the quantities offered:

$$\{b_1, \ldots, b_n\} \leq \{\max_{i=1,\ldots,r} S_{i,1}, \ldots, \max_{i=1,\ldots,r} S_{i,n}\}$$

How much an agent can sell depends on the quantities in its possession:

$$\{s_1, \ldots, s_n\} \leq S_i(t)$$

For every consumer the expenditures on consumables $m_{l_1}, m_{l_2}, \ldots m_{l_n}$ must generate at least c units of labour:

$$\{m_{l_1}, m_{l_2}, \ldots m_{l_n}\} \cdot \{o_{l_1}, o_{l_2}, \ldots o_{l_n}\} \geq c$$

These consumables are removed from the system without resulting in labour.

The last condition **E** prescribes forced, non-productive consumption and prevents the model from declining into the trivial equilibrium null state of no production equaling no demand. It implies that a portion of the consumption happens regardless of the labour market being in short supply or saturated. All other activities are motivated by excess demand.

With the actions of agent A_i defined as above, its state S_i is updated as follows:

$$S_i(t + \Delta_i) = S_i(t) + (o - i) \cdot z_i(t)$$

The Δ_i indicates that the system is in an intermediate state between $S(t)$ and $S(t+1)$. Once all agents have had their turn, time t is increased to $t+1$. Thus $\sum_{i=1}^{r} \Delta_i = 1$.

In the execution of $z_i(t)$ the state of other agents is possibly affected as well, since they serve as suppliers or buyers and thus resources have to be exchanged. Appropriate agents (who have the required products or cash) are selected randomly and the exchanges are made. This implies that when purchase or sale z_k involves agent $A_{j(k)}$,

$$S_{j(k)}(t + \Delta_i)$$

$$= S_{j(k)}(t) - (o-i) \cdot \{0, \ldots, 0, z_k, 0, \ldots, 0\}$$

$$\forall k \in (m+1, \ldots, m+2n)$$

A complete update U_i of the state of the system as a result of the actions of agent A_i thus consists of the following $2n+1$ instantaneous updates:

$$\begin{cases} S_i(t+\Delta_i) & = & S_i(t) + (o-i) \cdot z_i(t) \\ S_{j(m+1)}(t+\Delta_i) & = & S_{j(m+1)}(t) - (o-i) \cdot \\ & & \{0, \ldots, 0, z_{m+1}, 0, \ldots, 0\} \\ & \vdots & \\ S_{j(m+2n)}(t+\Delta_i) & = & S_{j(m+2n)}(t) - (o-i) \cdot \\ & & \{0, \ldots, 0, 0, \ldots, 0, z_{m+2n}\} \end{cases}$$

where $S(t) = (S_1(t), S_2(t), \ldots, S_r(t))$, (see Fig. **2**). One complete iteration consists of this update sequence applied to each of the agents. For that purpose the r agents are put in a randomly ordered list $r_1, r_2, \ldots, r_r$ as Fig. **3** illustrates, and

$$S(t+1) = U_{r_r} \circ U_{r_{r-1}} \circ \ldots \circ U_{r_2} \circ U_{r_1}(S(t)).$$

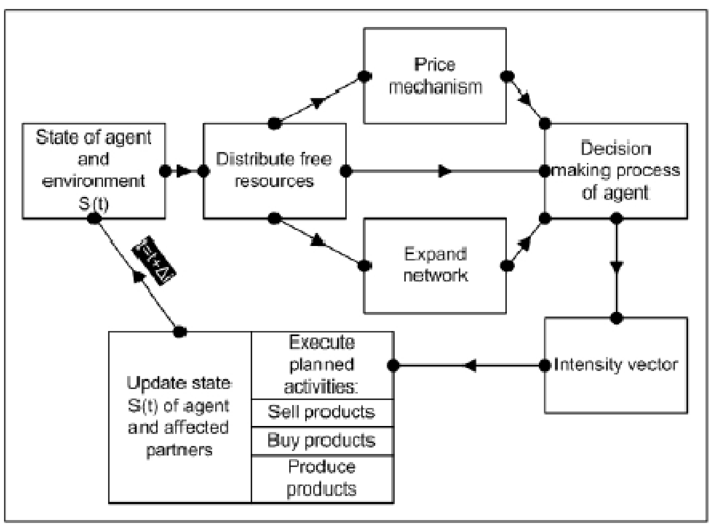


Figure 2: Activity cycle for one agent cycle. The cycle starts in the top left corner with state $S(t)$ and in particular state $S_i(t)$. After a number of steps the system is updated to $S(t+\Delta_i)$ and the cycle is repeated but with another agent j.

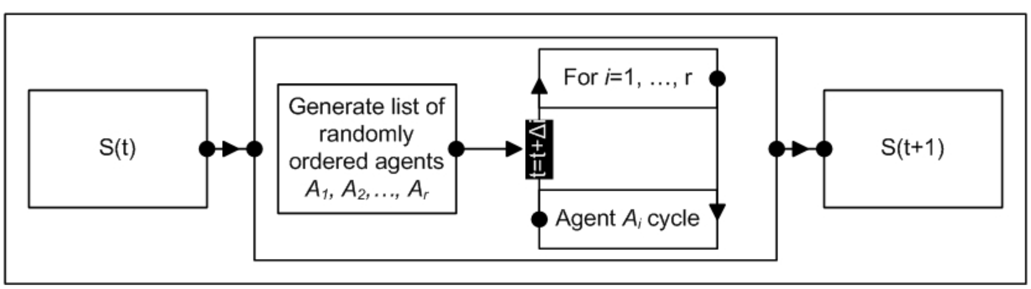


Figure 3: One iteration cycle. When all agents have had their turn the system has been updated to the next time step.

Furthermore,

$$z(t) = \sum_{i=1}^{r} z_i(t)$$

and

$$S(t+1) = S(t) + (o-i) \cdot z(t).$$

THE PRICE MECHANISM

The above algorithm defines one iteration. With a history of w iterations it is known which products and how much of each have been demanded during the previous w iterations $D(w,t)$, and also the current stock $\Sigma(t)$ is known. The price mechanism takes into account the balance for each of the n products:

$$B = (bal_1, bal_2, \ldots, bal_n) = \Sigma(t) - D(w,t)$$

Von Neumann's work shows there exists an equilibrium price p_{eq} that provides each production process with an equal return rate, β

$$\left(\frac{(p_{eq} \cdot (b-a))_1}{(p_{eq} \cdot a)_1}, \ldots, \frac{(p_{eq} \cdot (b-a))_m}{(p_{eq} \cdot a)_m} \right) = (\beta, \ldots, \beta)$$

where $(.)_j$ stands for the j-th element of the vector $(.)$. The mechanism to determine appropriate prices $p = \{p_1, \ldots, p_n\}$ is a linear programming procedure that minimizes the sum of the prices:

$$\min \sum_{i=1,\ldots,n} p_i \text{ under the conditions } \mathbf{P} \text{ and } \mathbf{R}$$

where

Prices are always greater than or equal to the equilibrium price:

$$p \geq p_{eq}$$

Let m_p be the set of production processes producing product i *for which* $bal_i \leq 0$ and let m_l be the set of production processes producing product j *for which* $bal_j \geq 0$. m_p is the set of manufacturing processes which result in a product of which there currently is a shortage, and m_l is the set of manufacturing processes which result in a product of which there currently is a surplus. In order to balance the market, processes in m_p should produce with a profit, *i.e.* a return rate higher than the interest rate β and processes in m_l should produce at a loss, *i.e.* a return rate lower than the interest rate. Condition R states that for each $j \in m_p$:

$$\frac{(p \cdot (b-a))_j}{(p \cdot a)_j} - \beta > 0$$

and for each $j \in m_l$

$$\frac{(p \cdot (b-a))_j}{(p \cdot a)_j} - \beta \leq 0$$

IMPERFECT KNOWLEDGE

In the model described above every agent could interact with any other agent in the system. A more realistic situation, however, is that agents can potentially interact with only a limited number of others. At any particular time some agents are not potential economic partners either because they are not known to the agent looking for suppliers or customers, or because they are inaccessible. This constraint of imperfect knowledge or access is implemented in the model through a network, to indicate which locations are connected with one another. In the structure used above to represent the agents in the model this can be represented in the connectivity matrix I_{ij} by setting some elements equal to zero to indicate that no interaction is possible between agents i and j.

When an agent does not have access to all of the products required for its technology, it will attempt to add a connection to its network such that formerly unavailable commodities become available. The implementation of such a process can follow different rules, for example:

- An agent can add the nearest supplier;
- An agent can add a supplier at random;
- An agent can ask an already existing contact to check its neighbourhood to find a supplier and if one is available this supplier can be added to the network.

Whichever process is used, to add agent j to the network of agent i the element j,i of I is updated by changing the 0 to 1.

The introduction of limited or local interaction results in local rather than global prices by limiting the balance

$$B_i = (bal_1, bal_2, \ldots, bal_n) = \Sigma(t,i) - D(w,t,i)$$

to the local demand $D(w,i)$ and local current stock $\Sigma(t,i)$ in the network of agent i. No other changes to the model are required.

CONSUMER BEHAVIOUR

A different model of consumer behaviour can be obtained by removing condition **E** above. In order to prevent the model from converging to the trivial stable 0-state, f_i is changed:

$$f_i(t) = \max_{z_i \in \mathbb{Q}^{m+2n}} g(t) \cdot z_i$$

subject to the conditions $\mathbf{B}, \mathbf{C}, \mathbf{T}, \mathbf{M}, \mathbf{O}$ *and* $\mathbf{F}$ where $g(t) = p(t) \cdot (o-i) - \beta$ with some of its elements replaced. The goal function $g(t)$ provides agents with an indication of the profitability (return rate higher than the interest rate) of each of the production processes. In case $p(t) = p_{eq}$, the return rate for each of the processes is equal to β, and thus none of the processes results in a profit higher than 0. As explained above, the price mechanism sets prices such that the appropriate elements of $g(t)$ are positive, and the remainder of the elements are less than or equal to 0. In order to drive the economy without exogenous consumption, the elements of $g(t)$ corresponding to consumption are replaced by positive numbers, disregarding the actual profits (which are based on the prices and ultimately shortage or surplus of labour) for these elements. This results in a permanent positive stimulus to consume. However, the intensity of the consumption depends on the available means, and thus the rate of consumption is not fixed. Either for a balanced labour supply or one that is unbalanced, consumers are driven to consume. Variations in the size of the inserted positive numbers allows the modelling of consumer preferences.

NEW OPPORTUNITIES AND NICHES

Generation of novelty into the model consists of two steps. First the technology matrices have to be expanded with columns or rows to add new technology or new products to the economy. Subsequently the technology or the product has to be applied by the agents.

EVOLVING TECHNOLOGY

For the implementation of evolving technology the matrices a,b , and thus i,o , are no longer fixed in time. In this case the matrices vary with time. This idea is not new; over the years it has been mentioned a number of times [24, 25, 26], but to the authors' knowledge it has never been implemented.

The idea of innovation used here is based on two concepts. The first one is the concept of preadaptation [18], where an existing object obtains a new trait that is not really a new trait but that becomes one in a changing context. The second concept is a novel combination of existing objects to create a new one.

A new technology can be added to the technology set by generating a new column in which the new transformation makes use of existing products. In addition to new technology based on existing products, it is also possible to introduce new columns that make use of new products. In such a system, the new product can have three functions: it can be an intermediate product to be used as a building block for more advanced products, it can be a consumable, which is thus destined for consumption and transformed into labour, or thirdly, it can be a product that serves as capital.

As in Schumpeter [27] and Mosekilde and Rasmussen [28], new technology is generated and introduced at random times. With a certain probability this new technology involves the creation of new consumables. If not a consumable, the new technology aims at producing products involved in the existing production process, be it capital or intermediates. In order to make something from which the system can benefit, we introduce new technology (*i.e.* a tool and the skill to use it) to produce something of which there is currently a shortage, in the same way an alternative in the form of coal was found for the wood that was in short supply. Surplus and shortage are defined as in the price mechanism.

The innovation algorithm pseudo code is given in Fig. **4**. The parameters q_1, q_2 and q_3 control the probabilities of innovation and the type of innovation.

```
if random ≤ q1 innovation:
  if random ≤ q2
    then generate new consumable:
         combine random products
         set the quantity of labour resulting from
         consumption of this good
    else generate new technology:
         select product that is in excess demand
         create a tool t
         if random ≤ q3
           then new tool requires new product:
                find list of products with
                appropriate factorization
                multiply these products to create
                new product
           else new tool t uses old products:
                find list of products with
                appropriate factorization
```

Figure 4: The innovation algorithm in pseudo code. *random* is a uniform random variable in the (0, 1) interval. q_1, q_2 and q_3 are parameters to control the probabilities of innovation and the type of innovation.

Suppose that *random* $\leq q_1$ such that new technology will be generated. The case of new consumables is explained in the next section. Once one of the products in short supply is identified, the input for the new tool t can be composed. More precisely, when the product in excess demand ped is factorized $ped = f_1^{q_1} \cdot f_2^{q_2} \cdot \ldots \cdot f_n^{q_n}$ for some n and $q_. \in \mathbb{N}$, it is clear which factors are required in the input. A random list of products $i_1, \ldots, i_m$ that contain those factors is generated such that the product of these inputs $i = i_1 \cdot i_2 \cdot \ldots \cdot i_m$ will be divisible by the sought after product ped, thus $\frac{i}{ped} = g \in \mathbb{N}$. There are then two possibilities: a draw of a uniform random variable is less than q_3 and the tool t sets the production of the required product ped, based on the new input i, or the uniform random variable is larger than q_3 and the tool requires the different parts $i_1 \cdot i_2 \cdot \ldots \cdot i_m$ and uses these without product i first being assembled. g is considered garbage and ignored.

DIFFUSION OF TECHNOLOGY

With the expansion of the technology matrices as in the previous section, new products and new skills are not yet introduced into the economy. The new products only appear when agents use the new skills and the new products. This section describes the process of the introduction and the diffusion of skills. The model design of this process is based on work by Bruckner, Ebeling and Scharnhorst [29].

Sorting through Bruckner's list of innovation processes, we obtain a list of five relevant types of events that are applicable here: innovation by a new agent, innovation by an existing agent, random imitation of technology, spawning of a new agent with existing technology, and imitation of successful technology by an existing agent. These events can be interpreted in such a way that they fit within the general framework provided by [29].

When a new technology becomes available, its establishment is affected by some of the existing technologies. Parameter A_{ij} describes the inclination of technology j to establish technology i by means of a new agent: $W(N_i + 1, N_j \mid N_i = 0, N_j) = A_{ij} N_j$

When an agent expands its skills, be it with a new technology or an existing technology, the agent innovates its production process. The choice of additional technology is affected by the existing technologies. Parameter M_{ij} describes the inclination of technology j to establish technology i by means of an existing agent: $W(N_i + 1, N_j \mid N_j) = M_{ij} N_j$

An increase in the number of agents using technology i independent of the state of the system: $W(N_i + 1) = \phi_i$

An increase in the number of agents using technology i due to self-reproduction or sponsoring by technology j: $W(N_i + 1, N_j \mid N_i, N_j) = A_i^{(0)} N_i + A_i^{(1)} N_i N_i + B_{ij} N_i N_j$

Agents imitate the successful technologies of other agents, and thus replace less successful skills. The parameters A_{ij} represents a measure of success and failure. Furthermore, the probability is influenced by the current size of the the technology fields. The larger the field, the more occurrences of the technology in question, and the greater the probability of replacement: $W(N_i + 1, N_j - 1 \mid N_i, N_j) = A_{ij}^{(0)} N_j + A_{ij}^{(1)} N_i N_j$ where N_i is the number of occurrences of technology i. The probabilities are calculated based on parameters $A_{ij}, M_{ij}, \phi_i, A_i^{(0)}, A_i^{(1)}, B_{ij}, A_{ij}^{(0)}, A_{ij}^{(1)}$, which are functions of measures for success, occupancy, and connections to other sectors. See also Bruckener, Ebeling, Jiménez Montana and Scharnhorst [30] for the definition of parameters.

THE COMPLETE MODEL

With the expanding matrices algorithm in place, the model can simulate the impact of the introduction of new technology and new products. In order to do so the different processes that have been described in this section are placed in a loop depicted in Fig. **5**. The agent cycle is repeated again until all agents have had their turn. After a random number of repetitions of the iteration-technology diffusion cycle, the matrices are expanded by generating new products and technology, after which the iteration and technology diffusion cycle is repeated again. This process simulates the introduction of new technology and products into a free-market economy, where the success and failure of the products and the technology is determined by the combined action of many local actors.

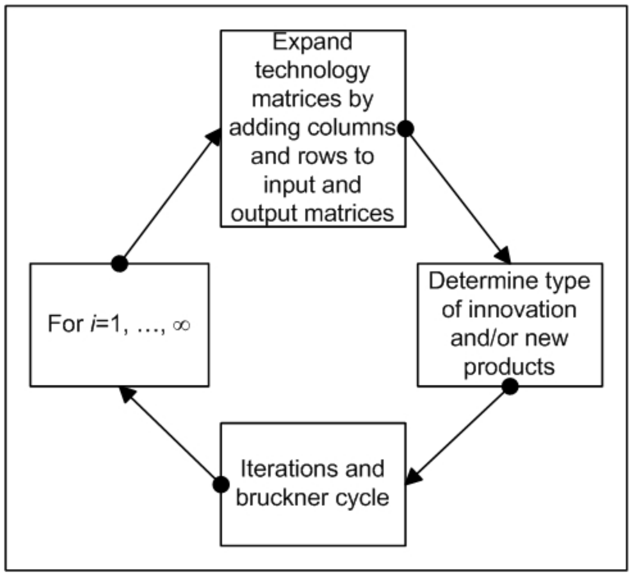


Figure 5: The full cycle of an evolving economy shows how all of the above cycles are combined into a routine that can be run indefinitely if time would allow.

THE APPLICATION OF THE BEDAU MEASURE

When one wants to apply the Bedau measure to assess the evolutionary qualities of a data set, two issues need to be resolved: what are the components that make up the innovations, and how should the evolutionary activity of a component be measured [21]? For the model in this chapter, an innovation consists of a new technology that appears as a new column of the von Neumann technology matrix. The second question to be answered is how to measure the contribution of a technology to the system. For the analysis of the fossil data sets [21] define the evolutionary activity of a component at time t by its age at time t, given that the component still exists at that time; and aggregate activity then is defined as the sum of evolutionary activity over all components. Indeed such a measure indicates the persistence of the component itself, as well as its importance to the subsequent components that make use of it.

For our technologies something similar was done. A technology is persistent if it is still present, and if it supports subsequent (or any other) technology, meaning that there is demand for its products. A technology is active because the market is conducive, which is caused by other components. The higher its output, the more the technology contributes to the system in terms of creating opportunities for other (old and new) technologies. We have defined the component activity of a technology at time t by the value of output generated by the technology during iteration t and not by the summation of activity from 0 to t. The aggregate activity at time t is the total output generated at time t by all technologies, the diversity at time t is the number of active technologies at time t and the mean activity is total output divided by the diversity. The results of this application of the Bedau measure is discussed in Section 5.

RESULTS

Fig. **6** displays the three Bedau indicators for a model run in which there is an increasing number of agents due to the possibility of the introduction of new agents. It shows an increasing diversity, while the added technologies result in a higher mean activity. That is, the output per unit of applied technology is increasing. This suggests that the application of additional technology has a positive effect on the production system. Thus, not just an increase in diversity because additional technology appears, but instead the increased diversity results in a higher output of the system. According to the Bedau indicators, this is the essence of evolutionary dynamics: the system generates new activities, and as the diversity increases, the new activities do not simply replace the existing ones, they have a positive effect on the system by increasing its output.

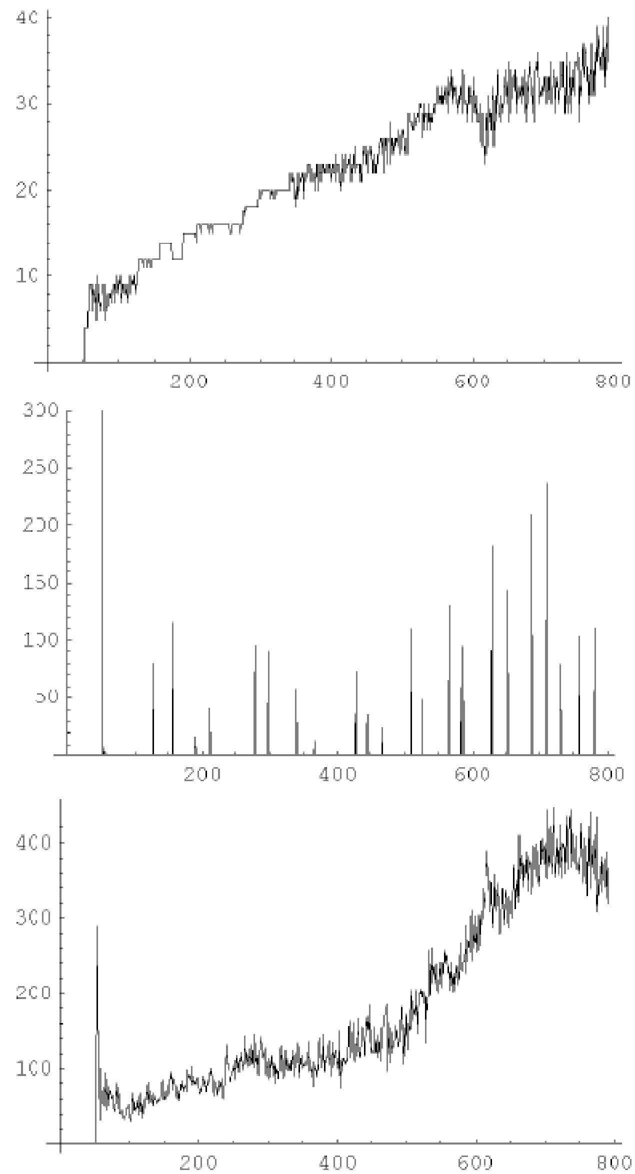


Figure 6: Bedau's measures applied to the economic activity of an agent-based model with the possibility of immigrants supporting the successful skills. The top graph displays the number of active technologies, the graph in the middle shows the activity of new technology, and the bottom graph displays the average activity per active technology. The x-axes shows the iteration number as the simulation progresses, the y-axes shows the number of active technologies (top graph), the value of generated output at the time of the first application of new technology (middle graph), and the generated value of the output per active technology (bottom graph).

Fig. **7** shows the results of a simulation in which no new agents are introduced to the system, and thus only the existing agents are capable of imitating successful skills. This is an important difference with the previously described simulation, because it prevents the model from having an infinite capacity to innovate. Innovations appear, but at the cost of existing technology, since the limited agent set has a limited capacity to apply technology. As a result technology is replaced and forgotten, not because of some prescribed extinction rate, but because of poorer performance. Regardless of the disappearance of some technology, the simulation shows that at random intervals there are signs of new activity; this new activity leads to an increase in diversity, and the increasing diversity results in a higher output. The results show that an old, successful technology, once applied by numerous agents can fade, and is replaced by a more advanced technology, with higher capacity, or better suited to the time, in that it requires inputs and produces outputs that are more in line with the demand. Here, the first signs of creative destruction are clearly visible, although calling it a gale is still a little far fetched. The model provides an appropriate platform for simulating open-ended evolving economies.

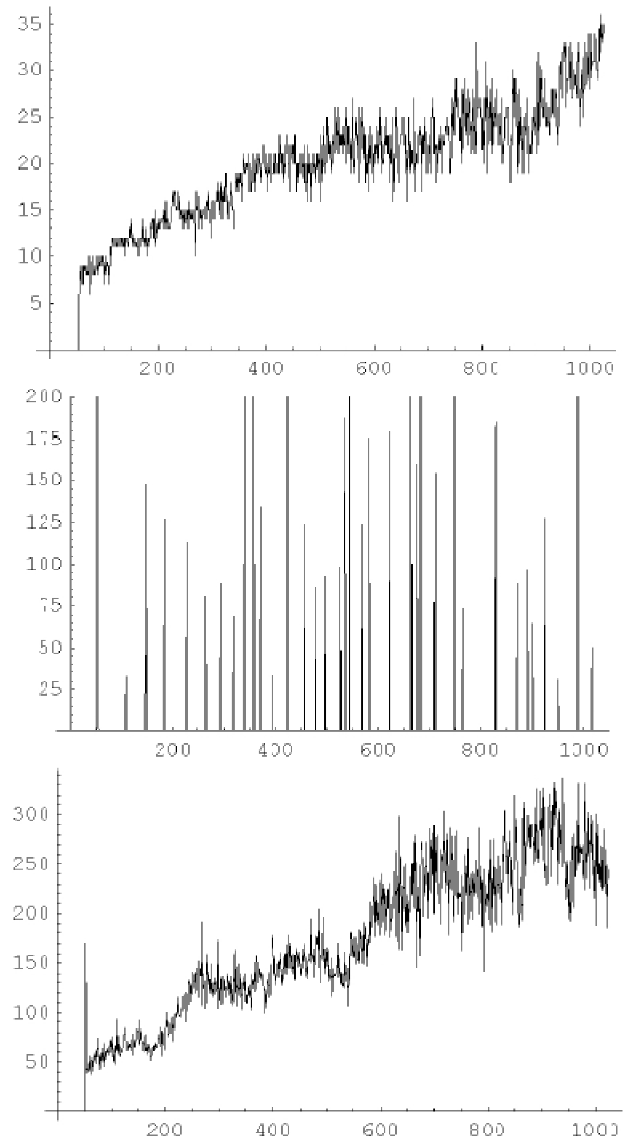


Figure 7: Bedau's measures applied to the economic activity of an agent-based model without the possibility of immigrants

supporting the successful skills. The top graph displays the number of active technologies, the graph in the middle shows the activity of new technology, and the bottom graph displays the average activity per active technology. The x-axes shows the iteration number as the simulation progresses, the y-axes shows the number of active technologies (top graph), the value of generated output at the time of the first application of new technology (middle graph), and the generated value of the output per active technology (bottom graph).

CONCLUSION

This paper presents a constructive agent-based model and the application of a measure to classify the generated evolutionary dynamics. The agent-based model simulates a market and production system, where the agents produce and exchange goods within their networks. Agents are profit driven, their production activity affects the supply and demand, this in turn again affects the set of prices, which changes the opportunities to realize a profit. On top of this agent-based model is a model that imitates a process of invention and innovation.

The results show that novelty in the form of new technology or new products does not only add to the diversity of the system, but adds to the output of the system as well. As the diversity of the system increases, the average output increases as well. This suggests that the dynamics generated by the model fall in Bedau's class 3 of unbounded evolutionary dynamics.

With the expanding and thus wide variety of technology present in the production system, a spatial agent-based modelling is an appropriate technique to study such systems; heterogeneous agents can mix and match different sets of technologies to find a combination of production means that works for their locale, and the agent-based production and market model provide an appropriate selection mechanism for novel technology and products.

ACKNOWLEDGMENTS

This work was supported by SSHRC grant No. 410-2004-1843 to R. White. We would like to thank R. White and W. Banzhaf for their very valuable contributions to the development of this work. We also like to thank ISEEE for their continued support of this research program.

REFERENCES

[1] R. K. Standish, “Open-ended artificial evolution,” International Journal of Computational Intelligence and Applications, vol. 3, no. 2, pp. 167–175, 2003.
[2] K. E. Boulding, “What is evolutionary economics,” Journal of Evolutionary Economics, vol. 1, no. 1, pp. 9–17, 1991.
[3] J. G. Lambooy and R. A. Boschma, “Evolutionary economics and regional policy,” Annals of Regional Science, vol. 35, no. 1, pp. 113–133, 2001.
[4] C. van Dinther, Handbook on Information Technology in Finance. Springer, 2008, ch. Agent-based simulation for research in economics.
[5] L. Tesfatsion, “Agent-based computational economics: Growing economies from the bottom up,” Artificial Life, vol. 8, pp. 55–82, 2002.
[6] J. Ferber, Multi-agent systems: An introduction to distributed artificial intelligence. Addison-Wesley, 1999.
[7] D. Marceau, Monitoring, Simulation, and Management of Visitor Landscapes. The university of Arizona press, 2008, ch. What can be learned form multi-agent system modeling.
[8] S. Heckbert and T. Baynes, “Agent-based modeling in ecological economics,” Ecological Economics Reviews. Accepted for publication, 2010.
[9] M. A. Janssen, Modelling in ecological economics. Edward Elgar Publishing, 2005, ch. Agent-based modelling.
[10] T. S. Schelling, “Dynamic models of segregation,” Journal of Mathematical Sociology, vol. 1, pp. 143–186, 1971.
[11] J. M. Sakoda, “The checkerboard model of social interaction,” Journal of Mathematical Sociology, vol. 1, pp. 119–132, 1971.
[12] E. Hatna and I. Benenson, Advanced Geosimulation Models. Bentham Sceince Publisher, 2010, ch. Geosimulation of Income-based urban residential patterns.
[13] L. Tesfatsion, “Agent-based computational economics: modeling economies as complex adaptive systems,” Information Sciences, vol. 149, no. 4, pp. 262–268, 2003.

[14] J. Doyne Farmer and D. Foley, "The economy needs agent-based modelling," Nature, vol. 460, no. 6, pp. 685–686, 2009.
[15] M. Buchanan, "Meltdown modelling," Nature, vol. 460, no. 6, pp. 680–682, 2009.
[16] K. Suzuki and T. Ikegami, "Spatial pattern induced evolution of self-replicating loop network," Artificial Life, vol. 12, no. 4, pp. 461–485, 2006.
[17] T. Miconi, "Evolution and complexity: The double-edged sword," Artificial Life, vol. 14, pp. 325–344, 2008.
[18] S. A. Kauffman, Investigations. Oxford, New York: Oxford University Press, 2000.
[19] G. Nicolis and I. Prigogine, Exploring complexity. New York: W.H. Freeman and Company, 1989.
[20] P. Dittrich, J. Ziegler, and W. Banzhaf, "Artificial Chemistries A Review," Artificial Life, vol. 7, pp. 225–275, 2001.
[21] M. A. Bedau, E. Snyder, and N. H. Packard, Artificial life VI. MIT Press, 1998, ch. A Classification of Long-Term Evolutionary Dynamics.
[22] M. A. Bedau, J. S. McCaskill, N. H. Packard, S. Rasmussen, C. Adami, D. G. Green, T. Ikegami, K. Kaneko, and T. S. Ray, "Open problems in artificial life," Artificial Life, vol. 6, pp. 363–376, 2000.
[23] J. von Neumann, "A model of general economic equilibrium," Review of Economic Studies, vol. 13, no. 1, pp. 1–9, 1945-1946.
[24] J. M. Blatt, Dynamic Economic Systems. New York: M.E. Sharpe, Inc., 1983.
[25] O. Morgenstern and G. L. Thompson, Mathematical Theory of Expanding and Contracting Economies. Lexington, Massachusetts: Lexington Books, 1976.
[26] R. K. Standish, "The role of innovation within economics," in Commerce, Complexity and Evolution, ser. International Symposia in Economic Theory and Econometrics, vol. 12. Cambridge University Press, 2000.
[27] J. A. Schumpeter, The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle. Cambridge Massachusetts: Harvard University Press, 1961.
[28] E. Mosekilde and S. Rasmussen, "Technical economic succession and the economic long wave," European Journal of Operational Research, vol. 25, pp. 27–38, 1986.
[29] E. Bruckner, W. Ebeling, and A. Scharnhorst, "Stochastic dynamics of instabilities in evolutionary systems," System Dynamics Review, vol. 5, no. 2, pp. 176–191, 1989.
[30] E. Bruckner, W. Ebeling, M. Jim´enez Montano, and A. Scharnhorst, "Nonlinear stochastic effects of substitution -an evolutionary approach," Journal of Evolutionary Economics, vol. 6, pp. 1–30, 1996.