Pre-proceedings of

ABMUS2018
the 3rd International Workshop on
Agent-Based Modelling of Urban Systems

15th of July 2018
held in Stockholm, Sweden,
at the Federated AI Meeting (FAIM2018)

http://www.modelling-urban-systems.com/abmus2018

ABMUS2018 organising committee:
- Dr Koen H. van Dam, Imperial College London
- Dr Jason Thompson, University of Melbourne
- Prof Pascal Perez, University of Wollongong
- Dr Nick Malleson, University of Leeds
- Dr Alison Heppenstall, University of Leeds
- Dr Andrew T. Crooks, George Mason University

ABMUS2018 scientific committee:
- Prof Ana Lúcia C. Bazzan, Universidade Federal do Rio Grande do Sul, Brazil
- Gonzalo Bustos-Turu, Imperial College London, UK
- Prof Eric Cornelis, Universite de Namur, Belgium
- Prof Remy Courdier, University of Reunion, France
- Prof Stephane Galland, Universite de Technologie de Belfort, France
- Dr Arnaud Grignard, MIT, USA
- Dr Ed Manley, University College London, UK
- Prof Jörg P. Müller, TU Clausthal, Germany
- Prof Kai Nagel, University of Berlin, Germany
- Dr Dhirendra Singh, RMIT University, Australia
- Dr Samarth Swarup, Virginia Tech, USA
- Prof Robert Tanton, University of Canberra, Australia
- Dr Sarah Wise, University College London, UK

Table of contents

Multi-Agent and Smart Buildings: What? and Why?

Nicolas Verstaevel[1], Johan Barthelemy[1], Pascal Perez[1]

[1]SMART Infrastructure Facility, University of Wollongong, Australia
`nicolas.verstaevel@uow.edu.au`

Keywords: Smart Cities · Smart Buildings · Multi-Agent Systems.

1 Extended Abstract

The expansion of the Internet of Things (IoT) is linked with the development of Smart Cities. As sensors are becoming cheaper and more powerful, networks of smart objects are becoming wider than ever. The scope of applications of IoT in Smart Cities is huge, ranging from mobility to waste management [14]. Smart Buildings have a huge role to play in this transformation [8]. This revolution is socio-technical, as it involves the relation between building occupants and its technology. In this work, we present the benefits of Multi-Agent Systems (MAS) applied to Smart Buildings. First, we study the evolution of the Smart Buildings concept from fully autonomous entities without human action, to a more people-centric approach with human feedback, highlighting the increasing levels of complexity. Based on a review of the scientific literature, we discuss how the multi-agent paradigm can offer innovative solutions to cope with this complexity. We then conclude with some perspectives and issues that still need to be addressed.

Defining what a smart building is, is not an easy task, as scientific and industrial literature is proficient of applications of new technologies to building management [3]. Common applications of Smart Buildings involve energy management, modelling of occupant behaviour, and user comfort [12]. The term *Smart Building* can be found in literature since the 1980s. For instance, [13] proposed in 1988 the following definition: *"a building which totally controls its own environment"*. This definition emphasizes the notion of **control**, and proposes to empower buildings with the use of technology to monitor their environment and control its features, with an important focus made on energy management. In their review, [16] highlighted that most of the early definitions indicated to reduce human interaction (even to none).

A latter definition, proposed by [5], defines an intelligent building as *"one that is responsive to the requirement of occupants, organization and society. It is sustainable in terms of energy and water consumption besides being lowly polluting in terms of emissions and waste: healthy in terms of well-being for the people living and working within it, and functional according to the user needs"*. This definition introduces a shift from the fully automated building of

the early ages to a more **people-centric** approach. Indeed, the focus is made on people, and control is not the goal of Smart Buildings, but a means to design them. Furthermore, they promote a multi-disciplinary approach to deal with the socio-technical aspects of smart buildings. [3] goes further in this people-centric vision. They define Smart Buildings as *"buildings which integrate and account for intelligence, enterprise, control, and materials and construction as an entire building system, with adaptability, not reactivity, at the core, in order to meet the drivers for building progression:energy and efficiency, longevity, and comfort and satisfaction. The increased amount of information available from this wider range of sources will allow these systems to become adaptable, and enable a Smart Building to prepare itself for context and change over all timescales"*. They identify **adaptivity** (to different people's perception of comfort, changes in occupant or building use, varying occupancy data characteristics,...) as the core concept of their definition. [7] propose that Smart Buildings should go further by not only being designed with an adaptive people-centric approach, but that Smart Buildings have to put **humans in the loop**. They argue that *"most of the current studies and solutions developed for building thermal control have been designed independent of the occupant feedback"*, leading to occupant dissatisfaction and/or technological misunderstandings. Putting humans in the loop is probably one of the major challenge that Smart Buildings will face, which is also a key factor for Smart Cities: the association of technology, people and institutions [1]. It involves not only learning from its occupant, but importantly to learn with them. This requires to design local feedback loops. Those loops are not only offering the possibility of interacting with the system and to shape it to the occupants' needs, but it also enables a cooperation between the building, its occupants and its governance by providing adequate information. Explainability, the ability to explain any non-trivial decision system to users, is thus becoming an important issue in Smart Environment [15].

Smart Buildings, such as Smart Cities, are truly complex systems [10]. Indeed, if we set aside the social and organizational aspects of Smart Buildings and focus only on technology, Smart Buildings are equipped with numerous heterogeneous sensors and networks. Smart Buildings are open, meaning that new sensors and data can appear or disappear at any time, and they have to face non-linear and unpredictable dynamics. Putting humans in the loop to design adaptive people-centric control systems will add another layer of complexity. This involves designing systems equally complex, as expressed by Ashby's law of requisite variety claiming that *if a system is to be stable, the number of states of its control mechanism must be greater than or equal to the number of states in the system being controlled* [2]. This level of complexity makes those systems difficult to be designed through traditional methods, as not all states of the system are known *a priori*. Muti-Agent Systems (MAS), through their decentralisation of decision and self-organisation capacities, offer interesting solutions to overcome this complexity.

MAS are systems composed of multiple interacting and autonomous entities, with the agents acting and sensing within a common environment. MAS offer

a framework to model, study and control complex systems with a **bottom-up approach**. This approach, focusing on the entities and their interactions, is able to solve a wide variety of problems [17]. Due to the distribution of tasks within the agents composing a MAS, and the possibility to decentralise control and decision, MAS are more suitable to model and simulate complex systems than traditional approaches [11]. We can classify the MAS application into three categories:

– **Modelisation**, where the MAS paradigm is used to distribute the tasks and roles and reflect the topological organization of a structure. For example, [6] proposes a review of MAS for smart-microgrid energy management and operation. The review focus on Smart Grids, a new type of decentralised infrastructure to produce and control energy. The authors see in MAS a paradigm enabling the **autonomy** of mini/microgrid, and the capacity to deal with the distributed topology of the emerging smart grid systems by its **decentralised approach**.
– **Simulation**, where MAS is used to reflect the interactions between social entities and produces previously unseen behaviours. For instance, [4] use MAS to build a stochastic simulation of occupants in buildings to provide designers the means to evaluate the performance of their designs. According to the authors, the MAS paradigm is used to *"enable the definition of archetypes and archetypal behaviours to account for diversity between occupants, take into account social interactions between members of a population and the corresponding implications for their behaviours, and enable behaviours that are conditional on others having already been exercised or indeed on proximity to the building envelope or system, with corresponding implications for interaction probability"*. Thus, they emphasise the **interaction** component of MAS enabling them to model behaviours.
– **Problem solving**, where the MAS paradigm is used to tackle the dynamic context in which agents are evolving and to propose self-adaptive and emerging solutions to complex problems. For example, [9] use the MAS paradigm to optimise energy consumption and allocation through the exchange of energy between multiple small scale battery, each one being managed by a small community. In this application, the MAS paradigm enables to solve a global problem (energy management) at a **local level** (a battery).

The MAS paradigm offers an intuitive and natural way to model entities and their interactions, focusing on the social components of the system. The decomposition of the problem into a set interacting agent allows the design of innovative solutions, with emergent features, and address a wide range of issues. The review of the literature illustrates this variety of applications within the Smart Building concept and highlights interesting perspectives for cross-domain applications.

In this paper, we identify some of the challenges of Smart Buildings relying on new adaptive people-centric control systems with humans in the loop. Those systems are complex and MAS offers an intuitive approach to deal with their modelling, simulation and control. However, while the decomposition into agents

is at the core of the MAS paradigm, there is a lack of frameworks and tools to link the macro level (where the problem is expressed) of a MAS and its micro level (at the agent level). The variety of applications of the MAS paradigm for Smart Buildings illustrates how the paradigm enables multidisciplinary research. This results in a great opportunity to trigger discussions among different fields in order to better understand their respective needs and to build the missing macro-micro link.

References

1. Albino, V., Berardi, U., Dangelico, R.M.: Smart cities: Definitions, dimensions, performance, and initiatives. Journal of Urban Technology **22**(1), 3–21 (2015)
2. Ashby, W.R.: Requisite variety and its implications for the control of complex systems. In: Facets of Systems Science, pp. 405–417. Springer (1991)
3. Buckman, A., Mayfield, M., BM Beck, S.: What is a smart building? Smart and Sustainable Built Environment **3**(2), 92–109 (2014)
4. Chapman, J.e.a.: On the multi-agent stochastic simulation of occupants in buildings. Journal of Building Performance Simulation pp. 1–18 (2018)
5. Clements-Croome, D.: Sustainable intelligent buildings for people. In: IBEA Conference 2011 Proceedings: Innovation and Integration-Science, Technology and Policy in the Built Environment. p. 205 (2012)
6. Coelho, V.N., Cohen, M.W., Coelho, I.M., Liu, N., Guimarães, F.G.: Multi-agent systems applied for energy systems integration: state-of-the-art applications and trends in microgrids. Applied energy **187**, 820–832 (2017)
7. Gupta, S.K., Kar, K.: Human-in-the-loop thermal management for smart buildings. In: Intelligent Building Control Systems, pp. 191–217. Springer (2018)
8. Kylili, A., Fokaides, P.A.: European smart cities: The role of zero energy buildings. Sustainable Cities and Society **15**, 86–95 (2015)
9. Mahmood, D., Javaid, N., Ahmed, I., Alrajeh, N., Niaz, I.A., Khan, Z.A.: Multi-agent-based sharing power economy for a smart community. International Journal of Energy Research **41**(14), 2074–2090 (2017)
10. Nigon, J., Verstaevel, N., Boes, J., Migeon, F., Gleizes, M.P.: Smart is a matter of context. In: International and Interdisciplinary Conference on Modeling and Using Context. pp. 189–202. Springer (2017)
11. Noël, V., Zambonelli, F.: Following the problem organisation: A design strategy for engineering emergence. In: Intelligent Distributed Computing VIII, pp. 311–317. Springer (2015)
12. Stojkoska, B.L.R., Trivodaliev, K.V.: A review of internet of things for smart home: Challenges and solutions. Journal of Cleaner Production **140**, 1454 – 1464 (2017)
13. Stubbings, M.: Intelligent buildings: An ifs executive briefing (1988)
14. Talari, S., Shafie-khah, M., Siano, P., Loia, V., Tommasetti, A., Catalão, J.P.: A review of smart cities based on the internet of things concept. Energies **10**(4), 421 (2017)
15. Wolter, D., Kirsch, A.: Smart environments: What is it and why should we care? KI - Künstliche Intelligenz **31**(3), 231–237 (Aug 2017)
16. Wong, J.K., Li, H., Wang, S.: Intelligent building research: a review. Automation in construction **14**(1), 143–159 (2005)
17. Wooldridge, M.: An introduction to multiagent systems. John Wiley & Sons (2009)

Agent-based modelling of a neighborhood's transition towards gas-free heating

Graciela del Carmen Nava Guerrero[1], Helle Hvid Hansen[1], Gijsbert Korevaar[1], and Zofia Lukszo[1,2]

[1] Energy and Industry Section, Faculty of Technology, Policy and Management,
Delft University of Technology, Delft ZH 2628BX, the Netherlands
[2] Politecnico di Torino. DENREG Group of Energy Department. Torino, Italy.
g.d.c.navaguerrero@tudelft.nl

1   Introduction

The EU aims at reducing greenhouse gas emissions to 80% by 2050, through domestic reductions [1]. As the heating and cooling sector accounted for 50% of the EU's annual energy consumption in 2016 [2], reducing its emissions is of great importance. Opportunities include adoption of efficient heat technologies and buildings' insulation [1, 3].

1.1   Purpose of this paper

We present an agent-based modelling (ABM) process for a neighborhood's thermal energy transition. We study a residential area where heating is provided by on-site technologies fueled by natural gas, as is often the case in the Netherlands. We explore how the adoption of buildings' insulation and heat technologies can enable the neighborhood to reduce its natural gas consumption. Hence, the question: How can a Dutch neighborhood transition from natural gas-based to natural gas-free heat supply over the coming years while meeting the *neighborhood'*s heat demand? We build on the theories of socio-technical [4] and complex adaptive systems (CAS) [3] to answer this question.

We define they key performance indicators (KPIs) as natural gas used for heating purposes [MWh] and expenses [Euros]. We explore the KPI values as a function of households' characteristics: value orientation (environmental, social or financial), time horizon to assess investments, ability to compare combined investments, and whether they invest on individual or building-wide solutions. We also explore how different energy retail prices and biogas as a partial substitute of natural gas influence the KPI values.

1.2   Modelling questions

Our modelling questions are the following:
1. Which combinations of household's characteristics lead to low natural gas consumption and low expenses at the end of the simulation?
2. What are promising combinations of technologies and insulation levels with which low natural gas consumption and low expenses were achieved?
3. How would the cost of heat supply be affected by promising combinations of technologies and insulation levels?

1.3   Relevance for the ABMUS 2018 workshop

We build on the work by [4] to design in socio-technical systems for the energy transition. We account for technology, institutions, actors and their interactions. In addition, we discuss how we cope with the double challenge of representing a complex system in the simplest possible way while producing quantitative results to support decision making. Our approach relies on the co-development of a simple agent-based model to define the problem with stakeholders, identify influential factors and processes, and discuss the key insights that they wish to obtain. Then, the model will be iteratively extended to refine its quantitative capabilities to support decision making.

2   Research approach

The theory of CAS and ABM are applied to study the neighborhood's heat provision as a socio-technical system. While the neighborhood's objective is to reduce greenhouse gas emissions at low cost, decentralized decisions by households determine whether this objective is met. These decisions depend on households' heterogeneity, which calls for a bottom-up approach to study the problem. Furthermore, properties of households can change over time as they learn and adapt. Learning may happen as households observe how investment decisions and external factors influence their own and their neighbor's KPIs. Adaptation may happen when households use their knowledge to adjust their decision rules. In this work, we focus on the study of bottom-up decisions. We intend to account for learning and adaptation in future work.

2.1   Model conceptualization

We model the neighborhood as households living in apartments, with an association of house owners (AHO) for each building. Each apartment has a natural gas condensing boiler and the same level of insulation. Households and AHOs are model agents. The environment, external to agents, tracks time, sells electricity, natural gas, biogas, heat technologies and improvements in dwelling's insulation, and sets their price. Theoretically, a household can decide to improve its apartment's insulation and replace its boiler with a micro combined heat and power unit, aerial or geothermal heat pump or electric

radiators. Technologies vary in efficiency and costs (capital and operation). Insulation can be improved to two levels, each with its own costs.

Households' decisions are based on whether they have the ability to assess combinations of insulation and heat technologies or can only assess them independently (AAI), time horizon as years into the future that they consider when assessing investments (HRZ), and value orientation (ORI), which can be environmental (Env), social (Soc) or financial (Fin). Households have perfect knowledge of their own heat demand, insulation level, heat technologies, budgets, as well as the environment. They can make simple financial cost-benefit analyses (CBA) of possible investments. Households do not account for biogas in their CBAs, as they do not know its cost or availability in the future. Households select their preferred option based on their CBAs and value orientation (ORI). After investing, a household cannot invest for a period equal to their HRZ, to represent limited financial resources.

When households invest with their neighbors on building-wide solutions, they allow their AHO to decide on their behalf (IWN=1). We assume that either all or none of the households living in the same building invest with their neighbors, that AHOs are able to assess combined investments (AAI=1), that all or none of the apartments will be insulated or will replace their boilers with a building-wide technology. Available building-wide technologies are combined heat and power units and electric heating. AHOs have perfect knowledge of all households in the building. Like households, AHOs base their decisions on their properties (AAI, HRZ, ORI). Changes in the values of KPIs as a result of the decisions of both AHOs and households are influenced by external factors defined in the model environment.

In this first iteration of the model, which will enable a discussion with stakeholders regarding future research steps, we make assumptions to represent the research problem in an intentionally simplified way. These include households having identical purchasing power, time horizon, number of residents, deterministic demand and operation, small natural gas condensing boilers and fixed decision rules by households. Further, we limit the choices of technologies that are available to both households and AHOs. We assume that AHOs are environmentally oriented and aim at reducing natural gas.

2.2   Experimental scenarios

The model was built in NetLogo [5] and is currently parameterized based on desk research, with real data considered for future applications. Table 1 presents the variables that were used to generate experimental scenarios.

We observe changes in the values of KPIs over 20 years. We test combinations of 5 populations of households with a fraction of all households with each combination of ORI: pop1=[Env:0.33, Soc:0.33, Fin:0.33], pop2=[Env:0.50, Soc:0.25, Fin:0.25], pop3=[Env:0.25, Soc:0.50, Fin:0.25], pop4=[Env:0.25, Soc:0.25, Fin:0.50], pop5=[Env:1]; 3 populations of households with percentages of IWN=1: 100, 50 and 0%; 3 populations of households with percentages of AAI=1: 0%, 50% and 100%; 6 HRZ for all households: 1, 5, 10, 15, 20 and 30 years; 3 dGp cases: -.025, 0 and .025; 3 dEp cases: -.06, 0 and .06; 3 Bp cases: 1.5*Gp, Ep and 2*Ep; 2 AB cases: 0 and 1.

Table 1. Groups of variables that will change during the simulations

Variables	Possible values
IWN: whether households invest individually or with their neighbors, allowing the AHO to decide.	0 or 1
AAI: ability to assess combinations of investments.	0 or 1
HRZ: time horizon when assessing investments.	Number of years
[ORI]: environmental (Env), social (Soc) and financial (Fin) value orientation.	"Env", "Soc" or "Fin".
dGp and dEp: annual changes in natural gas and electricity retail prices (Gp and Ep, respectively).	Negative and positive fractions.
Bp: biogas retail price, as a function of Gp and Ep.	Function of Gp and Ep
AB: adoption of biogas, where 1% of the natural gas used the previous year is replaced by biogas this year.	0 or 1

3   Expected results and discussion

When making decisions individually, we expect households with longer HRZ and higher environmental values to be the first ones to replace their boilers or improve their apartments' insulation. These choices may also lead to higher expenses. We expect financially oriented households to keep their boilers for a longer time, and possibly insulate their apartments in order to reduce their expenses. Furthermore, we expect to find possibilities to achieve relatively low natural gas consumption and relatively low expenses when households invest with their neighbors and try to minimize natural gas consumption through building-wide solutions.

Insights from this study will be used to explore opportunities to steer neighborhoods towards natural gas-free futures in two related ways. First, by discussing modelling concept and outcomes with stakeholders to co-define a research agenda. Second, to identify functional requirements of a model with quantitative capabilities.

References

1. EC Press release on heating and cooling. http://europa.eu/rapid/press-release_MEMO-16-311_en.htm, (last accessed 2017/05/3.
2. Balaras, C.A., Gaglia, A.G., Georgopoulou, E., Mirasgedis, S., Saradis, Y., and Lalas, D.P.: European residential buildings and empirical assessment of the Hellenic building stock, energy consumption, emissions and potential energy savings. Building and Environment, 42(3), 1298-1314 (2007).
3. Holland, J. H.: Complex adaptive systems. Daedalus 121(1), 17-30 (1992).
4. Moncada, J.A., Park Lee, E.H., Nava Guerrero, G.D.C., Okur, O., Chakraborty, S.T., Lukszo, Z.: Complex Systems Engineering: designing in sociotechnical systems for the energy transition. EAI Endorsed transactions on energy web 17(11):e1, 1-13 (2017).
5. Wilensky, U. (1999). NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Understanding Input Data Requirements and Quantifying Uncertainty for Successfully Modelling 'Smart' Cities [*]

Nick Malleson[1,3][0000−0002−6977−0615], Jonathan A. Ward[2][0000−0003−3726−9217], Alison Heppenstall[1,3][0000−0002−0663−3437], Michael Adcock[3], Daniel Tang[4], Jonathan Coello[4], and Tomas Crols[1,3][0000−0002−9379−7770]

[1] School of Geography, University of Leeds, LS2 9JT, UK
http://geog.leeds.ac.uk/
n.s.malleson@leeds.ac.uk
[2] School of Mathematics, University of Leeds, LS2 9JT, UK
http://maths.leeds.ac.uk
[3] Leeds Institute for Data Analytics (LIDA), University of Leeds, LS2 9JT, UK
http://lida.leeds.ac.uk
[4] Improbable, 30 Farringdon Road, London, EC1M 3HE, UK
http://www.improbable.io

Abstract. Agent-based modelling (ABM) is ideally suited to modelling the behaviour and evolution of social systems. However, there is inevitably a high degree of uncertainty in projections of social systems – input data are noisy and sparse, and human behaviour is itself extremely uncertain – one of the key challenges facing the discipline is the quantification of uncertainty within the outputs of agent-based models. Without an adequate understanding of model uncertainty, or a means to better constrain models to reality, simulations will naturally diverge from the target system. This limits the value of their predictions. This paper presents ongoing work towards methods that will (i) allow real-time data to be assimilated into models to reduce the uncertainty of their predictions and to (ii) quantify the *amount* of data (including overall volume as well as spatio-temporal granularity and regularity) that are required for successful assimilation (i.e. to model the system within an acceptable level of uncertainty). Specifically, this project emulates a simple system of pedestrians who all move towards an exit. The paper reports on initial experiments to constrain the range of possible model outcomes using Bayesian inference techniques as implemented in a new probabilistic programming library. Ultimately the project aims to provide valuable information about the number and type of sensors that might be required to model movements of humans around real urban systems.

1   Introduction and Background

Individual-level modelling approaches, such as agent-based modelling (ABM), are ideally suited to modelling the behaviour and evolution of social systems. This is especially true in the context of modern 'smart' cities, where large volumes of data, supported by innovative 'big' data analytics, can be leveraged to better understand and capture the characteristics of the underlying systems. However, there is inevitably a high degree of uncertainty in projections of social systems – input data are noisy and sparse, and human behaviour is itself extremely uncertain – one of the key challenges facing the discipline is the quantification of uncertainty within the outputs of these models.

This work focusses on the simulation of *urban flows*, i.e. the movement of people around urban areas over relatively short time scales (minutes and hours). It presents ongoing work towards methods that will (i) allow real-time data to be assimilated into models to reduce the uncertainty of their predictions and (ii) to quantify the *amount* of data (including overall volume as well as spatio-temporal granularity and regularity) that are required for successful assimilation. In other words, given some human system and a simulation of that system (we study the movement of pedestrians in this case), how much data are required from the real system in order to prevent the model uncertainties from causing the simulation to rapidly diverge from reality? With too little data it will be impossible to reliably constrain the model to reality, but how much is too little? Is one well-placed footfall counter sufficient to capture the dynamics of the system, or in reality would it be necessary to track the actual movements of a large proportion of the individual people? The hypothetical model used here is a simple system of pedestrians, each of whom move from an entrance towards one of two exits. This is analogous to a train arriving at a train station and passengers moving across the concourse to leave. The model environment is illustrated in Figure 1.

Very limited prior work has been conducted in this area. Some authors have attempted to conduct data assimilation, but use agent-based models that are simple in the extreme [5, 7]. Here, a more advanced model will be used in order to test how well the various methods handle complex features such as emergence and feedback. Others have developed more complex agent-based models [1] but those must remain mathematically coupled to an aggregate proxy model which limits the flexibility of the underlying agent-based model. The most similar work is that of [6], who attempt to assimilate data into a model of peoples' movement in buildings. They do this by running an ensemble of models and re-starting each one using new input conditions that are created each time new data become available. The main difference between the approach in [6] and this paper is that here the aim is to eventually develop methods that are able to assimilate new data that automatically *alter the state* of the simulation while it is running. This is more analogous with data assimilation techniques in other fields [4].

Fig. 1: A snapshot of the simple, hypothetical model that is used here. Agents arrive from the green entrances on the left and move towards the red exits on the right.

2   Methods

2.1   Overall Approach

This paper presents a work in progress, whose overall proposed approach is to:

1. Develop a agent-based simulation of pedestrian movements from an entrance to an exit using a simple social forces framework (similar to [3]) to control the behaviour of the agents (see Figure 1);
2. Run the simulation to generate hypothetical 'truth' data, our equivalent reality created with synthetic data. We assume this equivalent reality represents the real underlying system. We sample from this in a manner analogous to that of using sensors of peoples' movements to sample from the real world;
3. Re-run the simulation with a different random seed and use samples of varying resolution from the 'truth' data to reduce the uncertainty of the new simulation;
4. Quantify the volume (total number of samples), granularity (amount of aggregation) and regularity (the number of samples per time period) that are required to successfully model the hypothetical 'real' data.

This paper presents the preliminary results up to stage 3, with the ultimate aim of quantifying the amount of data required to successfully constrain the simulation being immediate future work.

2.2 Preliminary Results: Constraining Model Uncertainty

Our initial experiments use a prototype probabilistic programming library to perform parameter estimation using Bayesian inference (see [2] for a useful discussion about probabilistic programming and Bayesian inference). This work is a precursor to performing data assimilation (i.e. step 4). The probabilistic programming approach treats the pedestrian simulation as a black box, whose only *input* is a list of random numbers that are used in the simulation whenever a probabilistic decision is made and whose *output* is the number of agents in the system at each time step. All internal simulation parameters are fixed. The simulation output is deterministic given the same input, but different inputs result in different outputs, and so stochastic simulations can be performed by choosing the input at random. Figure 2 illustrates the method.

We create truth data, i.e. agent counts over time, from a single model run using a model input (a list of random decimal numbers) sampled from a Gaussian distribution. Then, using a noisy observation of the truth data, we use the probabilistic programming library to create a Bayesian network from which we can compute the posterior distribution of the input given the noisy observation. We then sample from this distribution using Metropolis-Hastings, resulting in an ensemble of model realisations in which the uncertainty of the input, and hence the output, is constrained.
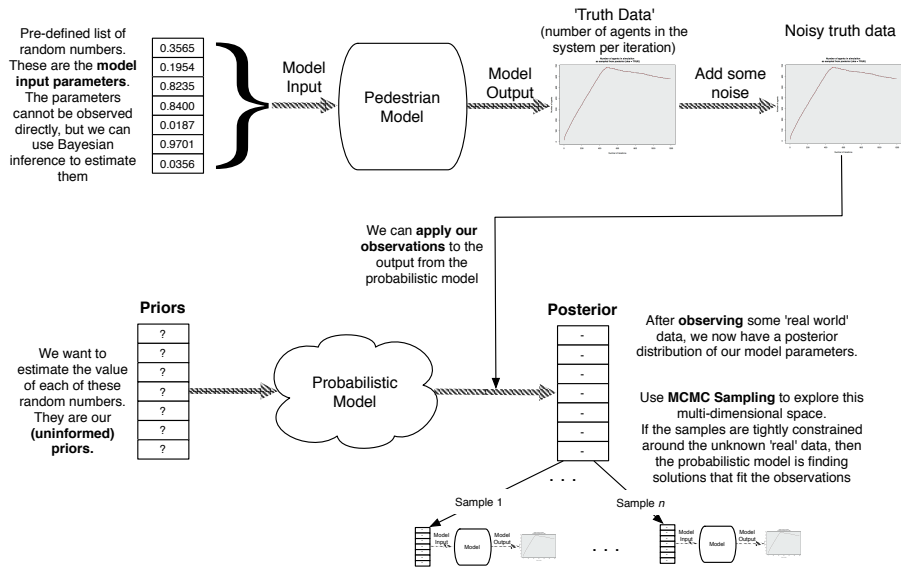

Fig. 2: An illustration of the procedure to perform parameter estimation.

Figure 3 compares the results of the sampling with and without the incorporation of the truth data. It is clear that when the Bayesian network makes use of the observations from the 'truth' model, the output of each model (i.e. sample of the posterior distribution) is much more similar to the truth data. In other words, the procedure is more accurately estimating the list of random numbers (the prior distribution) that were initially used as input to generate the 'truth' data. Although this result is not of value in isolation, it is very useful as a proof-of-concept. It shows that Bayesian inference on an agent-based model that has reasonably complex characteristics (namely the emergence of crowding due to agent interactions) is able to perform parameter estimation. More rigorous data assimilation is a relatively small step.

(a) constrained to observations   (b) without observations

Fig. 3: Results of sampling the posterior with observations (3a) and without (3b). When the 'truth' data are used to constrain the posterior distribution, the sampling routine is much better able to estimate the input model state, so the outcomes of the samples are much closed to the 'truth' data.

3   Link to ABMUS Workshop Themes

This paper addresses a core aspect of the ABMUS workshop theme; that of the *trust* that we can have in model outputs. The conference recognises challenges in "designing, developing and implementing **trusted** models that can be used by industry and governments to enhance decision-making". By adapting existing methods that are aimed at reducing uncertainty in models through the incorporation of up-to-date data, this work advances the methodology towards more a rigorous representation of real urban systems that will be more acceptable for policy use.

References

1. Flötteröd, G., Bierlaire, M., Nagel, K.: Bayesian Demand Calibration for Dynamic Traffic Simulations. Transportation Science **45**(4), 541–561 (Nov 2011). https://doi.org/10.1287/trsc.1100.0367
2. Ghahramani, Z.: Probabilistic machine learning and artificial intelligence. Nature **521**(7553), 452–459 (May 2015). https://doi.org/10.1038/nature14541
3. Helbing, D., Molnár, P.: Social force model for pedestrian dynamics. Physical Review E **51**(5), 4282–4286 (1995). https://doi.org/10.1103/PhysRevE.51.4282
4. Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press (2003)
5. Malleson, N., Tapper, A., Ward, J., Evans, A.: Forecasting Short-Term Urban Dynamics: Data Assimilation for Agent-Based Modelling. In: Proceedings of the European Social Simulation Conference, 25-29 September 2017, Dublin. p. 13 (2017)
6. Wang, M., Hu, X.: Data assimilation in agent based simulation of smart environments using particle filters. Simulation Modelling Practice and Theory **56**, 36–54 (2015). https://doi.org/10.1016/j.simpat.2015.05.001
7. Ward, J.A., Evans, A.J., Malleson, N.S.: Dynamic calibration of agent-based models using data assimilation. Royal Society Open Science **3**(4) (2016). https://doi.org/10.1098/rsos.150703

Steps Towards Scalable Agent-Based Simulation Model: Impact of the Time Scheduling Approach⋆

Tahina Ralitera[1], Nathan Aky[1], Denis Payet[1], and Rémy Courdier[1]

Laboratoire d'Informatique et de Mathématiques, University of Reunion Island,
France
{tahina.ralitera,denis.payet,remy.courdier}@univ-reunion.fr
nathan@smg.re

Abstract. In this extended abstract, we are interested in the topic of large scale urban and transport simulation applications. Most of the time, in these simulations, users have to find a compromise between the level of detail to take into account, the number of simulated agents and the computational resources available. We think that the time scheduling approach could be an alternative solution to explore. As illustration, we show how the use of an approach called the temporality model could increase the scalability of the SKUADCityModel while maintaining an acceptable level of performance.

Keywords: Scalability · Urban and transport simulation · Scheduler.

1 Introduction

Agent-based simulation models are often used as decision support tools. They intend to imitate complex phenomena that could occur in the real world, in order to evaluate their possible consequences. Thus, a simulation model should meet a set of constraints related to its field of application [5].

The transport and urban systems involve a large number of heterogeneous agents that operate in complex and very large sizes environments. That requires computational resources. Thus, when working on personal computers, the difficulty is to build a simulation model that can scale up while maintaining acceptable performance, short execution time and relevant results. Common solutions consist in finding a compromise between the level of detail taken into account, the number of simulated agents and the computation power available. However, in these kind of simulations, the level of details taken into account is important as it could affect the relevance and accuracy of the results [2].

We think that the time scheduling approach could be an interesting track to explore. This topic is relevant in the context of ABMUS Workshop because we tackle a problem of agent-based simulation of urban and transport systems' scalability.

2   The Temporality Model Approach

The simulation platform part that is responsible for the virtual time management is called the scheduler. This scheduler uses different approaches [1] such as the time-stepped, the event-driven and the mixed approaches. However, these conventional approaches may be inappropriate or may have limits in some cases (see Figure 1). These restrictions can become a barrier to the scalability.

The authors of [3] proposed an approach called the "temporality model". This approach intends to fulfil a set of criteria that the classical approaches of scheduler do not meet. So, we assume that it could be a potential solution to explore.

Table 1. Advantages and limitations of conventional scheduler approaches.

Approach	Advantages	Limits
Time-stepped	Easy to set up; Convenient in case of homogeneous agents.	Limited in case of large number and highly heterogeneous agents.
Event-driven	Convenient in case of highly heterogeneous agents.	No control over the simulation execution duration; The simulated time can take very complex forms; Calculations can become heavy.
Mixed	Use of different types of scheduler appropriate to each submodels.	No control over the simulation execution duration.

The temporality model approach tries to combine the advantages found in the three different approaches mentioned above:

- The periodicity and reducibility of the execution time that we can have when using the time-stepped approaches;
- The precision of the event-driven approaches;
- The adaptability to complex models of the mixed approaches.

The agents needs are expressed using a data structure called the "temporality". A temporality corresponds to a point on the time axis for which an agent wants to be activated. The agents are invited define their temporalities during the simulation initialization. Afterwards, if they need to adjust their behaviour, they will be able to redefine or create new temporalities at any time. The scheduler immediately processes the creations and the modifications, then updates the time axis. Thus, at a given time, the scheduler activates only the agents who need to be activated. This could reduce the computation resources required. Consequently, simulating a large number of heterogenous agents becomes easier.

Another particularity of the temporality model, approach is the **ability to add time constraints**. This is done by specifying three regulation properties: the minimum time-step, the default time period (used to activate agents that do not define any temporality) and the variability. This makes the difference compared to event-driven or mixed approaches. Users are able to reduce the simulation execution time by varying these parameters. In the extreme cases, if no agent defines any temporality, we automatically fall back into a time-stepped approach. On the other hand, if all the agents are complex and express a large number of temporalities, we end up with an event-driven type of scheduling.

3   The SKUADCityModel prototype

The SKUADCityModel is built upon the SimSKUAD simulation platform developed in our research laboratory. It is written in Java language and uses, by default, the temporality model approach. In our case study, experiments are conducted for individual transports in the city of Saint-Denis, the capital of the Reunion Island.

The agents and their environment models are based on the SmartCityModel simulation, developed in the Imperial College London, built upon the Repast Simphony simulation platform and described in [4].

In the SKUADCityModel, pretreatments such as adding a road cache are done during the initialization of the simulation. Thus, the simulation can hold a very detailed definition of the roads and the buildings without a great deterioration of performance. Figure 1 shows a part of the SKUADCityModel simulation GUI. The agents are in blue, the roads are represented by black polylines and the area of interests are in green or in red.
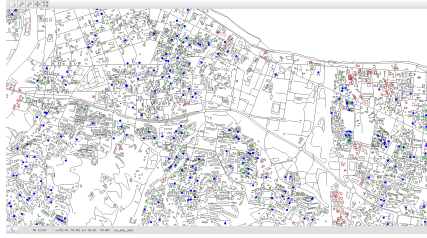
To show how the SKUADCityModel scales up according to the type of scheduler used, we implemented two different types of scheduler: a time-stepped approach and a temporality model approach. We choose to not make an implementation of the event-driven or the mixed approaches because they are not appropriate if we want the users to have control over the simulated time in order to reduce it for example (see Figure 1).

4   Comparison and results

The experiments are conducted on a personal computer with the following configuration: Intel core i5 (fifth generation), 16 gigabytes of RAM and a Solid State Drive. The comparison is based on the execution duration performance. Two types of manipulations are done (see Figure 2):

– **A variation of the experimental constraints** from 0% to 100%. In this way, we show how the simulation duration can be reduced depending on the time scheduling approach used.
– **A variation of the number of agents** from 1000 to 10000 moving over 12 hours. In this way, we show how the simulation model can scale up depending on the time scheduling approach used.

Figure 2 shows the results of the first experiments carried out. First, when we vary the experimental constraints (see Figure 2), the simulation durations can become shorter in the case of the use of the temporality model approach. Second, the results show that when using the time-stepped approach, the maximum number of agents supported by SKUADCityModel does not exceed 10,000. In the Figure 2, the yellow curve on 0ms indicates a crash due to an out of memory error. This is a problem we do not encounter, at least up to 10,000 agents, when we use the temporality model approach. Finally, the performances remain acceptable and the simulation execution times are less than five minutes.

Fig. 1. A Screenshot of a part of the SKUADCityModel user GUI.

Fig. 2. Execution duration performance of each scheduling approach.

5   Summary

In this extended abstract, we intend to show how the time scheduling approaches could play a role in urban and transport simulation scalability. First, we describe in a general way the bases of the temporality model approach. Second, as an illustration and an application, we built an agent based simulation called the SKUADCityModel. We run it using two different scheduling approaches: the temporality model and the time-stepped approaches. Finally, we made a comparison based on each simulation execution duration performance. First results are quite convincing. However a deeper analysis should be done with an even larger number of agents, a wider and more detailed environment and more evaluation criteria. For comparison, it could also be interesting to implement the temporality model in other simulation models built upon other platforms.

References

1. Fujimoto, R.M.: Time management in the high level architecture. Simulation **71**(6), 388–400 (1998)
2. Navarro, L., Flacher, F., Corruble, V.: Dynamic level of detail for large scale agent-based urban simulations. In: 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011), Taipei, Taiwan, May 2-6, 2011, Volume 1-3. pp. 701–708 (2011)
3. Payet, D., Courdier, R., Ralambondrainy, T., Sébastien, N.: Le modèle à temporalité: pour un équilibre entre adéquation et optimisation du temps dans les simulations agent. In: Systemes Multi-Agents, Articulation entre l'individuel et le collectif - JFSMA 2006 - Quatorzieme journees francophones sur les systemes multi-agents, Annecy, France, October 18-20, 2006. pp. 63–76 (2006)
4. Ralitera, T., Ferard, M., Bustos-Turu, G., van Dam, K.H., Courdier, R.: Steps towards simulating smart cities and smart islands with a shared generic framework - A case study of london and reunion island. In: SMARTGREENS 2017 - Proceedings of the 6th International Conference on Smart Cities and Green ICT Systems, Porto, Portugal, April 22-24, 2017. pp. 329–336 (2017)
5. Reynaud, Q.: A hybrid and flexible agent architecture for urban simulations: behavior composition and anticipation issues. Ph.D. thesis, Pierre and Marie Curie University, Paris, France (2014)

A framework to improve trust in Agent-Based Models of Human-Environment Interactions

Cédric Grueau[1] and João Araújo[2]

[1] Escola Superior de Tecnologia de Setúbal - Instituto Politécnico de Setúbal,
Campus do IPS, Estefanilha, 2910-761 Setúbal, Portugal
[2] Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Quinta da Torre,
2829-516 Caparica, Portugal
http://nova-lincs.di.fct.unl.pt/person/27

Abstract. Agent-Based Modelling (ABM) and simulation have gained popularity in disciplines dealing with complex systems, such as urban systems. Despite the increasing number of models built by experts and users, it is not always guaranteed that one is able to replicate the basic results of a research model nor to understand it. In this paper, we raise the questions that need to be answered to cope with ABM specification issues. We review some of the existing solutions that have been developed. We aim to address this problem by building a framework that includes a domain specific modelling language to respond to ABM documentation concerns. We show that ideas from the Web Semantic can enhance the understanding of ABM and improve users' trust in this kind of tools.

Keywords: Domain ontology · Geographic Information Science · Agent-Based Modelling · Domain Specific Modelling Language.

1 Introduction: why documenting ABM matters

To understand geographical problems such as sprawl, congestion and segregation, researchers have begun to focus on agent-based modelling (ABM) which allows one to simulate the individual actions of diverse agents, and to measure the resulting system behaviour and outcomes over time [5]. Modellers have been using ABM to simulate social interactions to understand and manage cities and urban infrastructure systems. [1], [2] and [3] are some examples of a broad range of applications.

As the number of models created by researchers increases, the opportunity to analyse a model and replicate its results naturally grows. Modellers are seeking to transfer the knowledge from one domain to another domain or to replicate an ABM and adapt it to their own case studies. It turns out that comparing different models representing a same empirical fact is a difficult task. To comprehend a model and assess the results of its simulation, these models need to be clearly described. This description is necessary not to only enhance communication, replication and comparison of models, but also to enable dialogue among disciplines [7].

2   Related work

For the last decade, initiative have been developed to enhance the description and communication of ABM. The ODD protocol [7] is a protocol originally formulated by ecologists that aims at providing a standard layout for describing individual and agent-based simulation models. It especially aims at documenting ABM for scientific publication and consists of several building blocks that facilitate writing and reading of model descriptions, using natural language. ODD is a real step forward communicating ABM. However, it is not a formal description aiming at being read by machine nor at being directly compiled to computer code. It is then still subject to interpretation. On the other hand, the degree of ODD's acceptance among experts has not been even between domains. In some domains such as Urban Systems the acceptance has been slow.

Another interesting initiative is the OpenABM (http://openabm.org), a repository of ABM contributed by users. OpenABM website is a portal specifically designed to facilitate the dissemination of simulation code, where contributors are required to upload the simulation code and describe the models using the ODD protocol. With about 500 models, at the time of this publication, OpenABM has now grown to a cyber infrastructure to preserve computational models and their digital context on agent based modelling. The sharing of models and computational code certainly contributes to facilitate models replication and validation. Nonetheless, the link between a model description with ODD and the respective simulation code is not verified and each user would need the programming skills to verify if description conforms to implementation.

Provenance in agent-based simulations is another way of tackling the specification of ABM. Provenance, which is a type of metadata, is the lineage of a data product or process: its creator, contributing processes, interactions, and data sources. In the context of agent based modelling, provenance captures state changes in individual agents and interactions through time [4]. In [8], authors investigated the role of provenance and proposed to record three types of provenance about a simulation: provenance about the social process of model development, about the execution of a model and about the history of a simulation. Authors concluded that as the number of simulation runs and agents in the simulation increases, these queries become exponentially complex and the application to more complicated model might be challenging.

3   A framework for the specification of agent-based models in socio-spatial systems

We believe that it should be possible to relate a model description (e.g., with ODD) to a simulation code. Many ABM practitioners are not computer scientists who can write and read computer programs. We think that this gap between description and implementation remains a challenge for many ABM users. To address this question we propose to adopt a research approach consisting of the development of a Domain Specific Modelling Language (DSML). DSMLs are

specification languages that offer, through appropriate notations and abstractions, expressive power focused on, and usually restricted to, particular problem domains [6]. On the other hand, in computing, linked data describes a method of publishing structured data so that it can be interlinked and become more useful through semantic queries. It builds upon standard Web technologies such as HTTP, RDF and URIs, but rather than using them to serve web pages for human readers, it extends them to share information in a way that can be read automatically by computers. Using ODD as a domain model, we aim at using it to compare the executable systems (source code for simulation) and the model description, to verify the conformity between both artefacts. Using this DSML, we could represent the model instances from the code as linked data and compare it to ODD representation and vice versa.

4   Conclusions

In this paper, we present the importance of agent-based model documentation and the challenge for practitioners to verify its conformity to models implementations. We identified that a gap between system representation and system implementation still exist for modellers, which represents a limitation to ABM communication and replication. Our goal is to provide a modelling language that would meet users needs and that would be straightforward enough to be adopted by the main users of the domain. We aim at presenting the first steps towards the development of this modelling language.

References

1. Banos, A., Genre-Grandpierre, C.: Towards New Metrics for Urban Road Networks: Some Preliminary Evidence from Agent-Based Simulations, pp. 627–641. Springer Netherlands, Dordrecht (2012)
2. Barile, F., Bove, L., Napoli, C.D., Rossi, S.: City Parking Allocations as a Bundle of Society-Aware Deals, pp. 167–186. Springer International Publishing, Cham (2017)
3. Barros, J.: Exploring Urban Dynamics in Latin American Cities Using an Agent-Based Simulation Approach, pp. 571–589. Springer Netherlands, Dordrecht (2012)
4. Chen, P., Plale, B., Evans, T.: Dependency provenance in agent based modeling. In: Proceedings of the 2013 IEEE 9th International Conference on e-Science. pp. 180–187. ESCIENCE '13, IEEE Computer Society, Washington, DC, USA (2013)
5. Crooks, A.T., Heppenstall, A.J.: Introduction to Agent-Based Modelling, pp. 85–105. Springer Netherlands, Dordrecht (2012)
6. Deursen, A.v., Klint, P., Visser, J.: Domain-specific languages: An annotated bibliography. ACM SIGPLAN Notices **35**(6), 26–36 (Jun 2000)
7. Grimm, V., Polhill, G., Touza, J.: Documenting Social Simulation Models: The ODD Protocol as a Standard, pp. 349–365. Springer International Publishing, Cham (2017)
8. Pignotti, E., Polhill, G., Edwards, P.: Using provenance to analyse agent-based simulations. In: Proceedings of the Joint EDBT/ICDT 2013 Workshops. pp. 319–322. EDBT '13, ACM, New York, NY, USA (2013)

Toward Simulation Science for Dense Space Management through Artificial Society Approach

Toshiyuki Kaneda [1], Manabu Ichikawa[2], Hikaru Uchida[3], Masakazu Takahashi[4], Reiko Hishiyama[5], Koichi Moriyama [1], Yuanyuan Liu [1], Sohei Inagaki [1], Tomohiko Kino [1], Youtaro Maekawa [1], Yoshimasa Matsumoto [1]

[1] Nagoya Institute of Technology, Japan
[2] Shibaura Institute of Technology, Japan
[3] Aoyama Gakuin University, Japan
[4] Yamaguchi University, Japan
[5] Waseda University, Japan

Abstract. Recent crowd accidents have posed a challenge to the complex science pedestrian simulation studies established during the 1990s. In this article, we describe a vision of our research initiatives aiming at establishing a problem-solving science based on pedestrian simulation using an artificial society approach for dense space management. To secure our initiative, we address a five-layered framework for integration of pedestrian agent models, and for our first key project refer to the 2001 Akashi pedestrian bridge accident as a specimen case. We will cover the three layers of Physical interaction, Personal spacing and Spatial cognition, and as the current position of our model development also show vision-driven agent simulations.

Keywords: artificial society, multi-layered agent model, Akashi pedestrian bridge accident, vision-driven agent

1   Our research question

Pedestrian dynamics research originating from complex system research in the 1990s has been developed as an agent-simulation model representing individual pedestrians by dynamic equations of micro-motives, and this has contributed to progress on the mechanism of crowd accidents at urban mass events. The absence of accidents over the intervening years could be seen as evidence of the successful application of counter-measures, but in the 2010s, systemic risks that cannot be handled by earlier approaches were seen in such incidents as the 2011 Love Parade disaster [1], the Shanghai stampede in 2014, and the Mina crowd accident near Jamarat bridge in 2015. These events are indicative of the growth of urban mass events, and there is a need for new approaches.

This situation as shown in Fig.1 involves two differing approaches. Pedestrian dynamics can be described as a Schelling approach in which micro factors contribute to unintended crowd behavior by manifesting the micro-interaction assumption (physical collision and social force) between agents. However, the systemic risk of recent years has

forced the introduction of an El-Farol approach in which each agent is influenced by feedback information about the state of their society. Our artificial society approach introduces both of them.

Fig. 1. Artificial society approach for agent-based urban modeling

In this paper, the spatial inhomogeneity of crowd density is called "nigiwai (dense, crowded or bustling)" and is a central research theme, but from the viewpoint of 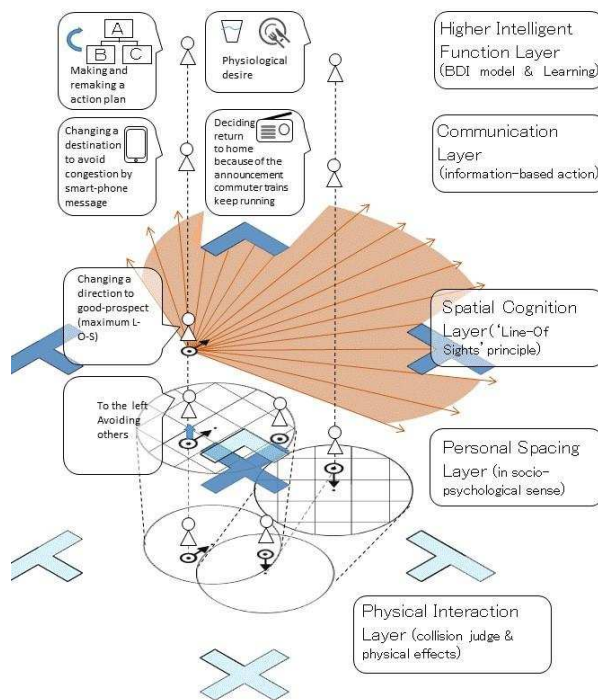dense space management, three more features are considered. (1) Mesoscale: heterogeneity, whether crowd-scale or space-scale, occurs on the mesoscale between the macroscopic and microscopic. This system is between the closed system and the open system. (2) Spatial form: whether it is a crowd accident or a purchase singular point (hot point) of a downtown area, it is influenced by the physical form of the space. (3) Citizen science: many citizens participate in the dense space, where they sense the conditions around them. We suggest that we consider diversified civic values in both the measurement data and analytical evaluation. In fact, research papers by Helbing et al. who analyzed the Love Parade disaster mention over 150 citizens evidence data files [1].

Our research initiative adopts an artificial society approach based on intelligent agents, but from the viewpoint of dense space management, the composition of the agent, such as the spatial recognition of an agent who is acting, message transmission/ reception, and micromovement due to their purchasing motive. For that reason, we are adopting a constructive methodology of exploring knowledge through simulation for refinement of the consequence space, without concern for any inevitable "built-in (or heavy)" model compositions. The simulation itself is a tool of deductive science, in the sense that it derives a possible consequence from a given set of conditions, but in our methodology, it is empirical science with the following two meanings. Firstly, we can only solve the problem by utilizing a wide variety of real-world information source data. Secondly, we are getting close to effective scientific knowledge only through the trial and error of model compositions and comparisons of various complexity. The core of our research question is exactly found by answering: how can we utilize diversified real-world data to approach the problem-solving? Moreover, how do we compose a wide variety of models to gain scientific knowledge?

2   A five-layered framework for integration of pedestrian agent models

The key feature of our research initiative lies in the artificial society approach which represents dense space which is composed of intelligent agents. The spatial behavior of the agent is composed of the following five ascending layers, and as we conduct further research into each layer and aim at an integrated model dealing with multiple layers, we approach the heart of the problem (see Fig.2).
(a) Physical Interaction Layer: handles physical consequences at the time of contact/collision among pedestrians or between pedestrians and obstacles such as walls.

Fig. 2. Five layered frameworks on spatial behavior of pedestrian models

A pedestrian walking direction is given as various conditions (exogenous variables).
(b) Personal Spacing Layer: deals with the effect of personal spacing raised in social psychology, in situations of local avoidance without physical interaction. The walking direction is endogenized by given waypoints.
(c) Spatial Cognition Layer: constructs a vision field with Line-Of-Sight (LOS) and handles direction changes by the "LOS principle" during exploratory behavior. The line of sight should generally be represented in the three-dimensional space. The destination is given, and the agent's waypoints are calculated as a result.

4

(d) Communication Layer: switches behaviors by acquiring information. "Information-based action" in a narrow sense includes any action intended to obtain information such as using personal information devices and broadly refers to any spatial behavior affected by such acquired information. The destination may be changed in response to results.

(e) Higher-order Intelligent Functional Layer: assumes a BDI (belief, desire, intention) model in artificial intelligence research, including emotions as well as planning and learning dynamics and these, may conflict with resolutions. Multi-purpose multi-stop behavior called shopping-around behavior in a downtown area can also be composed in this layer. The visit plan (a chain of destinations) may be rescheduled.

3. Akashi pedestrian bridge accident revisited – as a specimen case allowing a postmortem through simulation

As a key project of this research initiative, the authors focus on the Akashi pedestrian bridge accident in Japan as a specimen. The Akashi stampede accident occurring immediately after the end of the July 21, 2001 firework festival held in the seaside city of Akashi, resulted in a tragic catastrophe with 11 people dead and 247 injured. The disaster was caused by significant densification of the crowd due to two opposing crowd flows meeting on Asagiri pedestrian bridge as they moved between the nearby Asagiri station and the venue.

Early coverage of newspapers speculated on a domino effect caused by panic, but the official accident report published in January 2002 found no evidence of this and determined the accident be mainly due to a crowd avalanche or crowd quake that only occurs with a density of 10 or more pedestrians per sq.m. The report also pointed out the inadequacy of safety measures. After a legal dispute, in August 2005 the victims won a victory in court; the ruling ushered in a new era for Japan where prevailed, and in our countryan event organizer's responsibility for management can now be questioned.

The official report states the following. The width of the pedestrian bridge is 6 meters, whereas the width of the staircase section connected to the west side of the angled section at the southern end of the bridge is 3 meters. More stationary firework sightseers were on the south side of the staircase. Victims were concentrated on the southern end of the pedestrian overpass. As the shape of the pedestrian bridge formed a bottleneck, this zone became   densely packed with a density of some 13 people per sq.m. resulting in a crowd avalanche, which was ascertained as the main cause of the accident (see the right of Fig.3).

Furthermore, the crossing of the opposing flows was listed as the first of the trigger points leading to the crowd avalanche. "Sometime a little before the fireworks, a flow was returning to the station of the people watching fireworks on the stairs and pedestrian bridges. This flow revolving from the staircase to the east side of the pedestrian bridge compressed the other flow from the station to the venue in a way that pushed pedestrians over to the west side.
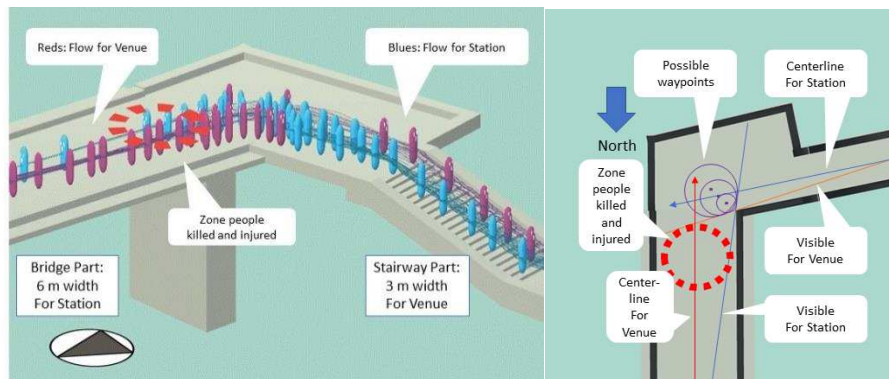
The flow toward the venue and the flow toward the station met and pressed against each other on a line diagonally heading north-east from the corner of the pedestrian bridge and the staircase.[2]

Based on this report, the authors had already conducted a simulation analysis [3]. By composing a model of the personal spacing layer in the two-dimensional space, the findings on the local densification at the southwest end of the pedestrian bridge were confirmed, and prevention measures drawn up (route separation on the bridge) for crossing flow conditions of low or medium density (with no effect of the physical body pressure, about 3 or less pedestrians per sq.m. density).

Subsequent research on the Akashi accident focused on high-density situations; mainly the stud wave phenomena [4]. In addition, simulation research to explore trigger conditions of accident occurrence in high-density situations has also progressed.

In our research project with such, rich data, we will plan a postmortem, or an ex-post verification through the composition of a multi-layer model.
Especially the aim is to establish whether the crossing of flows that occurred in the low and medium density situations was the accident trigger? In the official report, there is an estimate of the numbers of inflows and outflows of the pedestrian bridge with the average density rapidly increasing from 2.92 to 7.56 pedestrians per sq.m. from 19:00 to 19:30. In our research, using this estimate as a prerequisite, we attempt to reproduce the cross-flow in the low and medium density situation by simulation of the density rise process at the south end of the pedestrian bridge (accident occurrence zone).

Fig. 3. Akashi pedestrian bridge accident, July 21, 2001

At that time, in existing personal spacing layer models, to investigate if the position of the waypoint of the case played a major role in the composition of the crossflow, we apply a prototype application of the spatial cognitive layer model with three-dimensional space representation. Since the waypoints are in a visible relationship, endogenization of waypoints is one technical subject in our integrated model. Thus, this research project deals with the integration of three layers of the five layers mentioned in the previous section: the physical layer; the personal spacing layer; and the space cognition layer.

4. Vision-driven agents –the current position of spatial cognition layer model development

Here, to explain the current stage of our research project, the authors refer to vision-driven agents, which have been developed as an example of the spatial cognitive layer model.

A vision-driven pedestrian agent behaves based on information in their field of vision. It became known as a model class dealing with pedestrians' natural movement by EVA by Turner & Penn [5]. This natural movement is a behavior in which a pedestrian does not have any destination and only depends on information that appears in their field of vision, giving rise for example to such behavior s as free-mode exploration when entering a space for the first time. In the model, the vision field is divided into a plurality of Bins, and a Line-Of-Sight (LOS) segment represents each Bin is used as information. In other words, the distance to an obstacle at each viewing angle is used as information. As a behavioral algorithm, it consists of straight walking and turning direction selection, but variations are known depending on the direction change conditions and turning direction selection method [6].

Number	Exit and entrance name	Flowin	Flowout
1	City	20	24
2	Escalator	10	8
3	Station	20	24
4	Left shop	10	8
5	Central shop	10	8
6	Right shop	10	8
	total	80	80

Pedestrian traffic trace

Layout that interferes with visibility and passage

Layout that interferes with passage
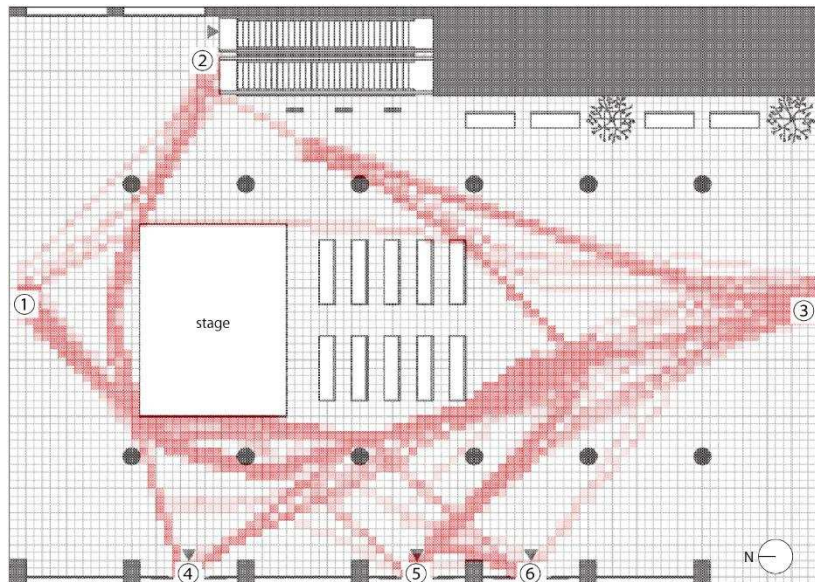
Obstacle

Research area

Fig. 4. Trails of the simulation results (Base case)

On the other hand, the navigated movement combines the vision-driven agent mechanism with destination information. Various technical methods have been reported according to the internal structure of the agent and the balance between free-mode exploration and destination orientation [6].

The vision-driven agent used here starts firstly with a natural movement mechanism, and then whenever it finds a direct destination, it moves straight towards that point. Besides, the direction change condition is set for each step probability 1/3, and the turning direction selection is set to the Longest Line-Of-Sight method. This agent model is simple and running in two-dimensional space.

(Base Case) The plaza is a rectangle, there are six entrance/exits, of which five entrances are facing the inside of the plaza and easy to see. Only the west-side escalator entrance/exit ② is facing the north, and it is difficult to see, so ② origin and destination flow lines draw arcs. Except for the cases (③ → ⑥) where the destination cannot be visually recognized and returned, the redundancy of the walking behavior is generally small (see Fig.4).

(Stage Case) A case where one stage is located on the north side of the plaza, and ten sets of long chairs are arranged on the south side. In the trips between ① and ③, the majority of agents bypass the west side. Also, at ② → ③ trips, they go round the stage counterclockwise. There are no agents passing through the stage and chairs which form a "mass." The other trends were the same as in the base case. The spatial cognitive layer model calculates direction change points (waypoints) for a passerby and those zones where they do not want to walk, whereas the personal spacing layer model is trying to follow a given set of waypoints. In our further research, we are planning to develop an integration model through the introduction of the third-dimensional space (see Fig.5).

Fig. 5. Trails of the simulation results (Stage Case)

5   Concluding remarks

In this article, we explained a vision of our research initiative aiming at an artificial society approach for dense space management. We also explained a five-layered framework for integration of pedestrian agent models to secure our initiative. We referred to the 2001 Akashi pedestrian bridge accident, as a specimen case for our flagship project, in that the three layers of Physical interaction, Personal spacing and Spatial cognition were covered. We also showed vision-driven agent simulations as our current position. The Akashi accident has many things in common with the Love parade accident, and the following implementations of 'Venue Capacity Evaluation' and 'Prediction of Visitors' are future tasks. An agent's communication layer and a Higher intelligent function layer are required for completing the artificial society approach. We have studied modeling research on 'information-based behavior' in tourist emergency evacuation [7], and a modeling research on a dynamic rescheduling function [8]. However, an integrated modeling with BDI model in mind is also a future task.

Acknowledgement

References

1. Helbing, D., & Mukerji, P.: Crowd disasters as systemic failures: analysis of the Love Parade disaster. EPJ Data Science, 1(1), 7 (2012).
2. Akashi City Summer Festival Accident Investigation Committee: Accident Investigation Report on the Firework Festival at the 32nd Akashi City Summer Festival, (2002, In Japanese).
3. Kaneda, T., & Suzuki, T.: A simulation analysis for pedestrian flow management. In Agent-Based Simulation: From Modeling Methodologies to Real-World Applications (pp. 220-232). Springer, Tokyo (2005).
4. Kaitsuji, M.: In order to enjoy the event with peace of mind. Kodansha editorial, Tokyo (2014, in Japanese).
5. Penn, A., Turner,T.: Space syntax based agent simulation. M. Schreckenberg and S. Sharma (Eds.), Pedestrian and Evacuation Dynamics, 99–114 (2001).
6. Kaneda, T., Morita, Y., Ohashi, S., & Mizuno, T.: Agent Simulation to Assess the Influence on Pedestrian Movement of the Lay-out of Furniture in an Open Space. CUPUM 2017: Planning Support Systems for Resilient and Smart Urban Futures (15th International Conference on Computers in Urban Planning and Urban Management) (2017).
7. Sonohara, M., Sakai, K., Takahashi, M., Kaneda, T.: A Study on Agent Modeling of Tourist Evacuation Behaviors in an Earthquake: A Case Study of an Evacuation Simulation of Himeji Castle. Kurahashi, S., Takahashi, H. (eds): Innovative Approaches in Agent-Based Modeling and Business Intelligence. Springer, (forthcoming).
8. Kaneda, T., Yoshida, T.: Simulating Shop-Around Behavior. International Journal of Agent Technologies and Systems, 4 (3), 15-28 (2012).

Combining agent-based simulation and road traffic emission calculations to evaluate the impact of land-use and infrastructure changes on air pollution in Beijing

Liu Yang[1], Lufeng Zhang[1] and Koen H. van Dam[2]

[1] University of Chinese Academy of Sciences, 100190 Beijing, China

[2] Imperial College London, London SW7 2AZ, UK

`k.van-dam@imperial.ac.uk`

Abstract. Air pollution from personal motorised transport has a big impact on quality of life in cities, as a function of human activities, land-use and the transport network. In this paper we simulate a neighbourhood in Beijing where an existing train line is being decommissioned, offering the potential to develop new public spaces, houses and offices, as well as new road infrastructure. Using an agent-based simulation model which generates activities for a synthetic population the baseline and two alternative scenarios are compared, illustrating the impact on local air pollution by translating the traffic volumes to emissions. This chain of models shows how different urban designs can be successfully evaluated, giving planners and local authorities insights in the effect of different plans.

Keywords: Land-use, transport, infrastructure, air pollution calculation, agent-based model, linear park

1   Introduction

In large cities, such as Beijing, transport is a dominant contributor to urban air pollution and a major risk factor impacting public health, not only due to the large quantity of its emissions but also because such emissions are close to humans, with much of the population exposed to dangerous pollutants [1] such as nitrogen oxides and fine particulate matter [2]. Road traffic emissions account for two-thirds and one-half of nitrogen oxides (NO_x) and total hydrocarbon [3]. Urban planners need to take this into account when designing or updating existing neighbourhoods, and evaluate people's exposure to dangerous pollutants such as nitrogen oxides and carbon dioxide and quality of life resulting from their design choices. To support this process, simulation tools can evaluate different scenarios against a range of performance indicators and targets. Given that physically changing an actual system would be costly and disruptive, building a model as a representation is economical and efficient [4]. Urban models aim to capture the complexity of cities, linking theory and practice [5]. In particular, agent-based models are helpful in testing alternative configurations of a system to support decision-making through the evaluation and prediction of various policy alternatives and impact on decisions by individual agents [6].

In earlier work [7] we simulated a synthetic population in Beijing, but had simplified the emission calculation and only considered the baseline. The contribution of this paper is a combination of the agent-based simulation model of the urban environment with an air pollutant calculator to generate insights on a number of urban designs. Next, the methodology is presented followed by the case study in Beijing.
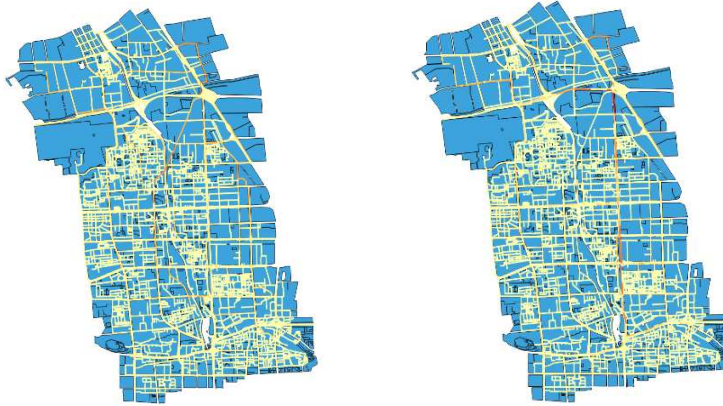
2   Methodology

In this research we link an agent-based model of a synthetic population of a sample district with the multi-spatial description of the transport infrastructure, land-use, and public open spaces in GIS. Based on the integrated design-modelling methodology presented in [7], we use a responsive design process of "simulating the existing system—urban design—simulating different design scenarios—evaluating their impacts on air pollution—redesign and decision-making". Firstly, the status quo of the road network and land-use (consisting of residential areas, workplaces, public open spaces, etc.) is represented at a self-defined sampling segment in QGIS as two separate layers with buildings/points of interest and a road network.

Using the GIS data, an agent-based model generates a synthetic population based on the socio-demographic data from the 6th population census of Beijing (2010). The agents travel for a workday, following their activity patterns, select destinations and deciding the time duration on different types of land-uses. In terms of route choice, the original model, which uses a shortest path algorithm, was upgraded by allowing the agents to consider the different travel speed on each road leading to a quickest path algorithm. Agents then react to the different urban plans and available transport infrastructure. The model yields result of traffic volume on each road segment, which is transformed to amounts of pollutants—CO_2 and NO_x—in the COPERT software, one of the most widely used emission models [8].

Vehicle data as input for COPERT was based on a Beijing Vehicle Activity Study [9], setting the type of simulated vehicles as 90% passenger cars and 10% light commercial vehicles, the most of which consume gasoline unleaded fuel and are subject to the China 4 and 5 Emission Standards (equivalent to Euro 4 and 5 Standards). It is worth noting that COPERT only uses average speed to estimate hot running emissions, and takes congestion into account during model development, resulting in a limitation that users cannot change the level of congestion [10]. To compare the impact of different road and public space design scenarios on air pollution, we zoom into a smaller area because the indicators of city-scale air pollution are complex and cannot be simplified as private car emissions. Finally, the air pollution under different scenarios is visualised in QGIS for decision-making support.

3   Case study and results

The Jing-Zhang Railway was located in the north-west Beijing and is being replaced by an underground high-speed rail. Two urgent issues have therefore appeared: how to redesign the linear spaces formerly occupied by rail track, and how to reorganize

Fig. 1. Traffic volumes generated by the agent-based simulation model for a synthetic population in Beijing using the shortest route choice algorithm (left) and fastest route (right).

the transport network previously cut by the rail. Since air pollution has been a crucial issue in Beijing, this paper uses air pollutant emission as a key indicator of evaluating designs for transport infrastructure, public space and buildings. We chose an area within a radius of 3-km from the railway track to validate the accuracy of the model against measured traffic volume, using the census and land-use input also used in [7]. Fig. 1 shows the traffic volume results generated from the original model using shortest path algorithm and the updated model using quickest path algorithm which are closer to the real-time traffic volume data. This data on simulated movement resulting from the agent behaviour can then be used as input for the COPERT software.

Having simulated the baseline scenario of the existing transport infrastructure and land-uses, we zoom in to an area within a radius of 1-km from the railway, to compare different accessibility and connectivity of the road networks, kinds of redesigns of the railway track and public open spaces, and multiple land-use redistributions. In scenario 1 the railway track is transformed to a linear park, adding east-west road connections to join the road networks lying on both sides of the rail, inspired by [11]. By combining the park with existing linear open spaces, a public space network is created. Knowing that linear parks could raise the land value of the surrounding areas [12], they are redeveloped as residential areas and workplaces. Scenario 2 transforms the railway to a main road, adding east-west roads, and only exploits the underutilized linear spaces along the line. Fig. 2 shows NO_x pollutant emissions of the baseline plan (the railway is shown as a dash line), followed by two redesign scenarios.

4

Fig. 2. NO$_x$ distribution from the COPERT model in the three design scenarios: baseline (left), scenario 1 (middle) and scenario 2 (right)

4　　Concluding remarks

The results from the case study show how the combination of models, simulating the urban activities for a given infrastructure and land-use input and calculating emissions and pollutants resulting from the generated trips, can give valuable insights in the impact of design choices on local air quality. Early results demonstrate that changes in land-use and road network have a direct impact on trips resulting in different air pollution levels, though these outcomes need to be validated before any design decisions could be made. With the same approach, other traffic simulation tools could be employed to take advantage of route planning algorithms with also include, for example, congestion levels, though for the purpose of this work it is essential to also include the activities themselves as they are dependent on land-use and infrastructure changes so the effect of urban planning decisions can be explored.

Next steps in our research are to feed the generated air pollution back into the agent-based simulation model and update the location and route choice algorithms to take differences in air pollution into account, for example allowing pedestrian agents to choose parks for leisure activities and routes with less pollution for their commute. This would be particularly relevant when incorporating mode choices and active transport modes. Exposure to air pollution by individual agents can this way be studied too, which allows comparisons between socio-demographic groups. Finally, developing urban plans in close cooperation with architects will allow testing this approach directly with potential users of the tools while comparing the outcomes with detailed measurements taken on the ground for the baseline scenario can aid further validation of the models. This bottom-up approach can then lead to decision-support tools embedded in the design tools of urban planners and transport planners and give direct feedback to high level plans by developers including local authorities.

References

1. Fenger, J., 1999. Urban air quality. Atmospheric Environment 33, 4877–4900.
2. Su, J.G., Apte, J.S., Lipsitt, J., Garcia-Gonzales, D.A., Beckerman, B.S., de Nazelle, A., Texcalac-Sangrador, J.L. and Jerrett, M., (2015). Populations potentially exposed to traffic-related air pollution in seven world cities. Environment international, 78, pp.82-89.
3. Oliver, Hongyan H., Kelly Sims Gallagher, Mengliang Li, Kongjian Qin, Jianwei Zhang, Huan Liu and Kebin He, "In-use Vehicle Emissions in China: Beijing Study" Discussion paper 2009-05, Cambridge, Mass.: Belfer Center for Science and International Affairs, May 2009.
4. Law, A. M., & Kelton, W. D. (1991). Simulation modeling and analysis (2nd ed. Vol. 2). New York: McGraw-Hill.
5. Batty, M. (2009). Urban modeling. International Encyclopedia of Human Geography. Oxford, UK: Elsevier.
6. van Dam, K. H., Nikolic, I., & Lukszo, Z. (2013). Agent-Based Modelling of Socio-Technical Systems, Springer.
7. Yang, L., van Dam, K.H., Anvari, B. and de Nazelle, A., (in print) Simulating the impact of urban transport infrastructure design on local air quality in Beijing, Social Simulation for a Digital Society: Applications and Innovations in Computational Social Science, Springer, to appear in 2018.
8. Gkatzoflias, Kouridis, Ntziachristos and Samaras (2012) COPERT 4 – Computer programme to calculate emissions from road transport, European Environment Agency.
9. Huan, L., Chunyu, H., Lents, J., Davis, N., Osses, M., & Nikkila, N. (2005). Beijing Vehicle Activity Study. Final Report, International Sustainable Systems Research, California, USA.
10. Smit, R., Brown, A. L., & Chan, Y. C. (2008). Do air pollution emissions and fuel consumption models for roadways include the effects of congestion in the roadway traffic flow? Environmental Modelling & Software, 23(10-11), 1262-1270.
11. Bahrami, F., & Rigal, A. (2017). Spaces of effort, exploration of an experience of mobility. Applied Mobilities, 2(1), 85-99.
12. Cervero, R. (2009). Transport infrastructure and global competitiveness: Balancing mobility and livability. American Academy of Political and Social Science, 626(1), 210-225.

Computational testing of urban planning policies performance: a case study of the Land Use Compatibility Matrix of Ensenada, México. [*]

Javier Sandoval-Félix [1][0000-0000-0000-0000], Manuel Castañón-Puga [1][0000-0003-2890-512X]
Carelia Gaxiola-Pacheco [1][0000-0000-0000-0000], and Eugenio Dante Suarez [2][0000-0000-0000-0000]

[1] Universidad Autónoma de Baja California, Tijuana, B.C. 22390, México {javier.sandoval10, puga, cgaxiola}@uabc.edu.mx
http://www.uabc.edu.mx
[2] Trinity University, San Antonio, TX 78212, USA
esuarez@trinity.edu
http://www.trinity.edu

Abstract. This paper shows an institutional effort to evaluate urban policies from a complex system approach, focusing on the rise of land-use specific attraction zones. These zones result from agglomeration processes permitted by an Urban Development Program that by doing so may inadvertently attract unfavorable urbanization. Unlike physical land uses, attraction zones appear, grow, diminish and rise again over time in the form of attraction pools. Understanding these phenomena is vital for policy makers.

For this evaluation exercise, an Agent-Based Model was designed to find industry-related attraction zones. Because they may rise regardless of land use regulations, public officials expected to identify places with a probability for law infringement or even discover unexpected opportunities and adjust public policies accordingly if that may be the case.

Findings indicate that the Urban Development Program is ill-suited regarding current industry development expectations. It also showed the roles that population density, topography, roads network, and residential land use play in Attractive Land dynamics. Lastly, the process of designing and running the model acted as an unexpected thought-provoking process for policymakers, as it helped them to understand and acquire new knowledge about Ensenada's industry phenomena.

Keywords: Agent-Based Modeling, Urban Simulation, Public Policy Analysis.

1   Introduction

An outstanding characteristic of the city of Ensenada, Mexico, is its spatial distribution of industrial activity. A series of historical factors ranging from non-compliance with urban regulations, the lack of primary roads and the cost of land have determined general dissemination of this activity, generating a low competitive profile of this sector compared to other cities in the region. At the same time this distribution has created severe urban problems like land use incompatibility, the invasion and damage to roads, long transit times and street obstructions among others. In a recent project called Sectoral Program of Industrial Development of Ensenada [1], it is estimated that 48% of the industries are in normative incompatibility with the Urban Development Program of Ensenada (UDPE) [2], 87% of the properties that host this activity are in lots less than one hectare contrary to the recommendation of 10 has [3] and 59% of this activity is on local roads. Together with the abrupt topography that characterizes Ensenada, has resulted in containing the lowest percentage of industrial use of all the cities of the state with 1.45% of the total urban area. Despite the current unfavorable spatial distribution of industrial activity, it is noteworthy that the city has had three iterations of its UDPE where clear regulations for the orderly distribution of land uses are established, but the industry has not occupied the territory as planned, but settling in certain areas stated as incompatible by the Program.

In this 2018, the municipal government began updating the UDPE, so one of the exercises that it carried out was to evaluate the pertinence of adapting the regulations regarding industrial land uses. To do this, it was considered necessary to know the probable scenarios of the spatial distribution of the industry if the application of current regulations were to continue as the city grew. Due to the fact of anarchic distribution of this use and the high demand for adequately sized and located industrial spaces, public officials concluded that the urban regulatory framework was not sufficiently established to face the urban challenges that this relevant activity means. Given the numerous future spatial distribution scenarios of the city as a consequence of a high number of compatibilities regulated by the UDPE, is that decision-makers wanted to adopt a methodology that minimized the uncertainty that always is present in a large number of results, and which of these scenarios should be taken into account in the event of a possible public policy modification. As an additional element, it was observed that each land use had specific location preferences. Given this high number of possible interactions between spatial distribution, compatibilities, and preferences, is that the municipal planning office decided to analyze the public policies of the UDPE from the perspective of complex systems, focusing on the issue of urban patterns through an agent-based modeling methodology. The objective of using this approach was to develop a model that would help in the identification and analysis of emerging industry-attractive vacant land as a result spatial agglomeration of favorable conditions for this use. Once these areas were obtained, called here Attractive Land, they were compared with current UDPE regulations to assess probable updates.

1.1 Complex Urban Systems

An important challenge in the context of urban development is how to design better public policies in an environment of uncertainty and slow progress in urban knowledge. In an institutional context where governments strive to order the city through urban development programs with a normative, static and top-down vision, with the particularity in Mexico of constant failure to comply with regulations, it is that there is a need to improve both the development processes of these programs and their implementation through knowledge that can shed new light on the functioning of cities. This perspective must recognize the city as a complex entity, conceiving it as a result of the interaction of a large number of environmental, social, economic and institutional components. This is the perspective of complex systems, which can be defined as a system in which large networks of components without central control and simple rules of operation give rise to complex collective behavior [4]. Although it has its own characteristics (Table **1**), this approach is not exclusive of a particular science but rather an inter / transdisciplinary exploration of nature in almost all scales and environments [5] challenging the traditional boundaries between physical, biological, ecological and organizational concepts, as well as between the human and natural sciences [6]. Since it can be applied to phenomena where a large number of interacting components are present, it has been used for the study of seemingly dissimilar fields such as markets [7], epidemics [8], social networks [9], traffic [10], crowds [11], public policies [12] and cities [13].

Table **1.** Characteristics, themes, methods and tools of Complexity.

Characteristics	Research topics	Methods	Tools
		Cellular automata	Computers
		Chaos	
	Self-Organization	Graphs	Modeling and
	Adaptation	Percolation	computational
	Emergence	Information theory	simulation
Anti-reductionist	Autopoiesis	Neural networks	
	System dynamics	Fussy logic	Programming
Holistic	Evolution	Genetic algorithms	languages
	Networks	Fractal geometry	
Multidisciplinary	Phase changes	Artificial life	Software
	Performance according to scale	Network analysis	
Non-linear	Learning	Multi-agent modeling	
	Information processing	Data mining	
Multiplicity of states	Path dependence	Artificial intelligence	
	Far from equilibrium dynamics	Game theory	
	Patterns	Numerical simulation	
	Segregation	Power laws	
		Machine learning	
		Monte Carlo simulation	
		Markov chains	

The complexity approach applies to different phenomena through similar topics like networks, self-organization, patterns, phase changes, inflection points and emerging processes have also been found in all of them. In the case of cities, complexity has functioned as a complementary approach to several urban disciplines, providing a theoretical basis to study and understand a variety of urban phenomena and properties that were previously considered independent of each other and interpreted by different theoretical bases [14]. The conception of cities under the complex systems approach is called Complex Urban Systems. It overcomes the limits of traditional urban planning (**Fig. 1**) since it unites aspects of cities and complex systems, where population an economic growth, the increase in levels of urbanization, the dependence on infrastructure and the role of technology in society are considered as elements that drive urban complexity [15]. This approach achieves a solid theoretical basis for a variety of urban phenomena that until then were considered independent from each other, in addition to new revelations about our understanding of cities that are a reflection of the basic properties of complexity [14]. What initially began with issues such as cities in equilibrium, scale laws, networks, exponential growth, logistic growth, non-linear behavior, chaos and bifurcation [13], has now advanced towards the conception of cities with evolutionary behavior and organization that change and adapt [16], the classification of cities according to power laws [17], the identification of a genetic code in cities based on functional and spatial structures [18] and cities as a hyper-network or "system of systems of systems" [19]. A sample of the subjects in which the city has been studied under this approach and that have brought new urban knowledge is shown in **Table 2**.
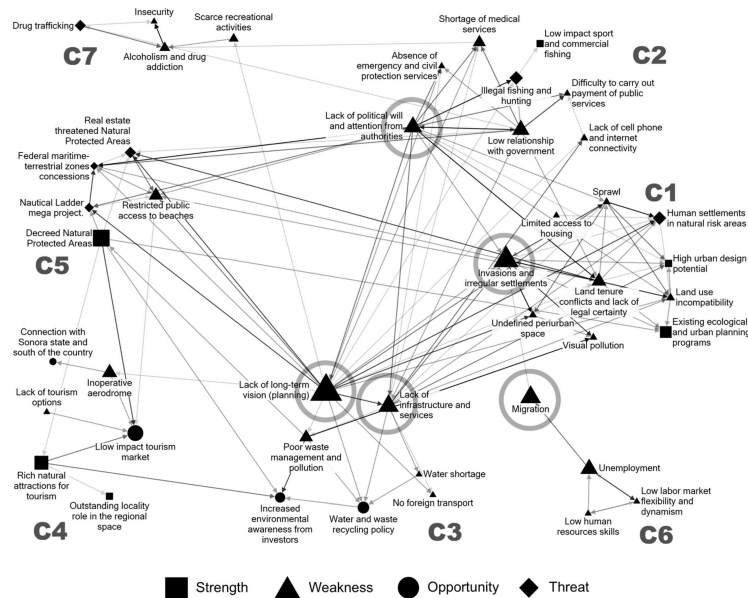

Fig. 1. Interaction between Strengths, Opportunities, Weaknesses, and Threats to the urban development of Bahía de Los Ángeles, Ensenada identified through network analysis [20]

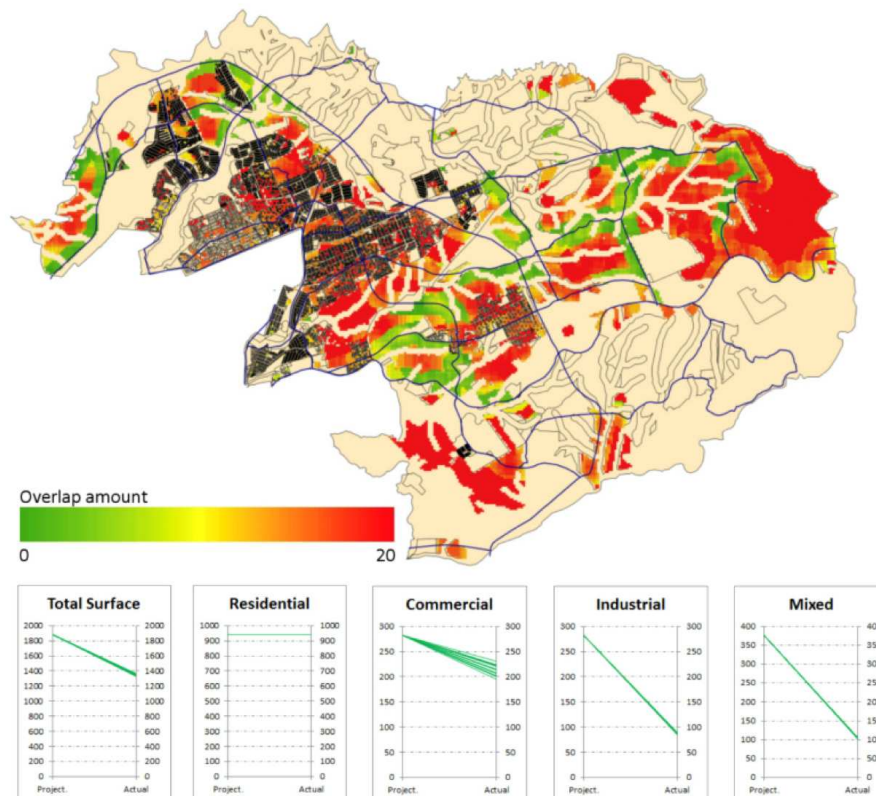**Table 2.** Complex urban systems research topics and methods.

Complex Urban Systems research topics	Methods
Land use spatial patterns	Cellular automata Agent-Based modeling
Social group segregation	Cellular automata
Economic performance according to scale	Power laws
Communication and services networks	Network analysis
Social networks	Network analysis
Urban emergent processes	Cellular automata Agent-Based Modeling Neural networks
Development based on positive feedback	Network analysis Agent-Based Modeling
Social self-organization	Network analysis Agent-Based Modeling
Phase transition towards new states of the system	Montecarlo simulation
Transportation, traffic and congestion	Network analysis Graph theory
Pedestrian mobility	Network analysis Graph theory Percolation theory
Crowds	Network analysis Graph theory Percolation theory
Residential mobility	Cellular automata Agent-Based Modeling
Spatial location of commerce and services	Cellular automata Agent-Based Modeling
Health and infectious vectors	Network analysis
Land market and real estate interactions	Cellular automata Agent-Based Modeling
Informal settlements	Cellular automata Agent-Based Modeling
Sprawl	Cellular automata Agent-Based Modeling
City size distribution	Power laws
Urban planning processes	Agent-Based Modeling Network analysis
Urban risk	Agent-Based Modeling Network analysis
Economic spatial distribution	Cellular automata Agent-Based Modeling Network analysis

6

One aim of urban analysis from a complexity approach is to discover new ways of addressing urban challenges in a context of uncertainty, posing as an alternative to the traditional process of public policy design that is unaware of interactions with domains foreign to the problem originally intended to address and that will operate in a highly complex space beyond what can be controlled or observed in a deterministic way [5]. It must provide knowledge about cities by identifying emerging behaviors and anticipating how they can be affected by the implementation of policies before their application in real life, as well as identifying new relationships between urban structure and dynamics that allow more successful solutions and sustainable. Thus, forecasting, explaining and controlling would have relevant conceptual implications for those who design and implement public policies [21].

Given the large number of variables, states, interactions, relationships, and rules to be considered in a complex system, it is asserted that without the existence of computers the field of Complexity would not have emerged [22]. With the formalization of an urban system through a computational model, it is possible to do a series of exercises to understand the system and to predict its possible future states that for decision makers in urban matters can be of great importance. Since the 1960s, with an initial focus on issues of urban transport and land uses, the processing power of computers has been used to implement complexity research topics through methods such as cellular automata (**Fig. 2**), [23], [24], chaos theory, neural networks, fuzzy logic, genetic algorithms, fractal geometry, artificial life, network analysis and multi-agent models, contributing to the understanding of cities or project their future behavior [22]. The application of these concepts has already been varied, but circumscribing the theme to urban dynamics similar to those observed in Mexico, hybrid models of logistic regression, Markov chains, and cellular automata have been used to simulate population growth, representation of dispersion drivers with the use of geographic automata [25], representation of dispersed growth through links of macro-scale processes with micro-dynamic processes [26], probabilistic maps of sprawl [27] and its modeling in Geographic Information Systems [28]. Such has been the recognition of the potential of simulation that it has been classified as the third way of doing science in addition to inductive and deductive processes [29], and has come to be called as Computational Social Sciences [30]. The usefulness of these tools lies in the fact that questions of the "what if" type can be posed as a way to analyze public policies [31] where they can be tested in-silico with different public policy options in a practical and economical way approaching its effects, effectiveness, and costs [16]. In a limited budget and resources environment where large urban projects become prohibitive because of their high cost, surgical and selective intervention becomes essential. Under the light of Complexity, it is glimpsed then that one of the keys of the administration of interacting and complex systems is that solutions are not imposed per se, but are extracted from the same system, to guide it towards one of the available states [31].

So, the usefulness of complex urban systems approach through computational simulation, apart from understanding and prediction [32], also lies in planning. On the one hand, it helps understand systems through characterization of its behaviors and thus identify essential properties [5]. Even the process of constructing the model can be as

enlightening as the use of the same model because it reveals what we know about its connections and causalities [33], and with this knowledge, we can design better public policies. On the other hand, it can be used to forecast unexpected scenarios that could result if public policies are applied. Regarding the latter, although the modeling does not indicate with absolute certainty the effects of a policy in real life, knowing the variety of these effects would ensure a better understanding of the whole as possible and thus ensure more robust public policies [5].

Fig. 2. Result of an urban simulation to identify urban growth impacts in Ensenada, México [23]

Even so, and despite the potential usefulness of the approach of complex urban systems and the increasing accessibility to high processing power computers, there are few examples of application in the area of public policy generation since there is not yet a coherent and unified set of concepts and tools for planners and formulators of public policies [12]. Moreover, although there are substantial advances in these fields, most of the studies are for urban contexts unrelated to the urban dynamics of Mexico, where the generation of this knowledge is still relatively new [34], literature devoted to Latin American cities is scarce, mainly focusing on squatters, so these studies are still considered "in their infancy" [35].

Urban patterns. The spatial rise, displacement, diminishment and rise again of Attractive Land pools relates to one of the most distinct aspect of complex systems: emergence. Defined as the rise of new and coherent structures, properties and patterns during the process of self-organization [36], it relates in turn to pattern formation, a topical area of complex systems.

In spatial terms, urban processes imprint distinctive patterns. The way in which these footprints physically manifest (form, arrangement, location, distribution, size, grouping, dispersion, density, connectivity and variability in time, among others) expresses information in a coded form of the system that generates them [37]. This suggests that urban patterns can be used to infer about their generating processes. However, although there are methodologies to characterize spatial patterns, simulating their emergent behavior in ways that reasonably capture their behavior remains a significant research challenge [38] and there is still a great deal of uncertainty regarding their spatial implication. It is thought that by shedding light on this issue, better planning principles will be possible and understanding their behavior in cities will allow planners and public policy makers to improve urban interventions [39].

2 Methodology

When simulation tools are required, the complexity approach uses multi-agent systems to formalize and test hypotheses about different aspects of their dynamics [40]. For this reason, the methodology in this research relies on Agent-Based Modeling (ABM) which focuses on individuals and on explaining how results can be achieved through their choices and interactions, which makes it a good method for policy analysis [5]. The present ABM was developed through the Netlogo, a widely used multi-agent programmable modeling environment [41].

Because ABM is generally more difficult to analyze, understand and communicate that other analytical models, the methodology will be described through the ODD (Overview, Design Concepts, Details) protocol [42]. The intended benefits of this protocol are to improve the writing and reading of model descriptions and to better communicate the theoretical background and assumptions of a model.

2.1 Overview

Purpose. The purpose of the model is to identify the emergence of specific spatial conditions of interest as a result of aggregation behavior in an urban policy context. It demonstrates how the sum of allowed urbanization can result in an attractive environment for prohibited urbanization. The case in point is vacant land that can attract industry (Attractive Land) by meeting requirements established from a developer's perspective regardless of compliance to the UDPE. These requirements include closeness to regional roads, workforce, urban equipment and not adjacent to residential areas. Due to the fact that the UDPE is currently under revision, this model is designed to be a tool to support the analysis and design of urban development policies by timely identifying

zones where urban regulations may be infringed or even to discover unexpected opportunities, and adjust public policies accordingly.

Entities, State Variable and Scales. This is a grid based model representing the spatial extent of the city of Ensenada, México. Each cell of the model represents 10,000 m2 or 1 hectare. The surface of the city's polygon in which the model takes place is 45,652 hectares and is taken from the Urban Development Program.

The entities of the model are of spatial nature and represent the city's land use with variables that change over time. This takes place at the cell level, so 1 hectare is the basic spatial expression of this entity. The basic time unit is a year and the model is run in a 10-year span. These scales have been chosen because they are the ones used when designing and reviewing urban development programs.

Low-level variables. This are variables that cannot be deduced from other low-level variables because they are elementary properties of the entity. Each cell of land contains the following low-level variables (**Table 3**):

Land use. It represents the activity that that takes place in the cell. It can be of 7 general types, or it can also be vacant, which means that there is no current activity. The 7 general types are subdivided into 111 specific land uses according to the UDPE. The 7 general types are Residential, Touristic, Industrial, Agro-Industrial, Commercial, Equipment, and Farmland. For visualization purposes, Industry land use appears as orange cells, and the rest of land uses are gray.

Attractive Land. Attractive Land is not a land use but vacant land suitable for industry from a developer's perspective. As such, in contrast to land uses that once they urbanize a cell is stays the same for the rest of the simulation, Attractive Land rises and diminishes on places that are urbanized or that aggregate favorable conditions for industry.

City subsector. The city polygon is divided into 5 main sectors and subdivided in 61 subsectors according to the Urban Development Program (**Fig. 3**). The city limits polygon is greater than the current urban surface, so it also contains natural areas. Each one of the 111 specific land uses are compatible, or not, with each one of the 61 subsectors. Subsectors and compatibilities are taken from the Compatibility Matrix of the Urban Development Program.

Developable?. The city's polygon is composed of areas that are developable or non-developable. Non-developable areas represent streams, farmland, geologic faults, natural land, steep slopes and land 200 meters above sea level. The rest of the surface is developable land, including vacant lots inside the city. These areas are obtained from the UDPE.

Potential. Random value that changes over time. It refers to the potential of vacant land to be developed. If this value is greater that an established threshold, the vacant

land holding this value is urbanized and stays urbanized. Value is estimated according to the Vicsek-Szalay model that is later explained.

Fig. 3. Official Sectors and Subsectors plan of the city of Ensenada [2].

Intersected by a street?. It species whether the cell is intersected by a street/road or not. The Street/road network is obtained from the UDPE and includes current and projected network.

Body of water?. It specifies if the cell is water.

Table 3. Entities and variables

Type	Entity	Variables	Expresion	Changes over time?
Grid cell	Land	Attractive Land	Alive Dead	Yes
		Land use	Vacant	Yes
			Residential Touristic Industrial Agro-Industrial Commercial Equipment Farmland	No
		City subsector	S1-S11, C1-C9, N1-N12, Ch1-Ch10, M1-M19	No
		Developable?	Yes No	Yes No
		Potential	Numerical value stated by submodel	Yes
		Slope	0-5% 0-10% 0-15% 0-30% >30%	No
		Intersected by a street?	Yes No	No
		Body of water?	Yes No	No
Environment	Urbanization threshold	Potential threshold	Numerical value stated by submodel	Yes
	Land use neighbor preference	Adjacency Matrix	Desirable Not desirable	No
	Urban Development Program compliance	% Lawful industry	Numerical value stated by model user	No
	Population	Growth rate	Numerical value stated by model user	No
		Number of inhabitants	Numerical value stated by model	Yes
	Public policies	Land Use Compatibility Matrix	Compatible Incompatible	No
		Population density	Numerical value stated by model user	No
		Slope restriction	0-5% 0-10% 0-15% 0-30% >30%	No

Land Use Compatibility matrix. Matrix that specifies compatibility between each one of the 111 specific land uses and 61 subsectors, resulting in 6,771 combinations. The concept of compatibility is not related to the adjacency of land uses, but to the permission of a land use to take place inside a city subsector, and is obtained from the Land Use Compatibility Matrix of the UDPE. In this sense, the compatibility of each cell with each one of the 111 land uses will depend on the subsector is in.

Adjacency matrix. Matrix that species desired/undesired adjacency between land uses. It is designed from empirical evidence of land uses from the city of Ensenada.

There are also environmental variables that are not of spatial nature:

Lawful industry%. From the total of times the model tries to grow industry in incompatible zones, this is the percentage of times the model will allow it. This does not necessary mean that industry will be built, because the terrain must satisfy a series of additional requirements for this land use.

Years of development. The years the model will run.

Population. Number of city inhabitants. The model starts with a fixed population number according to recent census data, and is projected in a time span equal to *Years of development* variable. The amount of land uses that the model outputs depends on population size which in turn depends on the Population density variable.

Public policies. Actions to regulate urban growth. They are:
Population density: Inhabitants per hectare.
Slope restriction: Angle of terrain. Permitted slopes upon which urbanization can occur. Slopes are obtained from the UDPE.

Process Overview and Scheduling. In this section the environmental and individual processes are presented as a concise overview of the model.

Setup. This is the first step of the model where pre-established numerical and spatial layers of data are loaded that together form the initial urban state of the model.

Assign initial Attractive Land. Based on the initial spatial conditions of the model, it identifies the initial Attractive Land.

Assign random value to cells. The model randomly assigns a value of 1 or -1 to a cell variable called Potential to all cells that are inside of the city polygon. This will be the base for the Vicsek-Szalay submodel which determines if a vacant cell has potential for urbanization.

Each land use has location requirements, mainly its compatibility to subsectors of the city and its preference to neighbor land uses (**Table 4**).

Table 4. Adjacency Matrix

| | | Residential | Touristic | Industrial | Agro-Industrial | Commercial | | Equipment | | Beach |
						Industry compatible	Industry incompatible	Industry compatible	Industry incompatible	
Residential		O	O	X	X	O	O	O	O	O
Touristic		O	O	X	X	O	O	O	O	(2)
Industrial		X	X	O	O	O	X	O	X	O
Agro-Industrial		X	X	O	O	O	O	O	O	O
Commercial	Industry compatible	(1)	O	O	O	O	O	O	O	O
	Industry incompatible	(1)	O	X	X	X	O	X	O	O
Equipment	Industry compatible	(1)	O	O	O	O	O	O	O	O
	Industry incompatible	(1)	O	X	X	O	O	O	O	O

X No neighbors accepted   O Neighbors accepted

(1) As long as there are two or more residential neighbors

(2) As long as there are one or more beach neighbors in a radius of 5 cells

Table 5. Land Use Compatibility Matrix. Rows are land uses and columns are subsectors. Compatibilities are green, incompatibilities are red [2].

Estimate amount of projected urban Surface. Based on initial population, growth rate, and population density, and amount of urban surface is projected for one year. The simulation stops once the current year equals the Years of development variable.

Compare Current vs projected urban surface. Projected urban surface is compared to growing urban surface. While the growing urban surface does not match projected urban surface, model will loop the following steps. If this condition is met, it exits the loop, advances by one year, plots and spatial pattern of land-uses is updated, and starts a new loop by estimating new projected urban surface.

Update potential variable for each cell. All cells update their Potential variable according to the Vicsek-Szalay model. Additionally, urbanized cells rise the Potential variable of neighboring vacant cells. If the final value of this variable is higher than an stablished threshold, only cells meeting this requirement proceed to the next step.

Assign random land-use interest to cellss with potential above threshold value. One of 111 possible land-uses is randomly assigned as a land use interest to the cells above threshold. It is assigned according to pre-established probabilities and represents an interest of the cell to develop that land-use. The probabilities are based on the proportion of land uses stated as current by the UDP.

Validate the assigned land use with the Compatibility Matrix. The model verifies with the Land Use Compatibility Matrix if the randomly assigned land use interest is compatible with the subsector it's in. If its compatible, proceeds to next step.

Verify neighbors. The model verifies with the Adjacency Matrix that neighboring requirements by the land use interest can be met.

Urbanize cell with land-use interest. If compatibility and requirements are fully satisfied, the cell is urbanized with the land-use interest, and Land Use, Developable?, Suitability, Farmland and Urban cell variables are updated. The last three are high-level variables that are deduced from low-level variables previously explained. Suitability establishes if a cell can be urbanized taking into account public policies, developable and vacant land and if its inside the city polygon. Farmland states if the cell is used for agricultural activities and Urban if the cell is occupied by a land-use. A global variable called New Urban Surface is also updated each time a cell is urbanized.

Compare current urban surface with projected urban surface. If current urban surface has not reached the projected urban surface, the loop remains by starting a new *Update potential variable for each cell* step. If it does, exits the current year loop and proceeds to the next step.

Update Attractive Land. Based on the current spatial configuration of land uses, the model updates location and amount of Attractive Land.

Compare current year with Years of development variable. If current year is the same as Years of development variable, the model stops and the spatial pattern of land-uses is updated, along with plots and monitors.

2.2   Design concepts

Basic Principles. The basic principle of this model is that of cellular automata and how the state of a cell is determined by the state of its neighbors, and by environmental variables that represent public policies, population growth and compatibility rules established by an official Urban Development Program.

Emergence. Vacant land is going to be attractive to one or several land-uses depending on its location with respect to spatial patterns of land-use that surrounds it and to environmental variables. The behavior in quantity and spatial distribution of Attractive Land will vary over time. A cell of land is attractive to more than one land-use at the same time, so the quantity and location of Attractive Land for industry will diminish as it is taken by other land-uses but also arise in other zones where favorable conditions aggregate.

Interaction. Interaction in the model is between land-uses and land-uses with its surroundings. This interaction is based on the neighborhood concept. Depending on its type, land-uses can or cannot be adjacent to each other, so once a land-use is stablished it creates conditions around it that are favorable to some land-uses and unfavorable to others. Also, once a land-use urbanizes a vacant cell, it changes its *Urban, Suitability* and *Farmland* variables and raises the *Potential* variable of surrounding vacant cells.

Stochasticity. A land-use interest is randomly assigned to cells that its Potential variable is above a threshold value. The frequency on which a specific land use is randomly assigned depends on pre-established probabilities.

Collectives. Attractive Land spatially aggregate according to industry-specific requirements, forming particular patterns and land-use spatial distributions.

Observation. The model outputs an image of land-use patterns of industry, Attractive Land and general land uses. It also plots the following data:
Total Attractive Land
Attractive Land State
Industry location differentiated by compatibility
Urbanized Attractive Land differentiated by:
Compatibility
Land use
% urbanized

The steep topography that surrounds the city dramatically limits its growth and has had a historical influence on current urban shape. Regarding the model, topography has a

16

relevant influence on the parameter space of the final urban shape, which is rather low. An overlay of 500 final urban shapes was analyzed, and it was detected that by approximately the 50th overlay no noticeable differences were identified beyond this point, so for output analysis, an exercise of 50 runs was made.

2.3   Details

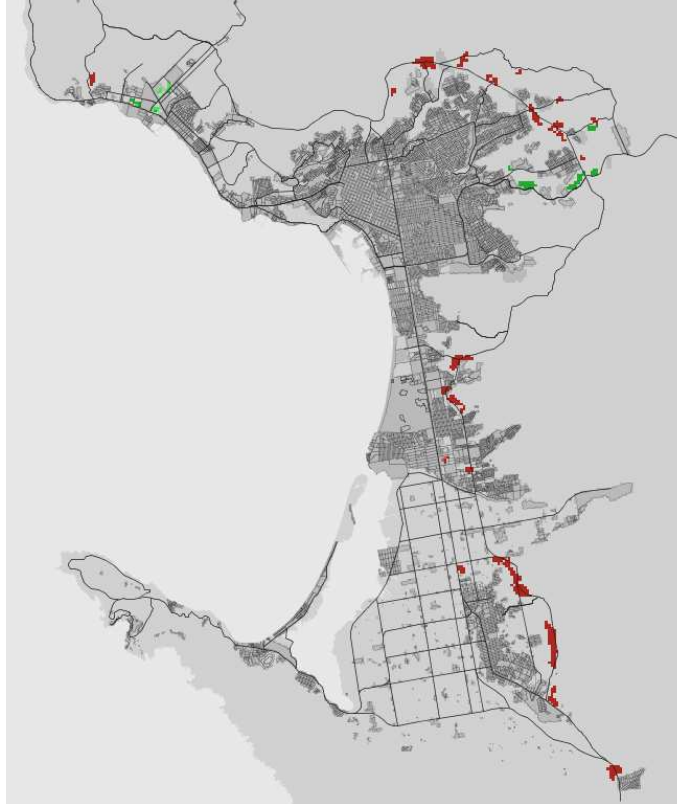**Attractive Land.** A cell in the model is considered as Attractive Land when it fulfills at least 7 of the following 9 requisites:

─ Land slope equal or lower than 15%
─ 0 neighbors of residential land use
─ 0 neighbors of touristic land use
─ 0 neighbors of commercial land use that does not want to be adjacent to industry
─ 0 neighbors of equipment land use that does not want to be adjacent to industry
─ A regional road within a 2-cell radius
─ 20 or more cells of residential land use within a 20-cell radius
─ 4 or more cells of industrial land use within a 2-cell radius
─ 2 or more cells of industrial land use in a 1 cell radius

A cell is no longer Attractive Land for industry once one or more of the following happens:

─ 1 or more neighbors of residential land use
─ 1 or more neighbors of touristic land use
─ 1 or more neighbors of commercial land use that does not want to be adjacent to industry
─ 1 or more neighbors of equipment land use that does not want to be adjacent to industry

Moore neighborhood is used when inspecting for neighbors.

Initialization. The model initially detects Attractive Land according to previously stated requisites (**Fig. 4**).

Fig. 4. Initial state of the model, showing the spatial extent of the city of Ensenada and Attractive Land for industry in compatibility (green) and incompatibility (red) with the UDPE.

Input Data. Initial values for simulation where as follows:

Initial population: 363,260. Taken from the most recent census data.

Growth rate*:* 1.7. Obtained from the UDPE.

Lawful industry%: 53. It was obtained from the Sectorial Industrial Development Program of Ensenada, which states that 47% of industry is located against urban regulations of the UDPE.

Years of development: 10. It is a time span in which it was considered by policy makers as adequate in a planning-revision context.

Population density: 35. It's the current Ensenada's average density according to the UDPE.

Slope restriction: 4. It corresponds to a slope range from 0 to 30%, which is close to current permitted urbanization on slopes not greater than 35% stated in the UDPE. The exact 35% value was not used due to technical challenges with available terrain data.

Threshold: 4. The value for the Vicsek-Szalay submodel was chosen after empirical observation of values that better approximated Ensenada's morphology.

Current land-uses, developable and non-developable land, slope, and street and road networks, including those that are on a planning phase, are loaded in the model. This data layers are loaded as shp (vector) or asc (raster) layers previously elaborated in a Geographic Information System software for the UDPE.

Submodels. *Vicsek-Szalay submodel.* This submodel updates all cells Potential variable. It is used to determine whether a vacant cell reaches the potential to be urbanized. For this to happen, the vacant cell potential must surpass an established threshold value. It is a cell based model that shows how highly ordered spatial patterns similar to those of urban development can emerge from a model that builds on spatial averaging [43]. This build patterns over time depending on a potential and threshold value. Potential can be translated to urbanization. This urbanization persists even if cell potential falls below the threshold value in the future. This pure averaging model is stated as

$$z_i(t+1) = \frac{1}{5}\left[z_i(t) + \sum_{j \in N_i} z_j(t)\right] + \epsilon_i(t)$$

It is a model that never moves to equilibrium because noise is constantly introduced into the system.

Besides the *Vicsek-Szalay* submodel, the Potential variable is also updated by neighboring urbanized cells as a way to replicate real-life situations where urban areas tend to urbanize adjacent vacant surfaces.

3   Results

Results are focused on analyzing the emergence of Attractive Land for the industry, as a result of agglomeration of land uses regulated by an extensive Land Use Compatibility Matrix established in an Urban Development Program. The objective was to identify where these areas arise and how their distribution and surface behaves over time. To do this, the model was developed to obtain the following data:
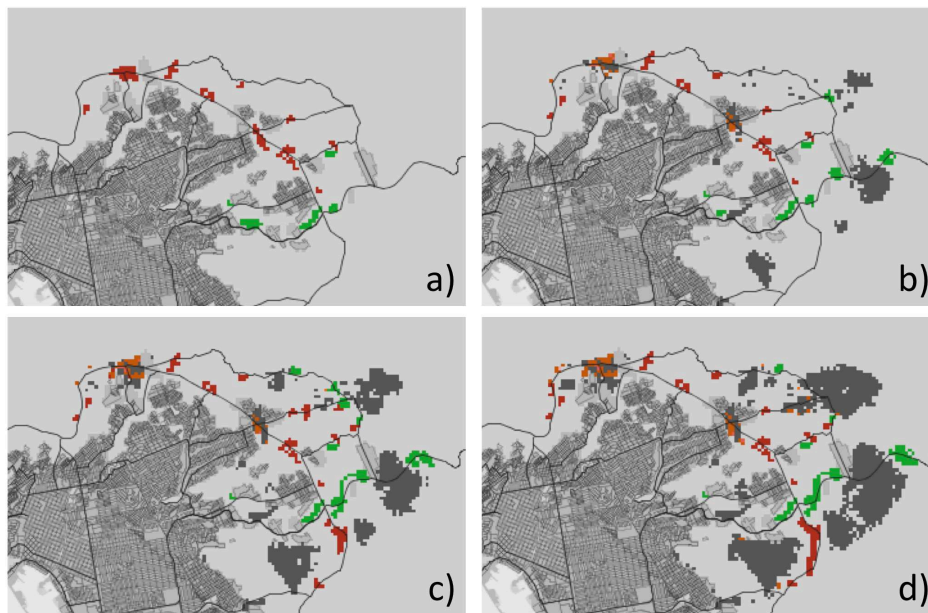   a)   Total Attractive Land.
   b)   Percentage of Urbanized Attractive Land
   c)   Urbanized Attractive Land differentiated by compatibility
   d)   Urbanized Attractive Land differentiated by land use
   e)   Industry location differentiated by location
   f)   Attractive Land State
   g)   Probability maps:
         a.   Attractive Land Formation
         b.   Urbanization by industry
         c.   Non-industrial land distribution
         d.   Alive or dead Attractive Land

Additionally, the previous data was also obtained under two scenarios: with variability in population density and variability in compliance with industrial regulations.
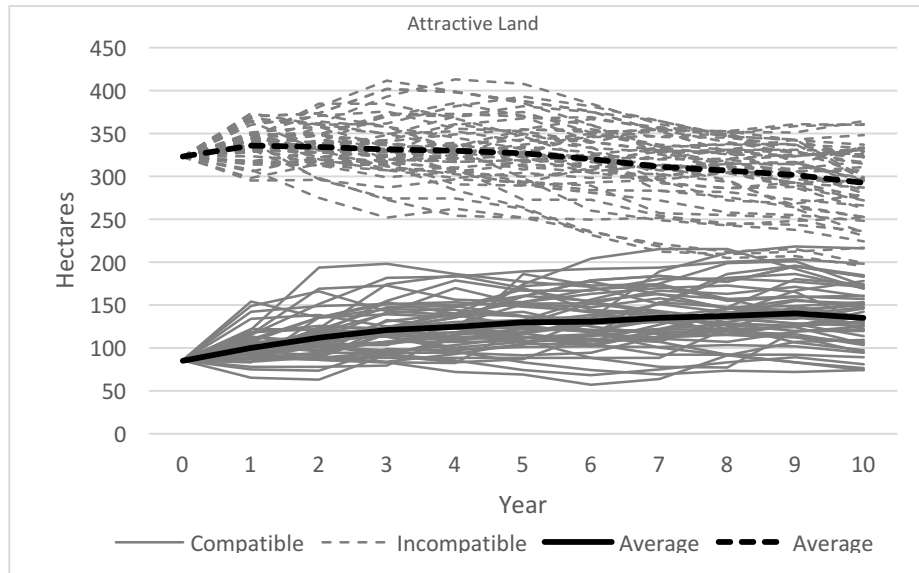
3.1 Total Attractive Land

This data measures the amount of Attractive Land in time (**Fig. 5**). It has always been considered by investors that the land available for industrial activity is very limited, so an interest in this research was to know what could be the availability of Attractive Land compatible with the UDPE. More important was to identify where could Attractive Land rise with incompatibility with the UDPE but attractive for this activity, and thus try to foresee possible normative violations.

In this aspect, it is observed that there will always be more Attractive Land in normative incompatibility than in normative compatibility (**Fig. 6**), but with a downward trend of the first and upward of the second until a time is reached when both types of Attractive Land descend uniformly. This shows that, as the city grows, the demand for industrial land use is higher than the zoning for this activity established by the city's UDPE. Something to note is that in the presence of low population densities between 10 to 20 inhabitants per hectare (**Fig. 7**), Attractive Land drops steeply in a period of between 4 to 7 years, probably due to the fact that in low densities the surface of urban land is higher and therefore the probability for the urbanization and disappearance of Attractive Land is greater.

Fig. 5. Evolution example of Attractive Land in a 10-year span. a) t=0, b) t=3, c) t=6, d) t=10. Green: compatible Attractive Land, red: incompatible Attractive Land, orange: industry.
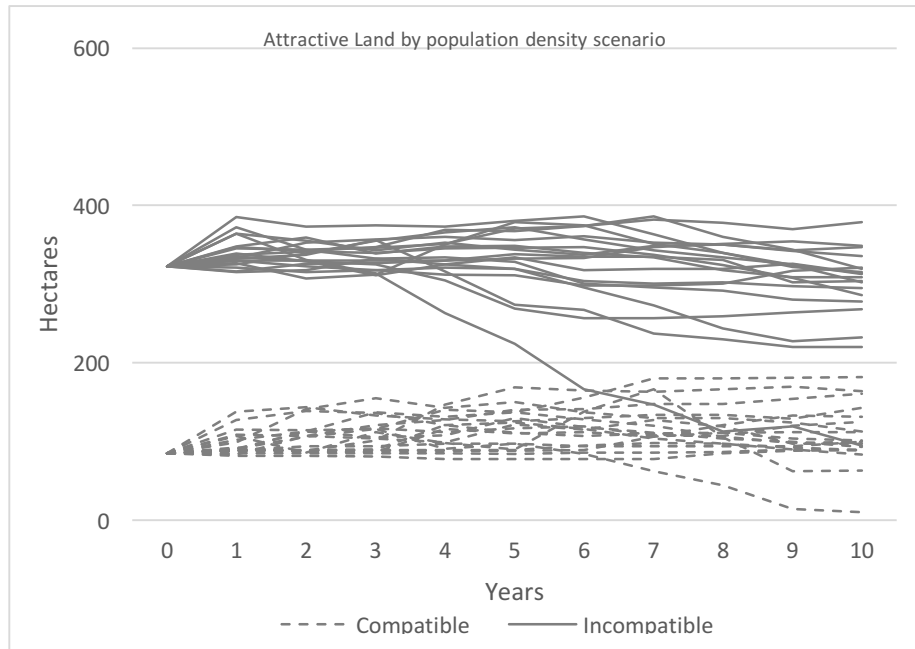
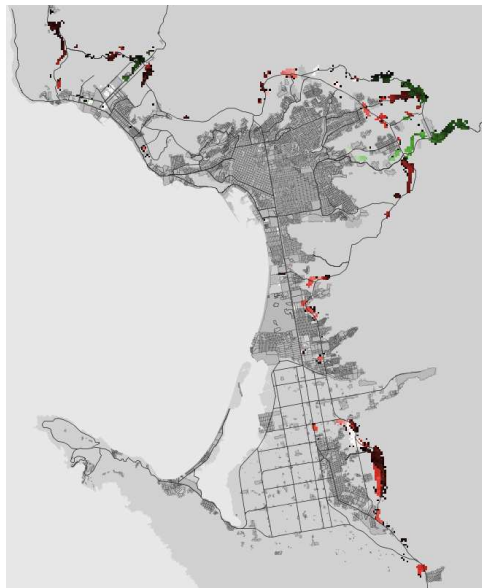**Fig. 6.** Behavior of Attractive Land in a 10-year span.

When studying the spatial average of all the executions of the model, it is observed the emergence of an Attractive Land corridor along a regional road that communicates the northeast part of the city (**Fig. 8**). This corridor presents a gradient that shows that the closer you get to the city, higher the probability of its formation. Together with the urbanization probability map that shows northeast growth, reflects a real-life city's situation where its growth in the last decade has been concentrated in that area. The limits of this corridor are only because the model is currently designed in such a way that it works only within official city limits, which means that there is a high probability that in real life it would extend beyond these limits. According to the Attractive Land State probability map, there is a possibility of its urbanization by other land uses and not only by industry.

The probability map about Attractive Land occupied by industry (**Fig. 9**) shows that this possibility is low in almost all parts of the city, even though there is ample opportunity for land occupation, mostly in incompatibility with the UDPE. There is a relevant exception to this condition in two zones, but to be able to develop them, the construction of regional roads is mandatory. It is observed that the development of these two zones located in the north and south ends of the city is not allowed by the UDPE despite that their location is desirable and contrary to the current situation where there are industrial uses in inner parts of the city.
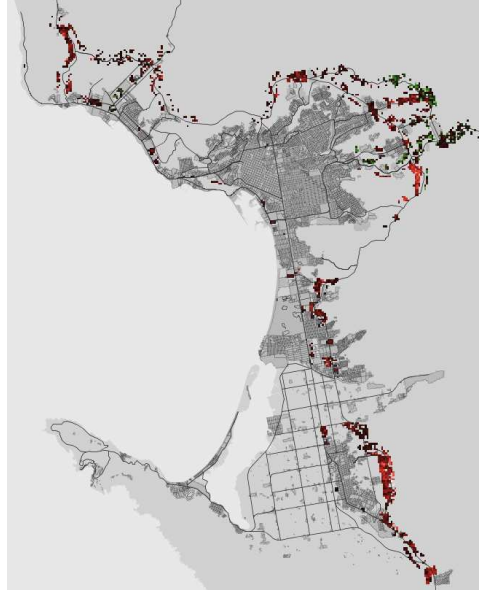
Fig. 7. Behavior of Attractive Land under population density scenarios in a 10-year span.

Fig. 8. Attractive Land formation probability map. Green: compatible with the UDPE, red: compatible with the UDPE. Darker shade: lower probability, brighter shade: higher probability.
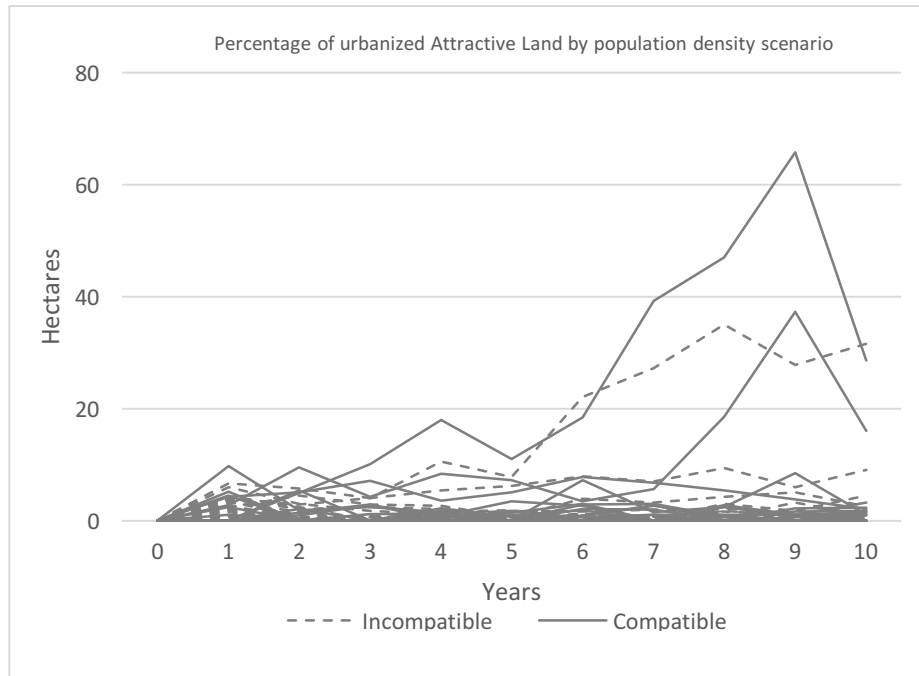
Fig. 9. Industry in Attractive Land probability map. Green: compatible with the UDPE, red: compatible with the UDPE. Darker shade: lower probability, brighter shade: higher probability.

3.2   Percentage of Urbanized Attractive Land

It outputs the percentage of Attractive Land that is urbanized in a given year. It was included in the model to know the dynamics of appearance-disappearance of Attractive Land and is complementary to the Urbanized Attractive Land output.

The behavior in each model run was so variable that it merited the individual analysis of each result, but relevant observations can be made in the context of density scenarios. In the presence of low densities of 10 to 15 inhabitants per hectare, urbanization of Attractive Land in a given year tends to increase by up to 65%, compared to up to 14% when 35 inhabitants per hectare, similar to Ensenada's density, was applied.

Fig. 10. Percentage behavior of urbanized Attractive Land in a 10-year span under different population density scenarios.
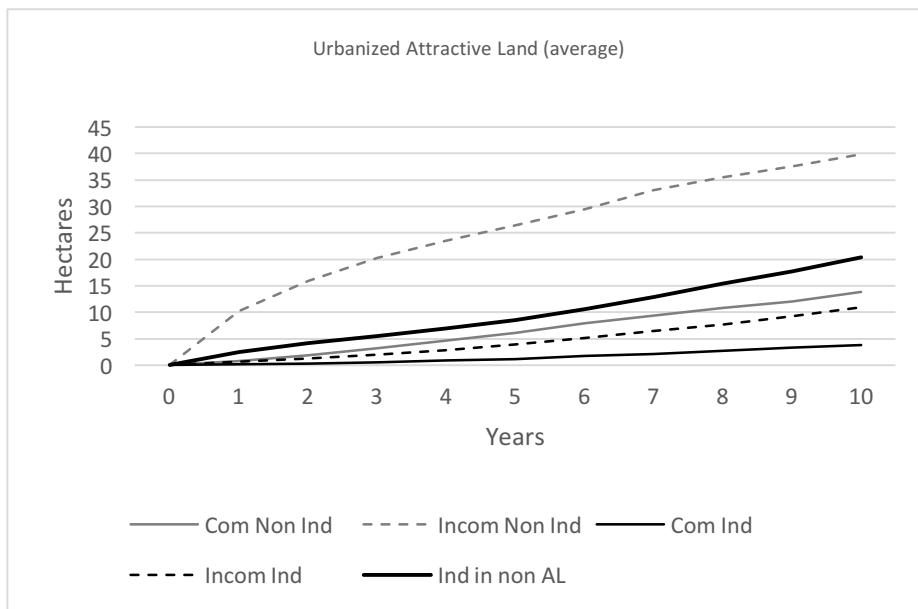
3.3 Urbanized Attractive Land differentiated by compatibility

This output quantifies the amount of Attractive Land that has been urbanized and that complies with urban regulations, but also identifies the one that does not comply and thus violates urban regulations. The intention of this indicator is to have support in the characterization of industry in normative incompatibility due to lack of appropriate industrial zoning or the lack of compliance with urban regulations, and in this way assess the need to add new zoning or complement industrial regulations.

Throughout the different executions of the model, it was observed that urbanization of Attractive Land by other land uses different to the industry predominates (**Fig. 11**), which reduces the opportunities for a favorable location. The next condition that predominates is industry outside Attractive Land, which shows that, although there is land in attractive conditions, the industry finally sits mostly on land where it is not normatively allowed. Although a land cost variable is not included in the model, it nevertheless relates to a real-life situation in which most of the current industrial activities sit in more unfavorable locations due to high land costs. Contrary to a desirable situation, and according to the model, the lowest occupation scenario was that of industrial occupation in Attractive Land with normative compatibility.
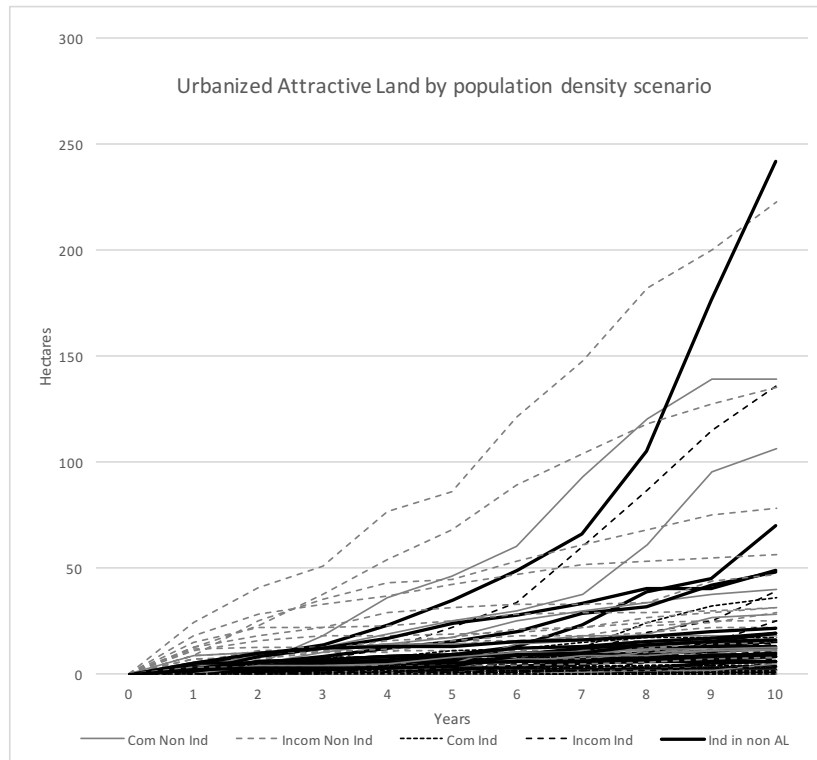
In a low population density context between 10 to 20 inhabitants/ha, it was observed that the urbanization of Attractive Land rises rapidly, and occupied by non-industrial

land uses (**Fig. 12**). The rise of industrial occupation outside of Attractive Land was also noted. It is observed that in densities of approximately 40 hab/ha or higher, the total industrial surface outside of Attractive Land is low which is a desired condition. This finding is significant since the current density of the city of Ensenada is 35 inhabitants/ha, which means that is currently at a threshold that, if it's surpassed by even lower densities, the industry will begin to appear more frequently in undesired places. This assumption is supported by the Sectorial Industrial Development Program of Ensenada [1] and the Partial Program of Urban Improvement of Downtown Seafront of Ensenada [44]. Between both programs, it is established that the downtown area has both a low density of 11 inhabitants/ha and a high presence of industrial use, even surpassing existing industry in peripheral areas that should be highly attractive in real life due to connectivity factors.

Fig. 11. Average behavior of urbanized Attractive Land in a 10-year span. **Com Ind:** Compatible Attractive Land urbanized by Industry, **Incom Ind:** Incompatible Attractive Land urbanized by industry, **Com Non Ind:** Compatible Attractive Land urbanized by other uses, **Incom Non Ind:** Incompatible Attractive Land urbanized by other uses, **Ind non AL**: Industry located outside Attractive Land.
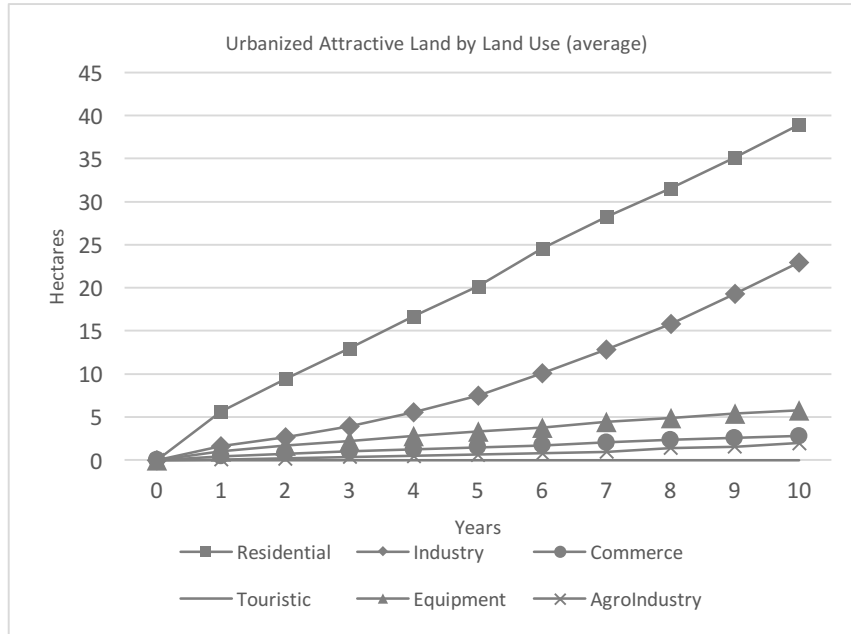
Spatially, there is a medium probability of Attractive Land urbanization in most of the northeastern sector of the city, and also in those northern and southern areas previously identified as having a high probability of forming Attractive Land but which is currently prohibited by the UDP.

Fig. 12. Average behavior of urbanized Attractive Land in a 10-year span under different density scenarios. **Com Ind:** Compatible Attractive Land urbanized by Industry, **Incom Ind:** Incompatible Attractive Land urbanized by industry, **Com Non Ind:** Compatible Attractive Land urbanized by other uses, **Incom Non Ind:** Incompatible Attractive Land urbanized by other uses, **Ind non AL**: Industry located outside Attractive Land.

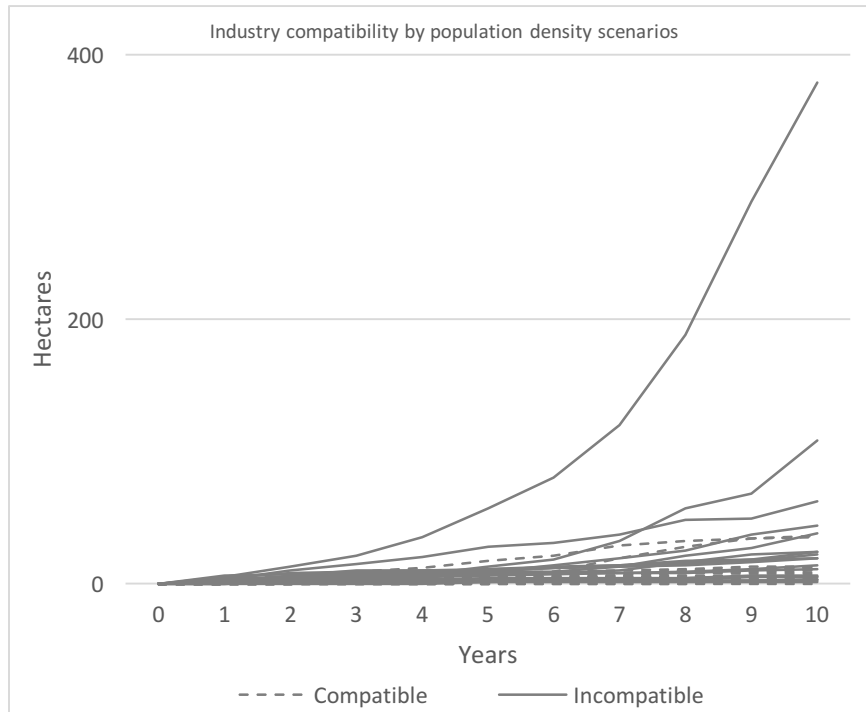### 3.4   Urbanized Attractive Land differentiated by land use

This output quantifies Attractive Land that has been urbanized, differentiated by type of land use. It is intended to understand if there is a competition for Attractive Land between industry and any other particular land use. Results outputted a current urban phenomenon of the city of Ensenada that is the diminishment of proper vacant land for the industry due to housing occupation. Traditionally it has been considered that the location of residential use has been an inhibitor in the growth of the industrial sector as they occupy zones initially intended for industry. The model outputs consistently, and with minimal variation between results, that housing is the land use that mostly occupies Attractive Land followed by industrial use (**Fig. 13**). The model replicates a clear competition for vacant land between these two uses that occur in real life, even recently reflected in relevant social activism.

Fig. 13. Average behavior of urbanized Attractive Land by different land uses in a 10-year span.

3.5 Urbanized Attractive Land differentiated by location

This output quantifies industrial surface that is compatible or incompatible with the UDPE. This indicator is intended to know the amount of industry that is located in subsectors that are not compatible with this activity. If the majority of industrial activity is outside compatible subsectors, it denotes either an insufficient industrial zoning or regulation breaches. The model showed that in low densities between 10 to 25 hab/ha, the amount of industrial surface in incompatible subsectors rises steeply in time (**Fig. 14**).
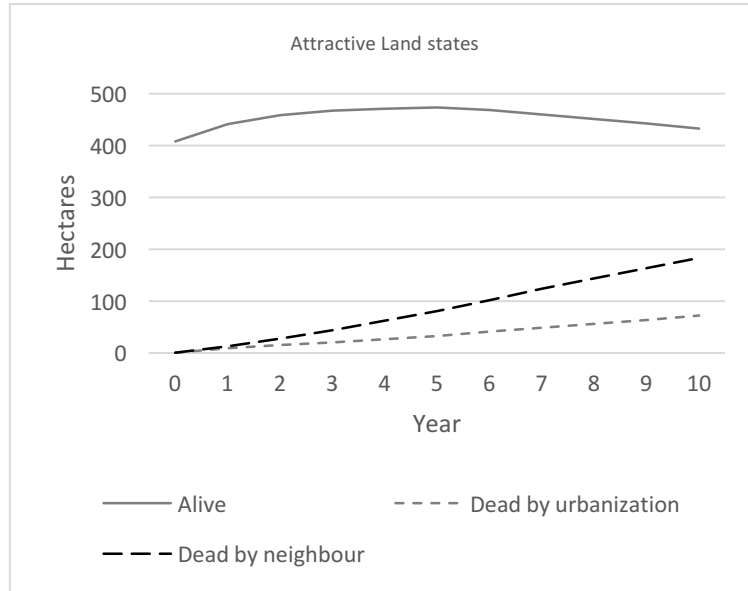
Fig. 14. Compatibility of industry under different population density scenarios in a 10-year span

In spatial terms, is of interest that the model outputted zones with a high probability of creating Attractive Land but conditioned to the noncompliance of the UDPE and the construction of a still projected road network.

3.6 Attractive Land States

This output quantifies the surface of Attractive Land that is "alive" (vacant and desirable) or "dead" (by urbanization or neighbors). For this part of the of study, Attractive Land is considered "alive" when is vacant and fulfills desirability conditions for industry, and "dead" to the one that at some time was alive but was finally urbanized or started having one or more neighboring land uses that are not desirable to be adjacent and thus stopped being alive. The interest in this indicator is to identify zones where there is a high urbanization of Attractive Land that can mean a high competition for land and consequently a need to regulate with more precision.

Fig. 15. Average behavior of Attractive Land states in a 10-year span.

Results show that there is always a higher amount of Attractive Land that is alive in the initial years but then invariably goes down as dead surface rises until they are inverted and there's is more dead than an alive surface. According to the model, the year in which these surfaces are inverted is density dependent, and it comes closer to the present as population densities are lowered. The rate at which alive surface diminishes is also density dependent. When density is low, it immediately goes down at the same rate that dead surface rises, but when density is high, the rates are significantly different as dead surface rises at a faster rate than the surface that is alive.

4   Conclusions

Through the use of the model presented in this paper, relevant topics that should be important for the UDPE update were identified for performance improvement of industrial land in the city. Some of these topics are only related to the industry, but others have a more generalized impact where industrial activity is just one of those affected.

First place, there is the issue of low population densities, which is a well-known theme in Ensenada's urban context. Through the model, it was identified that this is a factor that reduces industrial growth opportunities in a zone that is already limited by its abrupt topography. It has an important role in the gradual decrease of industry Attractive Land because as population densities drop it produces sprawl that rapidly engulfs Attractive Land. Even though population density has always been considered a

relevant issue for the city, it was always seen from the perspective of greater infrastructure demand or mobility, but the effects on a particular land use had never been identified, so this finding complements the understanding of low-density impacts in the city of Ensenada.

Just as low density is a limitation for the industry, so is the scarce road infrastructure. This has already been a well-known issue, but through the model, it can be observed how the formation of Attractive Land strongly depends on the completion of projects that complement the road structure of the city that remains incomplete to this day.

Another significant finding was the level of importance that non-compliance with UDPE regulations can have on industry growth. Although its insufficient application is well known, it was observed that the UDPE would have to be breached if growth was to be stimulated in areas where a high probability for Attractive Land for the industry was detected. This is further complicated because the model showed that there will always be more Attractive Land for the industry in subsectors where this activity is currently not allowed. Which forces to revise the UDPE under this perspective: or more detailed and strict regulations are elaborated that inhibits this dissemination, or some of these areas must be recognized as new zoning for industrial growth.

The need for greater normative specificity between land uses was also corroborated through the model since it demonstrated a high competition for vacant land occupation. It is particularly important since the model outputted areas with a high probability of forming Attractive Land near housing developments. It is these high probabilities and the steep topography that limits industry location that this model becomes a relevant justifier for redirecting the city's efforts in attracting generic industry. Instead, a focus should be made on drawing types of industry that have a higher possibility of carrying out their activities in the conditions mentioned above and without creating inhibitions with surrounding land uses.

References

(1)   IMIP. *Programa Sectorial de Desarrollo Industrial de Ensenada*. Instituto Municipal de Investigacion y Planeacion 2017, p 2017.

(2)   IMIP-Ensenada. *Programa de Desarro Urbano del Centro de Población de Ensenada 2030* **2009**, 117.

(3)   Economía, S. de. *D. Of. la Fed.* **2011**.

(4)   Mitchell, M. *Proc. Seventh Int. Conf. Lang. Resour. Eval.* **2009**, 367.

(5)   Alves, B.; Sakowski, P.; Tovolli, M. *Modeling Complex Systems For Public Policies*; Alves, B., Sakowski, P., Tovolli, M., Eds.; Institute for Applied Economic Research, 2015; Vol. 37.

(6)   Read, S. *Complex. Theor. Cities Have Come Age* **2012**, 105–127.

(7)   Filatova, T. *Comput. Environ. Urban Syst.* **2015**, *54*, 397–413.

(8)   Edmonds, B.; Meyer, R. *Simulating Social Complexity*; 2013.

(9)   Weiler, R.; Engelbrecht, J. *Cadmus J.* **2013**, *2* (1), 131–141.

(10)   Beheshti, R.; Sukthankar, G. *AI and Society*. 2014, pp 1–12.

(11)   Easley, D.; Kleinberg, J. *Science (80-.)*. **2010**, *81*, 744.

(12) Morcol, G. *Emerg. Complex. Organ.* **2012**, *14* (4), 1–16.

(13) Batty, M.; *Cities as Complex Systems: Scaling, interactions, Networks, Dynamics and Urban Morphologies*, *Work. Pap. Ser.* **2008**.

(14) Portugali, J.; Meyer, H.; Stolk, E.; Tan, E. *Complexity Theories of Cities Have Come of Age*; 2012.

(15) Atun, F. *Understanding effects of complexity in cities during disasters,* **2014**, *106*, 51–65.

(16) Allen, P. M. *Complexity Theories of Cities Have Come Age* **2012**, 67–89.

(17) Pumain, D. *Complexity Theories of Cities Have Come Age* **2012**, 91.

(18) Hillier, B. *Complexity Theories of Cities Have Come Age* **2012**, No. September, 129–152.

(19) Johnson, J. *Cities: systems of systems of systems Sustain. Dev.* **2012**, 153–172.

(20) Sandoval, J.; Castañón-Puga, M.; Gaxiola-Pacheco, C.; Suarez, E.; *Identifying Clusters of Complex Urban–Rural Issues as Part of Policy Making Process Using a Network Analysis Approach: A Case Study in Bahía de Los Ángeles, Mexico, Sustainability* **2017**, *9* (6), 1059.

(21) OECD; Global Science Forum. *Applications of Complexity Science for Public Policy: New Tools for Finding Unanticipated Consequences and Unrealized Opportunities*; 2009.

(22) Castelliani, B.; Hafferty, F. *Sociology and complexity science. A new field of inquiry*; 2009.

(23) Sandoval-Felix, J.; Castañón-Puga, M.; *Using an Urban Growth Simulator for Ensenada City Public Policy Analysis*, 1–12.

(24) Deep, S.; Saklani, A. *Egypt. J. Remote Sens. Sp. Sci.* **2014**, *17* (2), 179–187.

(25) Torrens, P.; Kevrekidis, I.; Ghanem, R.; Zou, Y. *Entropy* **2013**, *15* (7), 2606–2634.

(26) Lagarias, A. *Appl. Geogr.* **2012**, *34*, 146–160.

(27) Fang, S.; Gertner, G. Z.; Sun, Z.; Anderson, A. A. *Landsc. Urban Plan.* **2005**, *73* (4), 294–306.

(28) Sudhira, H. S.; Ramachandra, T. V; Jagadish, K. S. **2004**, *5*, 29–39.

(29) Axelrod, R. *J. Jpn. Int. Econ.* **2003**, *12* (3), 16–22.

(30) Conte, R.; Gilbert, N.; Bonelli, G.; Cioffi-Revilla, C.; Deffuant, G.; Kertesz, J.; Loreto, V.; Moat, S.; Nadal, J. P.; Sanchez, A.; Nowak, A.; Flache, A.; San Miguel, M.; Helbing, D. *Eur. Phys. J. Spec. Top.* **2012**, *214* (1), 325–346.

(31) Ball, P. *Why Society is a Complex Matter*; 2012.

(32) Gilbert, N.; Troitzcsch, K. G. *Simulation for the Social Scientist*; 2001.

(33) Furtado, B. A.; Sakowski, P. A. M.; Tóvolli, M. H. *Inst. Pesqui. Econômica Apl.* **2015**.

(34) Rueda, L.; Rico, D. *Sci. Tech.* **2007**, *XIII*, 421–426.

(35) Patel, A. V. Slumulation: An integrated simulation framework to explore spatio-temporal dynamic of slum formation in Ahmedabad, India, 2012.

(36) Goldstein, J. *Emergence* **1999**, *1* (1), 49–72.

(37) O'Sullivan, D.; Perry, G. L. W. *Spatial Simulation: Exploring Pattern and Process*; 2013.

(38) Alberti, M. *Advances in Urban Ecology*; 2008.

(39) McPhearson, T.; Pickett, S. T. A.; Grimm, N. B.; Niemela, J.; Alberti, M.; Elmqvist, T.; Weber, C.; Haase, D.; Breuste, J.; Qureshi, S. *Bioscience* **2016**, *66* (3), 198–212.

(40) Bretagnolle, A.; Daudé, E.; Pumain, D. From theory to modelling: urban systems as complex systems http://cybergeo.revues.org/2420?lang=es (accessed May 31, 2015).

(41) Netlogo website https://ccl.northwestern.edu/netlogo/ last accesed 2018/30/06

(42) Grimm, V.; Berger, U.; Bastiansen, F.; Eliassen, S.; Ginot, V.; Giske, J.; Goss-Custard, J.; Grand, T.; Heinz, S. K.; Huse, G.; Huth, A.; Jepsen, J. U.; Jørgensen, C.; Mooij, W. M.; Müller, B.; Pe'er, G.; Piou, C.; Railsback, S. F.; Robbins, A.; Robbins, M.; Rossmanith, E.; Rüger, N.; Strand, E.; Souissi, S.; Stillman, R.; Vabø, R.; Visser, U.; DeAngelis, D. *Ecol. Modell.* **2006**, *198* (1–2), 115–126.

(43) Spencer, D. *J. Archit.* **2009**, *14* (3), 446–450.

(44) IMIP. *Programa Parcial de Mejoramiento Urbano de la Zona Centro y Frente de Mar,* **2017**.

Simulating Urban Flows of Daily Routines of Commuters[*]

Tomas Crols[1,2][0000-0002-9379-7770] and Nick Malleson[1,2][0000-0002-6977-0615]

[1] School of Geography, University of Leeds, Leeds, LS2 9JT, UK
[2] Consumer Data Research Centre, Leeds Institute for Data Analytics, University of Leeds, Leeds, LS2 9JT, UK
t.w.crols@leeds.ac.uk

Abstract. Forecasts of the ambient population, the actual location of people throughout the day, could have different applications in fields that need to know how many people are at risk or in need of something at any time. These include crime science, transport research, or modelling exposure to air pollution. Recently, human movements can be tracked thanks to a diverse range of 'big' datasets having a high spatial and temporal resolution. Examples are mobile phone locations, footfall cameras, geolocated social media updates or transport smart cards. In this presentation, we discuss the development of an agent-based model of the ambient population at an individual level in an urban environment. Ultimately, the goal is to calibrate the model with different big data streams in real time by using dynamic data assimilation techniques. At the moment, we have developed a model of recurring activities of individual commuters (working, shopping, leisure, etc.) that we calibrated with a recent British time-use survey and validated with hourly Wi-Fi sensor footfall data in the town centre of Otley, West Yorkshire, UK. The validation dataset captures mobile phones passing by at different well-chosen crossings and buildings. The results show that modelling the behaviour of commuters only is not sufficient to capture the evolution of the ambient population.

Keywords: Agent-based modelling, Big data, Ambient population

1 Introduction

The 'big' data revolution has an impact on the scale, applicability and calibration options of numerous urban models [1]. Agent-based models are most suitable to quantify the location and behaviour of an urban population [2]. The aim of our ongoing work is to build an agent-based model of the ambient population, the number of people at any location in an urban environment at any time of the day. Several studies in other disciplines can benefit from such population forecasts with a detailed time scale, including crime research [3, 4] and modelling exposure to air pollution [5]. This presentation

focuses on the development of the agent-based model, an application to commuting in a town in Yorkshire, calibration of the model parameters with a time-use survey, and validation with a large dataset of footfall.

2   An agent-based model of the ambient population

In our model of the ambient population, agents are driven by intensities to do different daily routines, reflecting which goals they want to achieve first [6]. We break this behavioural framework of intensities up into time intensities and background intensities. Time intensities are about the time of day and day of week when it is more likely that agents do specific activities. Background intensities change depending on the recurrence pattern of every activity. Generally, they slightly increase when an agent does other activities and decrease faster when the activity is actually being performed.
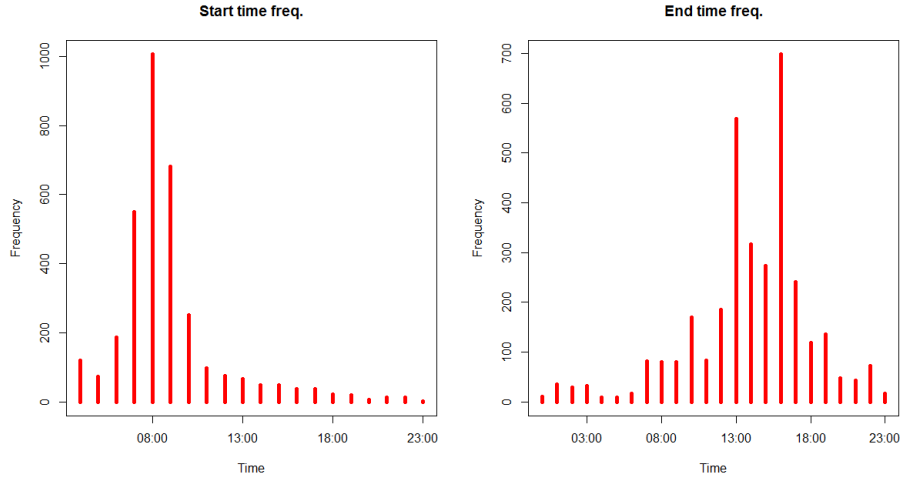
At present, we model the behaviour of individual commuters. Data on commuting between output areas, the lowest level of statistical units in the UK, were obtained from the 2011 UK Census [7]. In the model, the agents can be at home, work in their office, visit restaurants for lunch or dinner, go shopping, or do leisure activities (going out or sports). Commercial building functions were extracted from OpenStreetMap. The model code is developed in Scala with GIS functions and general agent classes based on the (Geo)MASON library [8].

We have applied this model to the town of Otley, West Yorkshire, UK, and its surrounding suburbs.

3   The UK Time Use Survey 2014-2015

The intensity parameters of the model are calibrated with the UK Time Use Survey 2014-2015 [9, 10]. The survey has 8278 full records of respondents who have kept diaries of all their activities per 10 minute interval during two days. We extracted time intensities and recurrence patterns for the activities of commuters on workdays. Not surprisingly, many commuters begin their workday at the office between 7 and 9 in the morning (see Fig 1), and stay on average for 8 hours. Part-time working seems to be much more popular in the morning than in the afternoon. Around 30 % of the commuters go shopping on an average workday (up to 40 % on Friday), and around 20 % do sports (only 15 % on Friday). Only a limited group (6.5 %) goes for lunch in a restaurant or cafe.

The data of the time use survey were used to manually calibrate the model by changing the parameters that determine the time intensities and the recurrence pattern of the activities.
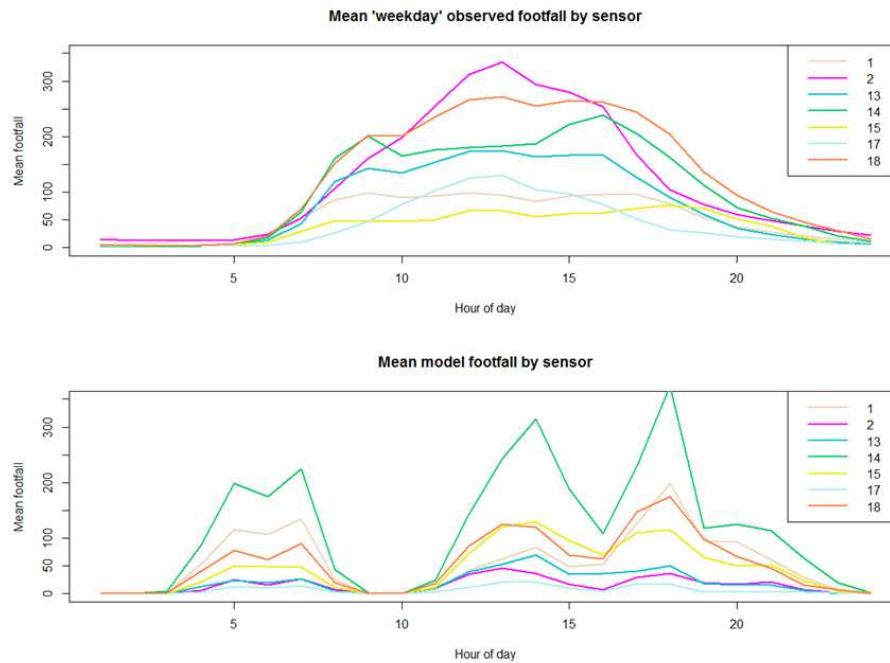
Fig. 1. Distribution of start and end times of working at the office in the UK, according to the UK Time Use Survey 2014-2015.

4   Validation with Wi-Fi footfall data

We obtained a dataset with hourly footfall for a period of two years (August 2015 –July 2017) at different key locations in Otley. The data has counted every Wi-Fi enabled mobile phone passing by. No personal data are collected and phones do not get an identifier to follow their path between different cameras, which means that there are no major concerns about the privacy of the people being tracked.

Even though activities of commuters are only a subset of the recorded mobility, we expected peaks in the model and the validation dataset to be similar. But except for one footfall device (with id 14, see Fig. 2), no clear morning or evening rush hour peaks could be detected. The model results obviously have these peaks since only activities of commuters were taken into account. This leads to the observation that a general model of the ambient population needs a much bigger effort to capture the behaviour of different groups with more unpredictable travel patterns, like unemployed or retired people, and employees with highly variable schedules.

In the future, we intend to further generalise the model by working with households and including a more diverse spectrum of daily activities. The calibration will be automised and would ideally lead to the development of live footfall predictions. Such live predictions can be based on data assimilation techniques that were developed for weather models. Initial tests with a simple footfall model are promising [11].

4

Fig. 2. Comparison between observed and modelled footfall at seven locations in Otley on an average workday (Monday and Friday were excluded).

References

1. Crooks, A., Malleson, N., Wise, S., Heppenstall, A.: Big data, agents and the city. In: Schintler, L.A., Chen, Z. (eds.) Big data for regional science. Routledge, Abingdon / New York (2018).
2. Burger, A., Oz, T., Crooks, A., Kennedy, W.G.: Generation of Realistic Mega-City Populations and Social Networks for Agent-Based Modeling. The 2017 Conference of the Computational Social Science Society of the Americas, Santa Fe (2017).
3. Malleson, N., Andresen, M.A.: Exploring the impact of ambient population measures on London crime hotspots. Journal of Criminal Justice, 46, 52–63 (2016).
4. Kounadi, O., Ristea, A., Leitner, M., Langford, C.: Population at risk: using areal interpolation and Twitter messages to create population models for burglaries and robberies. Cartography and Geographic Information Science, 45(3), 205-220 (2018).
5. Park, Y.M., Kwan, M.-P.: Individual Exposure Estimates May Be Erroneous When Spatiotemporal Variability of Air Pollution and Human Mobility Are Ignored. Health & Place, 43, 85–94 (2017).
6. Bratman, M.E., Israel, D.J., Pollack, M.E.: Plans and resource-bounded practical reasoning. Computational Intelligence, 4(3), 349-355 (1988).
7. Office for National Statistics: Census: Special Workplace Statistics (United Kingdom) (data collection). UK Data Service Census Support, https://wicid.ukdataservice.ac.uk (2011).

8. Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K., Balan, G.: MASON: A Multiagent Simulation Environment. Simulation, 81(7), 517-527 (2005).

9. Gershuny, J., Sullivan, O.: United Kingdom Time Use Survey, 2014-2015. UK Data Service, Study Number 8128 (data collection). Centre for Time Use Research, University of Oxford, Oxford (2017).

10. Morris, S., Humphrey, A., Cabrera Alvarez, P., D'Lima, O.: The UK Time Use Survey 2014 – 2015, Technical Report. Centre for Time Use Research, University of Oxford, Oxford, http://doc.ukdataservice.ac.uk/doc/8128/mrdoc/pdf/8128_natcen_reports.pdf, last accessed 2018/03/28 (2016).

11. Ward, J.A., Evans, A.J., Malleson, N.S.: Dynamic calibration of agent-based models using data assimilation. Royal Society Open Science, 3(4), 150703 (2016).

Estimating Shared Autonomous Vehicle Fleet Size to Meet Urban Daily Travel Demand

Wenwen Zhang[1], Henning S. Mortveit[2,3], and Samarth Swarup[3]

[1] School of Public and International Affairs,
[2] Department of Mathematics,
[3] Network Dynamics and Simulation Science Lab, Biocomplexity Institute,
Virginia Tech,
Blacksburg, VA 24061, USA.
{wenwenz3,hmortvei,swarup}@vt.edu

Abstract. Shared autonomous vehicles (SAVs) present the possibility of greatly reducing the number of cars in use, and consequently the required parking space. We present a methodology to estimate the required SAV fleet size to meet travel demand for a region, and develop a detailed synthetic population model where we model every individual in a city, along with typical weekday activity patterns to estimate the travel demand. We combine this with a simulation of SAV routing to determine the fleet size needed to satisfy all trips with small waiting times. Our results show significant reductions in both the number of vehicles on roads and parking demand in cities, which would result in substantial savings.

Keywords: Shared autonomous vehicles · synthetic populations · multiagent simulation · transportation · traffic

1 Introduction

Urban spaces and lifestyles are expected to be radically transformed in the near future with the availability and increasing use of Autonomous Vehicles (AVs). This will reduce congestion and increase safety, and potentially change travel behavior by reducing the need for vehicle ownership and promoting ride sharing. The benefits of AVs are expected to be more significant once the technology is combined with the shared economy. Shared Autonomous Vehicles (SAVs) are an envisioned new door-to-door mobility service, which is expected to be more affordable [5], reliable [7] and environmentally friendly [8] than conventional vehicles and privately-owned AVs.

Therefore, in the present work, we focus specifically on the possibility of using a fleet of shared autonomous vehicles to serve the entire daily travel demand of a region. We use the city of Richmond, VA, USA as a case study, though our methodology is broadly applicable. We develop a system that combines a detailed discrete event simulation methodology for modeling the use of shared autonomous vehicles with a data-driven synthetic population model of Richmond where every person is represented along with typical weekday activity patterns.

The simulation allows us to come up with a realistic estimate of the size of the fleet required as well as the parking demand for the SAV fleet.

2   Shared autonomous vehicle management

We used an agent-based SAV model, which we call the SAV management model, to estimate SAV fleet size and curbside parking footprint for the city of Richmond. The model inputs include (1) synthetic trips, with origin, destination and departure time, and (2) Origin-Destination (OD) travel time matrix. The synthetic trips and travel time matrix are obtained using the synthetic population model and routing algorithm discussed in the next section. The design of the SAV model is elaborated below.

There are three types of entities in the simulation model: (1) client entity, (2) vehicle entity, and (3) queue entity. Clients generate call events, once they decide to move from one place to another. When handling the call events, the SAV system will assign empty vehicle entities within a 2-mile travel distance to fulfill the trip, which leads to pick up and drop off events. After dropping off the last client entity, if the vehicle entity is not assigned for other services, the system will initiate a relocating event to determine whether the vehicle is in an SAV surplus area and needs to relocate to under-served areas. The vehicle will eventually park if it is in an under-served area and is not assigned to service. The simulation starts with no vehicles in the system and keeps adding vehicles into the system once the client has waited for more than 10 minutes. This allows us to estimate the number of vehicles needed to satisfy the travel demand with a maximum waiting time of 10 minutes (or any other chosen threshold).

3   Travel Demand and Routing

We use a detailed, data-driven modeling approach to generate the estimated travel demand for the city of Richmond, VA, known as a synthetic population model [1]. The steps in this process are:

- **Baseline population synthesis:** We begin by creating a set of disaggregated agents with multiple demographic variables attached to them, using the methodology of Beckman et al. [4]. The demographic data are drawn from the 2015 American Community Survey (ACS) for the state of Virginia.
- **Activity assignment:** In this step we use data from the National Household Travel Survey (NHTS)[4] to assign a daily activity sequence to each agent in the synthetic population constructed in the previous step [9].
- **Location choice:** In this step we assign a location for each activity for each agent. Road network data are used to create home locations. Locations for other activities are then assigned using business and school location data, using a gravity model.
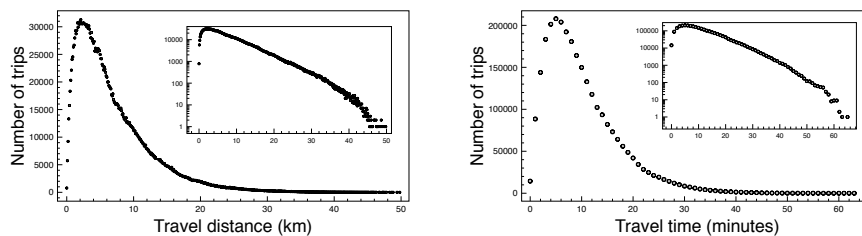
---

[4] http://nhts.ornl.gov/

Fig. 1: Travel distance distribution (left) and travel time distribution (right) over all trips, as generated by the routing model. The inset shows the y-axis on a log scale in each case.

Routing

For the routing work described in this paper, we used a regular-expression constrained variant of the Dijkstra's shortest path algorithm developed in [3] with optimization and parallelization implemented in [2].

Network description. The transportation network covered the Richmond Urban Area and was extracted from the HERE StreetMap Premium data set for the U.S., version 2017 (Q2).

Routing. The study included over 2.8 million trip requests, extracted from the synthetic population. The computations were split onto 360 cores on a computational cluster, finishing in roughly 22 compute days.

Figure 1 shows the distribution of travel distances and the distribution of travel times. For the purpose of visualization, travel distances are rounded to the nearest tenth of a kilometer and travel times are rounded to the nearest minute. Both distributions have exponential tails, as shown by the log-linear insets. Our travel demand does not include long-distance travel, since this is not assumed to be serviced by the city's SAV fleet.

4   Results

We applied the models to the City of Richmond, Virginia, USA. The SAV model was implemented with SAV market penetration rates of 1%, 10%, and 100% to explore the impact of market size on the fleet requirement and correspondingly the parking footprint. The results are tabulated in Table 1. The results suggest that the larger the market size, the more efficient the SAV system, as the number of trips served per SAV increases from 19 trips per SAV per day in the 1% market penetration scenario to 22 trips per SAV per day in the 100% scenario. Additionally, the results show that the parking demand is linearly correlated with the number of SAVs in the system. Despite the fleet size of the system, the parking demand is, on average, approximately 5-6 spaces per SAV running in the system.

The vehicle replacement rate of the SAV system is approximately 5.4. The Richmond synthetic population includes a total of 370,130 households and 671,165

Table 1: Model Results by Market Penetration Rate

Market Penetration	1%	10%	100%
Vehicle Fleet Size	1514	13759	126710
Served Trips per SAV	19	21	22
Parking Demand	9311	71546	620879
Parking Demand per SAV	6	5	5

adults, who generate 2.9 million trip requests on a daily basis. In the U.S., the average vehicle ownership per adult is approximately 0.99 [10]. The metropolitan area of Richmond approximately owns 664,453 private vehicles. Our results show that 169,882 vehicles are sufficient to serve the region to ensure that all clients can be picked up within the 10-minute waiting window. In other words, the replacement ratio is approximately 5.4 (i.e., 664453/126710).

The parking demand in the study area also declines dramatically after the introduction of SAVs. In the U.S., approximately four parking spaces are needed to accommodate the parking demand of one privately owned vehicle [6]. Therefore, at the 100% market penetration level, the SAV system holds the potential to reduce over 85% of the existing parking spaces. The reduction in parking space would translate into significant savings in city expenditures.

References

1. Adiga, A., et al.: Generating a synthetic population of the United States. Tech. Rep. NDSSL 15-009, Network Dynamics and Simulation Science Laboratory (2015)
2. Adiga, A., Marathe, M.V., Mortveit, H.S., Swarup, S., Wu, S.: Modeling urban transportation in the aftermath of a nuclear disaster: The role of human behavioral responses. In: Proceedings of the Conference on Agent-Based Modeling in Transportation Planning and Operations. Blacksburg, VA (Sep 30 - Oct 2 2013)
3. Barrett, C., Jacob, R., Marathe, M.: Formal-language-constrained path problems. SIAM J. Comput. **30**(3), 809–837 (2000)
4. Beckman, R.J., Baggerly, K.A., McKay, M.D.: Creating synthetic base-line populations. Transportation Research A – Policy and Practice **30**, 415–429 (1996)
5. Burns, L.D., Jordan, W.C., Scarborough, B.A.: Transforming personal mobility. The Earth Institute **431**, 432 (2013)
6. Chester, M., Horvath, A., Madanat, S.: Parking infrastructure: energy, emissions, and automobile life-cycle environmentalaccounting. Environmental Research Letters **5**(3), 034001 (2010)
7. Fagnant, D.J., Kockelman, K.M., Bansal, P.: Operations of shared autonomous vehicle fleet for Austin, Texas, market. Transportation Research Record: Journal of the Transportation Research Board (2536), 98–106 (2015)
8. Greenblatt, J.B., Saxena, S.: Autonomous taxis could greatly reduce greenhouse-gas emissions of US light-duty vehicles. Nature Clim Chng **5**(9), 860–863 (2015)
9. Lum, K., Chungbaek, Y., Eubank, S.G., Marathe, M.V.: A two-stage, fitted values approach to activity matching. Int J Transportation **4**(1), 41–56 (2016)
10. Santos, A., McGuckin, N., Nakamoto, H.Y., Gray, D., Liss, S.: Summary of travel trends: 2009 National Household Travel Survey. Tech. rep. (2011)

Evidence for the 'safety in density' effect for cyclists; validation of agent-based modelling results

Dr Jason Thompson[1,2], Dr Jasper S. Wijnands[1], Dr Suzanne Mavoa[3], Katherine Scully[1] and Professor Mark Stevenson[1]

[1]University of Melbourne, Transport, Health and Urban Design Research Hub, University of Melbourne, Australia, Ph +61 3 83447782
[2]University of the Sunshine Coast, Human Factors and Socio-Technical Systems Research Unit
[3]University of Melbourne, School of Population and Global Health

Abstract. Time-gap analysis of cyclists passing through an intersection was conducted using five hours of video-observation of a single intersection in Melbourne, Australia, where motorists were required to 'yield' to oncoming cyclists. Results demonstrated that potential collisions between motor-vehicles and cyclists reduced with increasing cyclists per minute. These results successfully validate previous 'synthetic' evidence gathered using agent-based models, supporting evidence of a proposed causal mechanism related to safety in density (SiD) rather than safety in numbers, per se. Results suggest that increased cyclist safety may be achieved through creating high-density strategic cycling corridors. This study highlights the utility of using agent-based models to identify data that are needed to be collected to advance understanding of public health and urban policy issues.

Keywords: Cycling, safety, traffic, health

1 Background, Method, Results and Conclusions

The safety in numbers (SiN) effect for cyclists is a widely referenced and observed, but poorly understood, phenomenon [1-3]. Although many academic and applied studies cite SiN as a potential solution to car vs cyclist crashes [3-7], there is little definitive evidence to guide policy-makers or transport planners in how to use SiN to create a safer cycling environment beyond simply encouraging 'more cyclists' into the system.
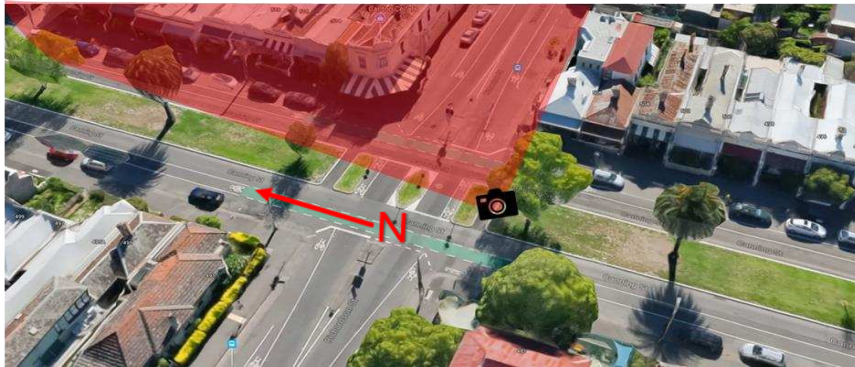
Recent results of agent-based models (ABMs)[8-10] have shown the SiN effect can be replicated in simulated systems that lead to the formation of higher-density cyclist groups. This suggests the safety effect is perhaps driven by a simple spatial arrangement of cyclists, rather than a result of learned behaviour by drivers. While acknowledged as a potential mechanism, this 'safety in density' hypothesis has been criticised[11] from the perspective that no in-situ empirical evidence exists that cyclists 'cluster' in the real

world. The focus of this study was therefore to observe micro-level interactions of cyclists and vehicles at an intersection mirroring that created in prior ABMs to determine how cyclist density is associated with potential crash risk.

1.1 Method

Data collection occurred through recording 5 hours of video of naturalistic traffic behaviour at an inner-city Melbourne cross-intersection (see Figure 1). Coding of time-stamped video footage of car and cyclist interactions enabled the calculation of variables, including:

- the time gap of approaching cyclists to the intersection at the instant when drivers arrived at the intersection (sec);
- number of cyclists each driver gave way to before moving away from the intersection (n);
- time gap between each cyclist passing through the intersection (sec); and
- frequency of cyclists passing through the intersection (cyclists / min).

Fig. 1. View of the selected intersection showing orientation, location and coverage of the video camera.

1.2 Results

The frequency of potential collisions (y) (accepted time gaps of < 4.2 seconds) was modelled using the number of cyclists per minute (x) as a single explanatory variable. Potential collisions and the number of cyclists were counted each minute across the total five hours of captured video footage, resulting in an appropriate sample of cyclist density and potential collision counts (n = 385). The potential collision count was modelled using Poisson regression.

Parameter estimates are presented in Table 1 indicating that the number of cyclists per minute was strongly associated with potential collision risk for cyclists. Figure 2 shows this relationship in more detail, demonstrating that as the number of cyclists per

minute moving through the intersection increased, the risk of collision per cyclist decreased dramatically up to the point of 8-10 cyclists per minute. This effect of cyclist spatial density on collision risk was independent of behavioural adaptation by drivers.
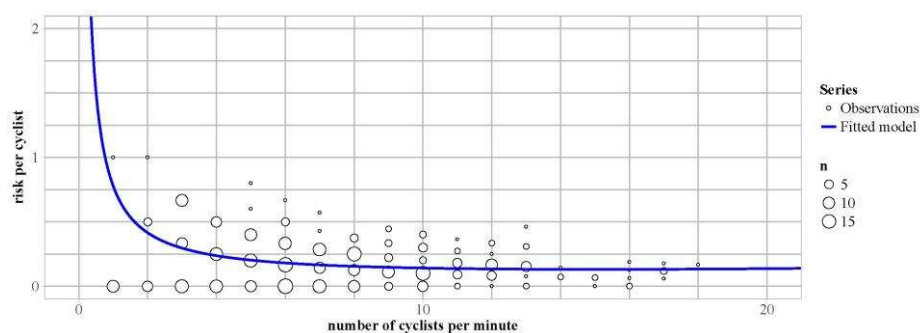
Table 1. Parameter estimates of Poisson regression.

	Estimate	Std. error	z value	P(>|z|)
α	-0.32	0.10	-3.31	.000
β	0.07	0.01	6.65	.000

The resulting estimated risk per cyclist (r) was:

Equation (1)

$$r = \frac{e^{-0.32+0.07x}}{x}$$

Fig 2. Potential crash risk with increasing count of cyclists per minute.

1.3   Conclusions

Public health researchers have identified the potential of agent-based models to highlight data needed to understand aspects of public health issues. However, few examples exist where empirical data collection has occurred as a result of synthetic model development, which can then be used to inform future modelling exercises. This study provided empirical validation of the previously modelled 'Safety in Density' hypothesis' operation in a real-world situation. It demonstrated that reduced crash risk was associated with reduced time-gaps between cyclists passing through intersections, which prevent motor-vehicles from attempting to move between on-coming cyclists (gap rejection). Using a methodology based on prior theoretical experimentation using ABMs, this work has provided further support for a candidate causal mechanism underlying the widely observed general relationship between cycling volumes and safety; one that

4

has thus far eluded comprehensive explanation in the cycling safety literature. The results of this study will inform the next generation of ABMs (see Figure 3, below), focused on understanding mechanisms associated with cyclist safety at intersections, completing a loop in model development, implementation and iteration that has been called for within public health [12].

Fig 3. Prototype model of cycling safety developed from empirical evidence gathered from observed car / cyclist interactions.

References

1. Bhatia, R., Wier, M.: "Safety in Numbers" re-examined: can we make valid or practical inferences from available evidence? Accid Anal Prev 23(1), 235-40 (2011).
2. Christie, N., Pike, I.: Is it safe in numbers? Inj Prev 21(4), 276-7 (2015).
3. Elvik, R., Bjørnskau, T.: Safety in numbers: A systematic review and meta-analysis of evidence. Safety Science 92, 274-282 (2017).
4. Jacobsen, P.L.: Safety in numbers: more walking and bicyclists, safer walking and bicycling. Inj Prev 9(3), 205-9 (2003).
5. Robinson, D.L.: Safety in numbers in Australia: more walkers and bicyclists, safer walking and bicycling. Health Promotion J Austr 16(1), 47-51 (2005).
6. Fyhri, A., et al: Safety in numbers for cyclists – conclusions from a multidisciplinary study of seasonal change in interplay and conflicts. Accid Anal Prev 105, 124-33 (2017).
7. Tin, S.T., et al: Regional variations in pedal cyclist injuries in New Zealand: safety in numbers or risk in scarcity? Aust N Z J Public Health 35(4), 357-63 (2011).
8. Thompson, J., Savino, G., Stevenson, M.: A model of behavioural adaptation as a contributor to the safety-in-numbers effect for cyclists. Transportation Research Part A: Policy and Practice 85, 65-75 (2016).
9. Thompson, J., Savino, G., Stevenson, M.: Reconsidering the safety in numbers effect for vulnerable road users: an application of agent-based modelling. Traffic Inj Prev 16(2), 147-53 (2015).
10. Thompson, J., et al: Estimating the safety benefit of separated cycling infrastructure adjusted for behavioural adaptation among drivers; an application of agent-based modelling. Transport Research Part F: Traffic Psychology and Behaviour 49, 18-28 (2017).
11. Jacobsen, P.L., Ragland, D.R., Komanoff, C.: Safety in Numbers for walkers and bicyclists: exploring the mechanisms. Inj Prev 21(4), 217-20 (2015).
12. Tracy, M., Cerda, M., Keyes, K.M.: Agent-based modelling in public health: current applications and future directions. Annu Rev Public Health 39, 77-94 (2018).

Initial Results from an Agent-Based Simulation of Housing in Urban Beirut

Stefano Picascia[1,2], Ali Termos[3], and Neil Yorke-Smith[4,3]⋆

[1] Manchester Metropolitan University, UK
[2] University of Siena, Italy
stefano@cfpm.org
[3] American University of Beirut, Lebanon
aatermos@gmail.com
[4] Delft University of Technology, The Netherlands
n.yorke-smith@tudelft.nl

Abstract. The motivation for our work is to develop an agent-based model (ABM) to capture the existence of migrant and refugee flows, and to explore their effects on urban dynamics. We leverage an extant agent-based model founded on the rent-gap theory, as a lens to study the effect of sizeable refugee migration upon a capital city in West Asia. In order to calibrate and validate the simulation model, we construct indices for housing prices and other factors. Results from the model show the impact of migration shock on the housing market, and identify the relative efficacy of housing intervention policies. Our work progresses towards a tool for policy makers asking what-if questions about the urban environment in the context of migration.

Keywords: rent-gap theory · migration · agent-based modelling · Lebanon

1 Introduction

Previous works constructed a micro-level ABM, derived from the *rent-gap theory* (RGT)—an economic hypotheses on the dynamics of investment in housing [9,2]—in an effort to model the economic mechanics of property investment along with their effects on the cultural and social diversity of urban areas. Picascia [7], whose approach we follow, examined the price dynamics triggered by different levels of capital flowing in a city and the patterns of spatial inequality that may emerge. He developed a model of housing regeneration, and applied it to a major European city [6]. We went on to extend the ABM to include refugee migration, but only in the form of an abstract, stylized model [8].

Our current work goes beyond previous models [1,2,3,10,8], grounding our work in the city of Beirut, Lebanon. This is a densely-populated capital of a middle-income West Asian country which, according to official estimates [12,4], has experienced a refugee influx of approaching 50% of its population in the period 2012–2016, causing extreme pressure on housing and social fabric [4,11].

⋆ Corresponding author.

We build the first (to our knowledge) validated ABM of metropolitan Beirut using data for population, housing prices, and property conditions. Our contribution includes devising principled indices for these factors and incorporating UNHCR-supplied data for refugees. Our model, implemented in the ABM platform NetLogo [13], is adept to investigate the possible societal consequences and economic indicators of migration, being specifically designed to explore the interrelation of urban economic and cultural dynamics. Results of the simulation, briefly discussed here, exhibit how sizeable migration of low income populations into a city impacts prices, slum locations, population density, and segregation.

2   Simulation Model

Our agent-based model simulates urban dynamics at the level of an entire city. The entities represented in the model are: (1) individual locations (residential properties), defined by their market value, repair state, and population; (2) individual agents that represent households (whether indigenous households or refugees households), characterised by an income, mobility propensity, and cultural configuration; and (3) economic forces, represented principally in the form of exogenous 'capital' level, aiming at profiting from redevelopment/restoration of residential locations. The model in stylized form was described in our previous paper at ABMUS'16 [8].

As we went on to discuss at ABMUS'17 [5], there are multiple challenges in modelling Beirut using actual data. Lebanon has no official census since 1932, and in addition many of the current refugees are not legally resident. Reliable geo-referenced data relating to the variables relevant in the model are non-existent or not easily accessible. Prior to 2017, there is no official, published data about housing prices—neither for Beirut nor indeed for the country—leaving only estimates.

While open-source GIS data proves adequate for cartographic modelling, and while there are a basket of population estimates that can be weighed for demographic modelling, to obtain economic data we developed and validated a multi-variate time series econometric analysis. Further, to obtain data about the current maintenance condition of residential housing in Beirut, we undertook a property survey in summer 2017.

3   Summary Results and Analysis

Figure 1 shows example evolution of prices and population over time, in three scenarios. The simulated period corresponds to five years starting in 2012. These snapshots in the figure are example runs only. The top row shows a baseline scenario without refugees, the middle row shows limited but sizeable refugee influx, and the bottom row shows an ongoing influx. We ran the simulation 10 times for each parameter setting and performed statistical analysis.

Outcomes from the model exhibit that refugees have some surprising impacts on prices and slum locations, as well as impacting population density. Their presence can cause price spikes and invigorate or dampen the natural economic cycle,
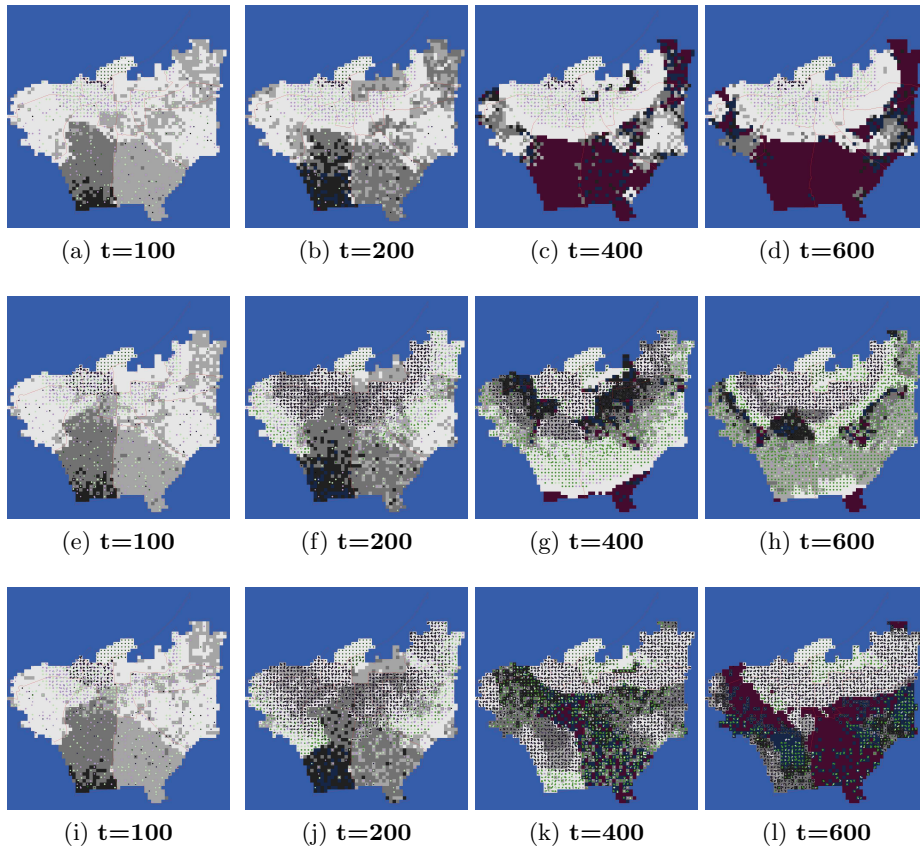
Fig. 1: City locations over time. Top row: no refugees; Middle row: refugee influx between $t = 90$ and $t = 180$ months; Bottom row: ongoing refugee influx from $t = 90$ months. Lighter colour indicates higher property value. Slum locations are dark blue, uninhabitable locations are dark purple.

which continues nonetheless with waves of decay and regeneration. We observe a further interesting effect depending on the capital in the economy and the rate of mobility, whereupon excessive shock of refugees—and in particular ongoing migration to the city—can render large sections of quarters uninhabitable, causing parts of both refugee and non-refugee populations to emigrate from the city.

References

1. Benenson, I.: Multi-agent simulations of residential dynamics in the city. Computers, Environment and Urban Systems **22**(1), 25–42 (1998)
2. Diappi, L., Bolchi, P.: Gentrification waves in the inner-city of Milan: A multi agent/cellular automata model based on Smith's Rent Gap theory. In: Innovations in Design & Decision Support Systems in Architecture and Urban Planning. pp. 187–201. Springer, Berlin, Germany (2006)
3. Diappi, L., Bolchi, P.: Smith's rent gap theory and local real estate dynamics: A multi-agent model. Computers, Environment and Urban Systems **32**(1), 6–18 (2008)
4. Government of Lebanon, the United Nations: Lebanon crisis response plan 2017–2020. https://data2.unhcr.org/en/documents/details/53061 (2017), accessed 27sep17
5. Picascia, S., Termos, A., Yorke-Smith, N.: Towards data-driven agent-based simulation: A case study of West Asian urban dynamics. In: Working Notes of the AAMAS'17 Workshop on Agent Based Modelling of Urban Systems (2017)
6. Picascia, S.: A theory driven, spatially explicit agent-based simulation to model the economic and social implications of urban regeneration. In: Proc. of the 11th Conf. of the European Social Simulation Association (ESSA'15) (2015)
7. Picascia, S., Edmonds, B., Heppenstall, A.: Agent based exploration of urban economic and cultural dynamics under the rent-gap hypotheses. In: Multi-Agent-Based Simulation XV. vol. LNCS 9002, pp. 213–227. Springer, New York, NY (2015)
8. Picascia, S., Yorke-Smith, N.: Towards an agent-based simulation of housing in urban Beirut. In: Agent Based Modelling of Urban Systems. vol. LNCS 10051, pp. 3–20. Springer, New York, NY (2016)
9. Smith, N.: Toward a theory of gentrification a back to the city movement by capital, not people. J. American Planning Association **45**(4), 538–548 (1979)
10. Torrens, P., Nara, A.: Modeling gentrification dynamics: A hybrid approach. Computers, Environment and Urban Systems **31**(3), 337–361 (2007)
11. UN HABITAT: Housing, land and property issues in Lebanon: Implications of the Syrian refugee crisis. http://www.alnap.org/resource/20198 (2014), accessed 27sep17
12. United Nations High Commissioner for Refugees: UNHCR – Lebanon. http://www.unhcr.org/pages/49e486676.html (2017), accessed 27sep17
13. Wilensky, U.: NetLogo. http://ccl.northwestern.edu/netlogo (1999)

Agent Based Modeling of the Australian Housing Market

Gideon DPA Aschwanden[1]; Jason Thompson[1]; Friedrich Burkhard von der Osten[2];
Viet Nguyen[1]; Zita Ulman; Brendan Coates[3]

[1] The University of Melbourne, Melbourne VIC 3000, Australia

gideon.aschwanden@gmail.com

Abstract. Factors and policy settings that drive affordability in the Australian housing market are hotly contested. This paper explores the use of agent-based modelling to quantify implications of policy changes on prices in the housing market with regard to ownership distribution and property price changes. The model containing heterogenous households, a development sector and financial sector is presented. The model enables experimentation exposed to system shocks (income, population, credit, policy and outlook) that cause fluctuations and long-term ripples in the market. Initial results show that policy adjustments, alone, had minor impact on both house prices and the distribution of ownership. Greater impact on the long-term property prices related to market sentiment. Further, changes to wage growth were amplified by a factor of 8 in household property prices.

Keywords: Agent Based Model, Housing Market, Policy Modeling.

1   Introduction

The Australian housing market has experienced significant price growth in the past 2 decades, creating problems for affordability for many people attempting to enter the market. Between 1992 and 2015 the ratio of investment loans in housing has gone from 15% to 56% making owner occupied housing loans the minority ((ATO), 2016). In public policy discussions regarding the reasons for such rises, two tax regimes in Australia are often highlighted: Negative Gearing (NG) and Capital Gains Tax Discounts (CGTD).

Using an agent-based model (ABM) containing a set of heterogenous households that make individual investment decisions, this project set out to build a model of the Australian housing market that could be used to investigate the impact of NG and CGTD, as well as other environmental and financial settings, on house prices.
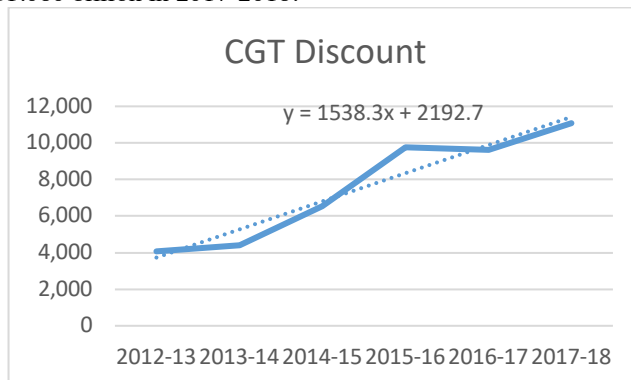
2

1.1 Policies

Even though the model can be used to investigate a range of environmental parameters and changes to it the focus of this project are the two policies that are targeted at housing: Negative Gearing and Capital Gains Tax Discount. These tax concessions are aimed to promote investment in housing markets.

Negative Gearing

The Australian tax code allows to deduct losses from investments into housing from gross income. This is not unique but exists in New Zealand, Canada and Japan and as a fraction in the U.S., Germany, France, Switzerland and Sweden. The costs on society accumulated in 2016 $2 billion or 5% of the budget deficit (GRATTAN INSTITUTE, "Hot property: Negative gearing and capital gains tax reform,"(Daley and Wood, 2016)). Arguments are made that this tax stimulates the housing supply and with it the number of rental properties. The counter argument focuses on the increased housing prices that remove access to real estate for many low and medium income households (Berkovec and Fullerton, 1992).

Capital Gains Tax Discount

Capital Gains Tax (CGT) in Australia are discounted 50% if the assets are held for longer than 12 months. Even though this is a tax code not aimed at the housing market its implications are. In the fiscal year 2001-02 Australia had a net capital gain from CGT $3.657 Billion while the CGT Discount was estimated to be $1.81 Billion (https://www.business.unsw.edu.au/About-Site/Schools-Site/Taxation-Business-Law-Site/jattavolumes/3_Kenny_JATTA_vol1_no2.pdf) the forgone tax expenditure rose to more than doubled in 10 years to $4.070 billion in 2012-13 and nearly tripled again to $11.080 billion in 2017-2018.

CGT Discount

$y = 1538.3x + 2192.7$

| | 2012-13 | 2013-14 | 2014-15 | 2015-16 | 2016-17 | 2017-18 |

Fig. 1. Forgone Capital Gains Tax due to discount in Australia in Billions per year
(https://static.treasury.gov.au/uploads/sites/1/2017/06/Tax-Expenditure-Statement-2016.pdf)

With both tax schemas having significance (Laidler, 1969), this project aims to model quantitatively the implications of both negative gearing and capital gains tax on

the real estate market. Since both are taxes schemas targeting the individuals and depend on the income of individuals an agent-based modeling framework is chosen that can incorporate this.

1.2 Modeling Framework

The ABM was built around a set of individual households making decisions based on a bounded view of a dynamic housing and financial environment. The environment consists of a developer sector agent that produces new properties to fulfill demand and responds to free liquidity, a financial sector that provides liquidity to individual households, a population model that adds new households according to population growth estimates, an alternative investment market that consists of a stock exchange and bond market.

1.2.1 Households

Households constantly evaluated their financial situation and reacted in a manner designed to optimize their long-term position (Rosen, 1985). Every household has a different investment horizon that adjust the perspective and calculation methods that drive decision. The investment horizon ranges between 0 - 40 years for renting and 0 – 30 years for investing households. Investment options ranged from keeping assets in the bank, or investing in government bonds, property or the stock market. For each alternative each household is calculating the net present value depending on their investment size, time horizon and a limited view of the transaction history.

$$NPV_{Buy} = Benefits_{buy} - oneoffcosts_{buy} - ongoingCost_{buy}$$

$$Benefits_{buy} = P_{value,H} * Building_{depreciation,H} - Mortgage_H - oneofcosts_{sale,H}$$

$$ongoingCost_{buy} = \sum_0^H \begin{pmatrix} mortgageCost_H \\ +maintenance_H \\ +councilTaxes_H \\ +insurance_H \end{pmatrix}$$

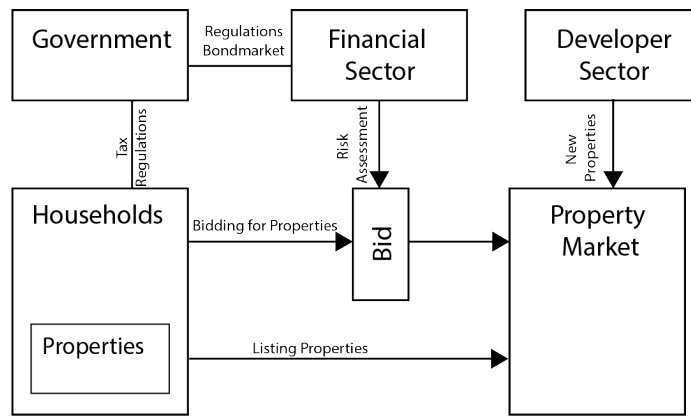$$P_{value,H} = P_{value,0} * \left(\frac{1 + appreciation}{Inflation} \right)^H$$

With $oneoffcosts_{buy}$ being the deposit and the transaction costs involved in the purchase, $ongoingCost_{buy}$ is the sum of all costs, adjusted for inflation, involved in

4

holding the property for every year until the investment horizon is reached, while the price of the property at the end of the investment horizon H is calculated as the current value multiplied by the expected appreciation and adjusted for inflation. The method to calculation expected appreciations are explained bellow with the bids.
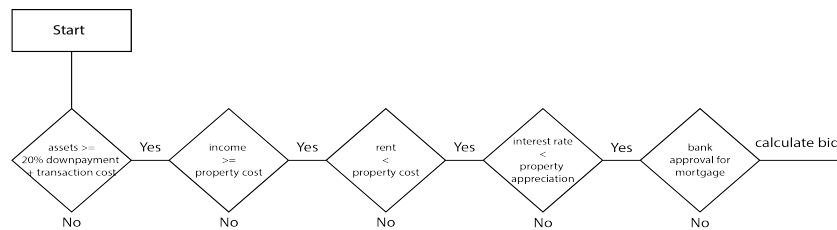
Home ownership required the capacity of households to make a down-payment of 20% of the property value, an ability to service mortgage and associated transaction costs (conveyancing, legal fees, etc.).

Fig. 2 shows each household's decision-making process. The developer, financial sector agents, and government provide the environmental backdrop of the agent's decision process.

Fig. 2. Simplified Agent Relation Diagram of the Model

Time was discretized by weeks and the model populated by 10000 households that represented a cross-sectional estimated distribution the current Australian income distribution. At each time step, households passed through a decision process where their total net present value (NPV) (i.e., financial position) was calculated.

Fig. 3. Household decision parameters for buying a property

Households had two methods to calculate their maximum bid for a particular property based on appreciation; taking into account geographic proximity $bid_{location}$ or the market development since the last transaction bid_{time}.

$$bid_{location} \; = \; P_{t-1} \left(\frac{1}{n} \Sigma_{-\frac{n}{2}}^{+\frac{n}{2}} a_i + \beta \right)^t$$

With P_{t-1}denoting the last known transaction price, n the number of properties in the geographic vicinity with their last annualized appreciation a_i and β was the outlook of the household calculating the bid.

$$bid_{time} \; = \; P_{t-1} \left(\frac{1}{n} \sum_{-i}^{o} a_i + \beta \right)^t$$

With a_i denoting the average annualized appreciation of all sold properties at time i.

1.2.2 Financial Sector

The financial sector assessed each household's capacity to purchase properties. If a household attempted to register a bid on a property, the financial sector checked that the purchase price matched the household's ability to make the 20% down-payment, and was also able to service the mortgage.

1.2.3 Developer Sector

The developer sector agent added properties at a predefined rate of 2% per annum.

1.2.4 Property Market

The property market collated all properties on offer as well as bids proposed by households. The market attempted to clear all properties on offer, starting from the property with the lowest reservation price. The implemented auction process is using a '*Vickrey Auction*' also known as a *sealed-bid second-price auction* (Vickrey, 1961). This leads to properties being sold to the highest bidder for the price of the 2ᵈ highest bidder's limit.

When a household bought a property, the transaction was enacted immediately, including all associated costs (e.g., stamp duty, legal fees, etc.). Upon purchasing a property, the involved households reevaluated their financial positions and either withdrew from the market that are (reduced asset liquidity due to purchase) or placed further bids on subsequent properties.

1.3 Validation

The model is validated by comparing the house prices generated in the model to the house price index in Australia. The quarterly house price index for Australia has been moving in the last 3 years between -0.7% and +4.1% (http://www.abs.gov.au/ausstats/abs@.nsf/mf/6416.0) which is within the bandwidth of the model output. We have to acknowledge that the comparison between averages is not in the spirit of this modeling technique, doesn't align with the goal of this project and might even hide underlying discrepancies.

2 Parameterization

2.1 Model Parameters

Global The model was set to run for a period of 20 years with each year split into 52 ticks. The ratio between households and properties was initialized at 1:1 and their absolute number was 10000. Mean rates of return for alternative (non-housing) financial markets were taken from the Reserve Bank of Australia and Australian Stock Exchange.

Households Households were a synthetic population that reflected the distribution found in the Australian census. For each household, an initial amount of assets was tracked and compared to assets available to the household at each time point. Households also had income which increased annually with a global parameter of wage growth (Chatterjee et al., 2016). Household expenses were the combined sum of living and rental or mortgage costs. Each household also had a parameter representing risk appetite that guided their decision to invest in higher yield, higher risk markets vs. low risk, low yield markets.

Properties. Only two attributes defined a property; price and location. Internally, the annualized return of the property in the previous transaction was tracked. Externally defined parameters and their resulting values were also tracked, including property maintenance and depreciation.

Government. The parameters Government could set were the taxes and the tax brackets for income and their deductions. A 'Capital Gains Tax Discount'(CGTD) applied upon sale of a property, and 'Negative Gearing' (NG), which was applied annually.
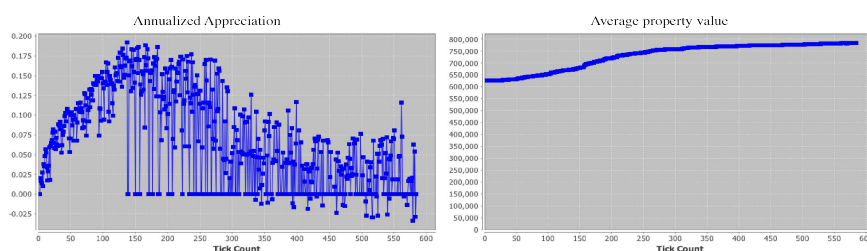
Policy Shock. This scenario was our main test case where the two tax policies of CGTD and NG were adjusted. The scenario varied the tax deduction on salary related to NG investment properties from 100% to 0% and removes the CGTD from 50% to 0%.

Changing Interest Rates. Manipulating interest rates explored the impact of an increase in the bond market with parallel movement of the mortgage interest rate.

3 Results

3.1 Distributions in Steady State

After initializing the model with the information available from ABS and HILDA the model stabilizes after ~550 ticks. During this period property prices increase and the ownership structure changes and moves from homeownership to more renting and investing households. Increasing prices indicated there was additional liquidity in the market that drove prices higher and is due to the modeling where no preference towards liquid assets is given.

Fig. 4. Left: Anticipated Annual Return in % for each week; Right: Property price development

3.2 External settings and Tax Policy Shocks

Effect of income: Changes to income growth had a profound impact on both home ownership distribution and property prices, resulting in new households being priced out of the market and becoming renters. The property price increases outstripped wage growth by a factor of 8. This implied that even if wages increased at a high rate, new households may still not be able to get onto the property market because they may be unable to raise an initial down-payment.

Effect of market sentiment: The model is currently using an equal random distribution of households having an optimistic outlook of the market vs a pessimistic one. This outlook changes the amount the households are willing to bid. By adjusting the center of this distribution households either bidding higher or lower on average. Adjustments to this value has a very strong impact on housing prices with a sustained impact.

Policy Changes: The adjustments tested in this model limited to NG and CGT discounts. After the initial stabilization of the model the parameters are shocked to the point of removing both deductions. The result is a decrease in average housing prices over a longer period of time of ~10%.

4 Discussion

Investigating changes to the tax regime through the lens of our model it becomes evident that the tax regime's impact is limited in comparison to global parameters. Modeling shows a promising investigation path to policy changes. Furthermore, a dynamic model, like the presented ABM, allows to explore different implementation strategies.

The model at hand also provides insights in macroeconomic parameters and their impact on different parts of society. As presented earlier adjustments to the income growth have an impact on the housing prices, what this model allows is the investigation into different income strata and the impacts they feel. A future implementation, for example, is tracking the lost taxes by income decile for each tax adjustment.

Even though the model can be deployed for many cases it is not an accurate representation of the society as an atomic representation of individuals. The results have to be read as a dynamic behavior of the whole and the individual results not as values but als trajectories.

5 Future Work

In a future iteration the presented model will be extended to include changes to population growth, wage growth and supply shock. This would better reflect the international influences that are present in the economy including population pressures due to displacement.

The presented model is already a powerful tool for further investigation. With the atomic knowledge of individual household's investigations into social mobility can be conducted. Many policies are guided by the aim of increasing the social mobility especially where households of lower income can move into the higher income and asset strata. The impact of specific policies though is clouded in reality by a mist of changing environmental parameters and policies. This tool presents a social petri dish where ceteris paribus the impact of changes can be investigated.

Another specific investigation would be looking at the changes to the GINI coefficient. This coefficient that measures the statistical distribution of income or wealth has been associated with life expectancy, child mortality and happiness. Investigating if any policy leads to an accumulation of wealth would allow to show the model's importance for other fields like public health.

References

(ATO), A. T. O. 2016. *In:* FINANCE COMMITMENT, A. B. O. S. A. (ed.). Taxation Statistics, Table 1 Individual, ; Cat. No. 5609.0 Housing.

BERKOVEC, J. & FULLERTON, D. 1992. A general equilibrium model of housing, taxes, and portfolio choice. *Journal of Political Economy,* 100**,** 390-429.

CHATTERJEE, A., SINGH, A. & STONE, T. 2016. Understanding wage inequality in Australia. *Economic Record,* 92**,** 348-360.

DALEY, J. & WOOD, D. 2016. *Hot property: negative gearing and capital gains tax*, Grattan Institute.

LAIDLER, D. 1969. Income tax incentives for owner-occupied housing. *The taxation of income from capital*, 50-76.

ROSEN, H. S. 1985. Housing subsidies: Effects on housing decisions, efficiency, and equity. *Handbook of public economics*. Elsevier.

VICKREY, W. 1961. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16, 8-37.