Agent-Based Modelling of House Price Evolution

Terry Bossomaier, Siti Amri and James Thompson
Centre for Research in Complex Systems
Charles Sturt University
Bathurst NSW
Australia.

March 4, 2005

Abstract

Real estate prices are an intriguing mixture of factors. Spatial attributes combine with speculation and economic trends to create a highly dynamic and unpredictable phenomenon. We argue in this presentation that agent based modelling is the only general method for modelling price evolution.

We first survey existing methods based on statistics, neural networks and other machine learning methods. Then we describe our RePast based model which incorporates vendor/buyer agents and other human agents, such as marketing agents and financial organisations.

The primary decision making processes for all agents uses fuzzy logic. This has the advantage of capturing the semi-quantitative nature of human decision making. We describe the fuzzy logic tools we have added to RePast.

The decision factors fall into three types:

1. The spatial attributes. GIS data is used to determine for each property a range of attributes, including distance from amenities, such as parks, area, elevation, orientation and so on.
2. Environmental factors, such as flood risk (important for Bathurst, our initial area of study). The beauty of agent based modelling is that such factors can be included, based on information from other sources without using specific case data as would be required for many machine learning techniques.
3. Human factors. These include perceptions about the economy, interest rates, new developments such as factories and freeways, plus social trends in the desirability of house ownership and property investment.

Data has been collected and analysed for the city of Bathurst over a 12 year period. The dataset is partitioned into a parametrisation set, to optimise fuzzy rules, and a prediction/validation set to test the generality of the model.

The presentation will show the RePast model in action, describe our findings on spatial and other factors and discuss the general findings for agent-based modelling.

1 Introduction

The advent of Complex Systems Theory and Application as a discipline in its own right is changing the way we model many characteristics of the physical, biological and social world. Whereas aggregate models subserve many purposes, they often do not perform well for complex systems, characterised by many interacting entities. In the geographic world cellular automata models of urban development have been around for some time, such as the SLEUTH models of Clarke et al. [1] urban growth in Santa Barbara, or the fractal cities of Batty and Longley [2]. Other models of urbanisation or land transformation use data fitting techniques, such as neural networks [3].

The " new wave" of complex systems modelling in geography is what Benenson and Torrens call *geosimulation*, in a 2004 special issue of Computers, Environment and Urban Systems [4]. These are agent models, where the interacting entities include people, with all their inherent complexities and difficulties. For example, Arentze and Timmermans [5] develop a model of retail site development following interactions among numerous heterogeneous stakeholders such as planners, developers and suppliers.

Miller et al [6] build a very large scale model, ILUTE (Integrated Land-Use, Transportation, Environment) in which a large number of individuals are tracked through life, using a combination of rule-based decision making and random-utility functions. Their model, which includes real-estate values, is not yet fully calibrated or validated.

1.1 Agent Based Modelling

Brain Arthur was one of the pioneers of complex systems thinking in economics. He addressed the issue of bounded rationality in economic simulations and explored the idea of *calibrated agents*, where the agents were matched directly against the choice behaviour of human subjects determined experimentally [7]. Arthur's pioneering studies showed that such calibrated simulations may verge from optimal or equilibrium states.

Bonabeau [8] suggests several criteria for the use of ABM: capturing emergence; natural description; and flexibility. The methodology is already in use at NASDAQ and Ernst and Young for financial modelling [8]. All these criteria are met by house markets. The boom bust cycles, which everybody would like to predict, seem intractable, while as noted in section 1.4, house valuers still consistently make large errors and miss major trends. Attempting to model individual buyer behaviour is interesting, and the subject of much ongoing research, both within and across countries [9]. Flexibility is important across several dimensions, such as regional changes, government taxation, and external factors such as the attractiveness of the share market as an investment vehicle. ABM can graft on these different factors as required, while continuing to add more and more accurate vendor/buyer behaviour as research becomes available.

1.2 Housing Choice

Many factors contribute to housing choice [10]. Macro-level factors relate to general economic conditions. Micro-level factors usually include income, household composition, found to be particularly important by Yates [11], and other objective quantities.

But Coolen and Hoekstra [10] argue that there are other significant micro factors relating to individual values, which they assess using interviews and means-end theory.

For the purposes of an ABM there are limits to the complexity of choice, which the model can exploit. If factors are unknown, or vary widely across a given population group, then too many factors become impossible to optimise. Furthermore, any optimisation is likely to be data specific and to generalise badly [12]. Hence we aggregate factors as follows:

- **Macro-level factors** are condensed into the single parameter *market perception*
- **Spatial factors** are objective quantities which relate to house position, elevation, orientation and so on which we compute precisely from GIS data, although real world values may not perfectly reflect cognitive space [13]. In a qualitative study in the UK, Ireland and Australia, Daly et al. [9] found location and proximity to amenities to be the most important attributes of house value while the Militino study [14] also found spatial factors to be dominant.
- **Attitude factors** which define how eager a vendor is to sell or a buyer to buy and how greedy he or she is, i.e. how long accepting an offer will be delayed until it meets some internalised financial target.
- **Motivational factors** determine the reasons for buying and selling, principally first home ownership, trading up or down (usually linked to family size) moving into a neighbourhood and investing.

1.2.1 Fuzzy Logic

There are numerous ways of representing human decision processes, and neural networks might seem the obvious choice. However, they are something of a black box technique. In modelling social systems, we want to be able to introduce heuristics based on psychology, sociology and market research.

Fuzzy logic [15] is a good candidate, since it captures the way we think. Essentially our assessments are not binary but more diffuse – close to a school is not less than some exact distance, but a more malleable quantity. Fuzzy systems are used in GIS modelling [16], but also in the modelling of personality itself [17, 18]

1.3 Other Approaches

Other approaches to house price simulation include soft computing methodologies such as neural networks [19] for house prices in South Australia. Li and Revesz [20] develop state of the art spatial-temporal interpolation techniques and apply them to real estate in Lincoln, Nebraska. Militino et al. use a variety of data fitting approaches for Spanish communities [14]. Gelfand et al. [21] study Baton-Rouge, Louisiana by fitting spatio-temporal stochastic processes, while Case et al. [22] compare four different methods for Inner Fairfax County.However, none of the alternative approaches have the flexibility and the ease of examining different scenarios into the future as ABM.

1.4 Valuation

Although housing vendors and buyers have definite desires and goals in house purchase, the price they pay is strongly influenced by valuers. Trust in them may be misplaced. Daly et al. [9] uncovered some disturbing factors:

1. valuers frequently ignore buyer behaviour with an undue emphasis on physical attributes of the property
2. they are strongly affected, possibly pressured, by lenders who essentially desire confirmation of the loan they propose

These characteristics of vendors led many of them to court in the UK housing crash of the early 1990s, where many ordinary wage earners faced bankruptcy as their house value fell below their loan and lenders foreclosed. Daly et al [9] assert that little changed thereafter in valuer skill and methodology.

1.5 Bathurst as a Model Domain

Bathurst is Australia's oldest inland city, with a population of around 35,000. The university accounts for almost 4000 people during term time and thus has a significant economic and residential impact. Bathurst is especially interesting on the house price front. A decade ago, prices in several bands were comparable with Canberra for a given size of property and land, yet the boom in Canberra has not been realised in Bathurst. The long term trends, however, have some key factors:

- the pressures on Sydney are increasing and it is sometimes argued that Sydney is reaching the limits to growth;
- a recent report shows that Bathurst's infrastructure can support an almost doubling of the number of houses without significant development;
- the Great Western Highway to Sydney is continually improving. The distance to Sydney is 200km, but much of the road is single carriageway with 60 and 80km speed limits. Journey times could be cut by up to an hour when road widening is complete.
- the small size of the city means that variations between suburbs, although they exist, are relatively small. But there are some strong negatives. Some streets are highly prone to subsidence, while houses near the river are at risk of floods. A serious flood a few years ago caused 450 houses to be evaculated.

Perhaps, most important of all, is the development of a technology park. At the time of writing the feasibility study is complete, the appropriate land has been gazetted, and the initial funding is being organised. Some estimates put the job increase at over 5,000, adding pressure to the housing market. Labour markets are important factors in mobility and house price effects [11].

Closely related to the impact of the technology park, is what Richard Florida describes as the rise of the creative class [23]. The new generation of graphics designers, new media experts, software engineers and other knowledge workers, brings a new demographic to labour markets. Such people choose the area where they wish to live first, and the job second, and according to Florida, tend to be biased to outdoor recreation. Thus we can expect attractive regional areas to be favoured. Byron Bay in Northern New South Wales is reputed to have the largest community of new media workers in Australia, outside of Sydney.

The long-lived house price book of the last decade led to an obsession with investor property, and unreachable prices in many parts of Sydney. As a result Bathurst attracted investor interest – as a place where middle income people could afford an investment property.

2 Methods

The agent based modelling system we use is RePast, written in Java. Although RePast has a wide range of tools, for the present Project, additional software is required.

2.1 Stage 1: Preparing initial data

ARCInfo was used for processing spatial data from the city provided by Land and Property Information NSW. Each of the spatial attributes was calculated using ARCInfo. This information was then indexed using the $parcel\_id$ attributes present in the exported data. The spatial factors then influence how buyer and seller modify the price relative to the neighbourhood.

2.2 Stage 2: Determination of fuzzy rules for spatial attributes

In earlier work [24] we describe the process of fitting house prices to spatial data. Two methods were used for cross reference – feed forward neural networks trained using Levenberg-Marquadt and zeroth order Sugeno fuzzy inference systems. The fuzzy inference system was simplified using subclustering as described by Chiu [25]. Both gave comparable results and the fuzzy logic was transferred to the next stage.

The complexity of the neural and fuzzy modelling made the Matlab Fuzzy Logic and Neural Network toolboxes a practical choice for the spatial analysis. The integration with RePast presented some challenges, however. Both Matlab and Java operate within their own environments, in the case of Matlab a workspace operating on ASCII scripts (m-files) with access to compiled subroutine libraries, in the case of Java a virtual machine, which dynamically processes byte codes. It is feasible to call Java RePast library routines from Matlab at the top level. But it is not efficient to repeatedly call Matlab from within Java.

Thus a fuzzy logic toolbox in Java was sourced for RePast integration. The NRC FuzzyJ toolbox fulfilled our requirements. Integration within the agent framework is now straightforward. Each agent (buyer/vendor in the present simulation) now has a number of attributes that are fuzzy variables, as detailed in section. These variables are listed in tables 2, 3 and 4 and include eagerness to buy/sell and the agent's perception of the market.

The houses are represented in the system by a **Parcel** JAVA class (representing a house) that incorporates all the spatial information from the local council records and stores them as simple data types. This information is managed internally by an abstract **SpatialHandler** JAVA class that allows the agents to access the spatial attributes while maintaining the integrity of the information.

2.3 Stage 3: Allocation of agent characteristics

The characteristics of each buyer and vendor fall into two categories, the first objective, the second subjective.

The objective parameters include income, family data (kids) and various other factors, which are all obtained from Census data. The Australian Census operates on Collection Districts (CDs) of around 200 houses Thus for each CD we have the income and other statistics from which agents are created at random using the distribution for the CD. (The data is not available at the level of single households but only as CD aggregates).

In the current version of the model, we consider only buyers and vendors from within the Bathurst area, as defined by the Census. The effect of investment from Sydney and elsewhere, which has become particularly significant in the last 5 years) is the subject of future work.

The subjective parameters are the attitudes of individuals to negotiation and their perceptions, such as how they perceive the housing market. These parameters are set using normal distributions about average values of 0.5 in a range of 0-1.

2.3.1 Setting the Price

At present the price is set in a straightforward way. The sales within a 1km radius are averaged. This has the disadvantage that it does not correct for inflation or deflation which has already occurred in the neighbourhood. This price value is then combined with the interest and unemployment rates, obtained from national data, to create the market perception.

2.4 Stage 4: Assembling and running the model

The agents now need to be assigned to a spatial grid. The RePast GIS module is used to import the Bathurst maps directly, hence the agents are located on a heterogeneous network reflecting real data.

The interaction and visibility between agents are governed by the core functionality itself, which is divided up into several processes. Firstly a randomly uniform number of **Vendor** agents (the sellers) between numbers specified at runtime are created and linked to a **Parcel** (or house), specified at random, which they will be selling. Also their greed and eagerness factors are also generated as random numbers distributed uniformly between 0 and 1. The next action is to create the buyer agents with the same attributes as the vendors.

Once the agents have been created the agent interaction can commence. The decision systems used to determine the behavior is largely fuzzy logic based with a constant variation equation on such things as setting house price and determining a bidding price. Firstly the Vendor makes an estimation of what their house is worth by a weighted factoring in of their own houses last sale price and the recent sales in the area as stored in the systems transaction log. After evaluating their house the buyer uses its fuzzy rules system to determine if the market is supportive of them selling their house, at which point they decide wether or not to put their house on the market. That complete, now each buyer agent sequentially traverses the list of vendor agents for a house that can be afforded and then they make a **bid** for that house based on both socio-economic factors and the spatial variables of the house. These factors are examined by a hybrid system of fuzzy logic and static equations to produce a decision on whether the buyer will bid or hold out for a better deal. The vendor agent then assesses the offer made by the buyer, which is once again a fuzzy logic process, and either accepts or rejects the offer.

If a house has been on the market for longer than a simulated year then it is removed from the list to be randomly assigned to another vendor at a later time. (In keeping with actual market practices).

This cycle is repeated until all the houses extracted from the data provided by the local council have been sold (in order for a comparison to be made against the existing data).

The model is run with a time step of 3 months. Sales are shown on the Bathurst map and as cumulative statistics of house price growth average.

2.5 Determining market climate

The fuzzy variables are used in determining the climate of the housing market are given in table 1. The specified fuzzy rules are:

IF InterestRate is Low AND UnemploymentRate is Low AND MedianPrice is Hot THEN MarketPerception is Positive

IF InterestRate is High AND UnemploymentRate is High AND MedianPrice is Cool THEN MarketPerception is Negative

IF MedianPrice is Cool THEN MarketPerception is Negative

IF InterestRate is High AND UnemploymentRate is High AND MedianPrice is Hot THEN MarketPerception is Positive

IF InterestRate is Medium AND UnemploymentRate is Medium AND MedianPrice is Hot THEN MarketPerception is Positive

2.6 Deciding whether to sell

The fuzzy variables associated with the vendors decision to sell are given in table 2. The associated fuzzy rules are:

IF MarketPercept is Positive AND Eager is High THEN VendorIntent is Sell

IF MarketPercept is Negative AND Eager is Low THEN VendorIntent is Hold

Fuzzy Variable	Fuzzy set description	Membership function	Range
InterestRate	Low	ZFuzzySet	0 - 10
InterestRate	Medium	GaussBell	7 - 13
InterestRate	High	SFuzzySet	10 - 15
UnemploymentRate	Low	ZFuzzySet	0 - 10
UnemploymentRate	Medium	GaussBell	7 - 13
UnemploymentRate	High	SFuzzySet	10 - 15
MedianPrice	Cool	ZFuzzySet	-10 - 2
MedianPrice	Hot	SFuzzySet	0 - 10
MarketPerception	Negative	ZFuzzySet	0 - 1.5
MarketPerception	Positive	SFuzzySet	0.5 - 2

Table 1: **marketpercept** (used to determine if the market is positive or negative)

Fuzzy Variable	Fuzzy set description	Membership function	Range
MarketPercept	Positive	SFuzzySet	0.5 – 2
MarketPercept	Negative	ZFuzzySet	0 – 1.5
Eager	Low	ZFuzzySet	0–0.6
Eager	Medium	GaussBell	0.4 – 0.8
Eager	High	SFuzzySet	0.4 – 1
VendorIntent	Sell	Triangle	1 – 2
VendorIntent	Hold	Triangle	0 – 1

Table 2: **vsell** (used to determine if the vendor will put the house up for sale)

Fuzzy Variable	Fuzzy set description	Membership function	Range
MarketPercept	Positive	SFuzzySet	0.5 – 2
MarketPercept	Negative	ZFuzzySet	0 –1.5
Eager	Low	ZFuzzySet	0– 0.6
Eager	Medium	GaussBell	0.4 – 0.8
Eager	High	SFuzzySet	0.4 – 1
VendorIntent	Accept	Triangle	1 – 2
VendorIntent	Reject	Triangle	0 – 1

Table 3: **vaccept** (used to determine if the vendor accepts a buyers offer)

Fuzzy Variable	Fuzzy set description	Membership function	Range
MarketPercept	Positive	SFuzzySet	0.5 – 2
MarketPercept	Negative	ZFuzzySet	0 – 1.5
Eager	Low	ZFuzzySet	0 – 0.6
Eager	Medium	GaussBell	0.4 – 0.8
Eager	High	SFuzzySet	0.4 – 1
MakeOffer	Match	Triangle	1 – 2
MakeOffer	Nomatch	Triangle	0 – 1

Table 4: **MakeOffer** (used to determine if the buyer makes an offer on the house)

2.7 Accepting an offer

The decision of whether to accept an offer uses the fuzzy variables in table 3. The rules are as follows:

IF MarketPercept is Positive AND Eager is High THEN VendorIntent is Accept

IF MarketPercept is Negative AND Eager is Low THEN VendorIntent is Reject

2.8 Making an offer

The buyer makes a decision to either match the proposed price from the vendor, or reduce the price according to the following rules, with the fuzzy variables given in table 4. The rules are:

IF MarketPerception is Positive AND Eagerness is High THEN MakeOffer is Match

IF MarketPerception is Negative AND Eagerness is Low THEN MakeOffer is Nomatch

3 Discussion

From the information collected, a simplified model of the system has been developed. Its output includes a overview mapping of transcations in Bathurst (figure 1), with a

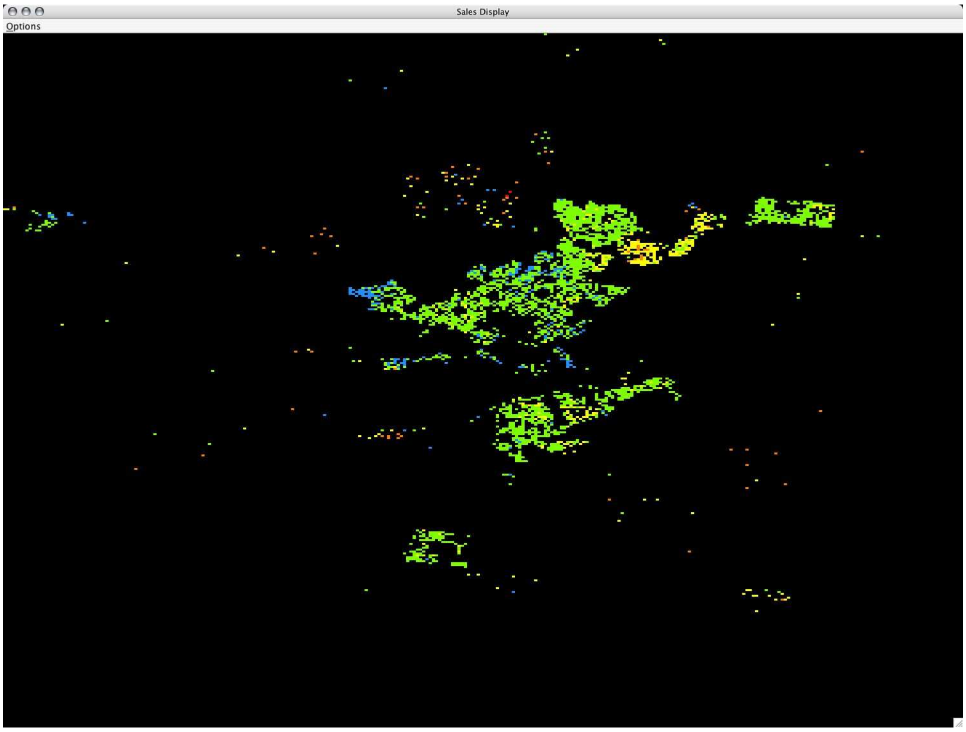

Figure 1: Transactions displayed on the map of Bathurst

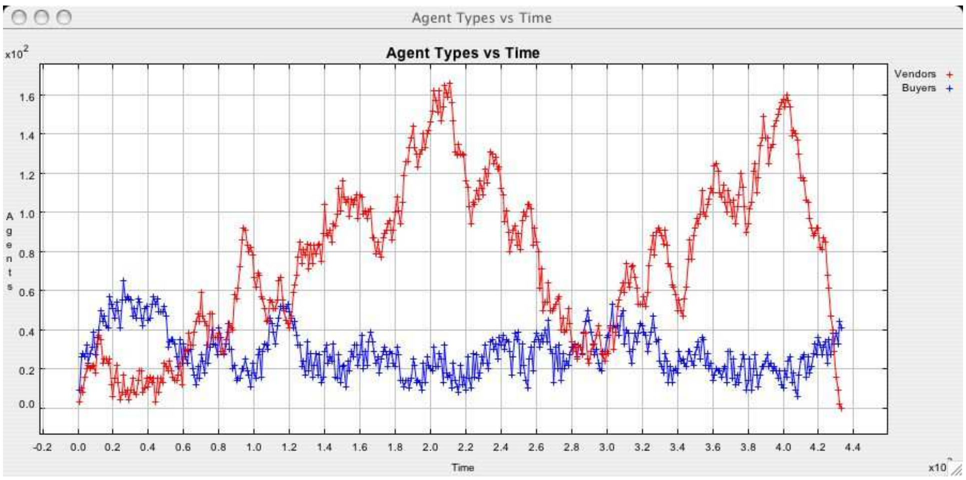

Figure 2: Number of agents as a function of time. The vendor agents accumulate when buyers are reluctant to sell

Figure 3: Average sales through a simulation with no inflationary pressures

graph showing the trends in agent population in figure 2, and finally a graph showing the average sales for the given time period in figure 3

We have described the first verion of our agent-based model of house prices. Real GIS data is combined with fuzzy logic to determine buyer and seller transactions.

The next generation of the model will optimise the fuzzy parameters against subsets of housing data using genetic algorithms, following a similar approach to the optimisation of cellular automata models of urban growth by Clarke and Goldstein [26, 27].

The next decade is likely to be quite interesting for the Bathurst housing market with our simulations able to show the growth for various scenarios in the offing.

References

[1] K.C. Clarke, S. Hoppen, and L. Gaydos. A self-modifying cellular automaton model of historical urbanization in the san. *Environment and Planning B*, 24:247–261, 1997.

[2] M. Batty and P. Longley. *Fractal Cities*. Academic Press, 1994.

[3] B.C. Pijanowski, D.G. Brown, B.A. Shellito, and G.A. Manik. Using neural networks and GIS to forecast land use changes: A land transformational model. *Computers, Environments and Urban Systems*, 26:553–575, 2002.

[4] I. Benenson and P.M. Torrens. Geosimulation: object-based modelling of urban phenomena. *Computers, Environment and Urban Systems*, 28:1–8, 2004.

[5] T. Arentze and H. Timmermans. A multiagent model of negotiation processes between multipe actors in urban developments: a framework for and results of numerical experiments. *Environment and Planning B: Planning and Design*, 30:391–410, 2003.

[6] E.J. Miller, J. Douglas-Hunt, J.E. Abraham, and P.A. Salvini. Microsimulating urban systems. *28*, pages 9–44, 2004.

[7] W.B. Arthur. On designing economic agents that behave like human agents. *Evolutionary Economics*, 3:1–22, 1993.

[8] E. Bonabeua. Agent-based modelling: Methods and techniques for simulating human systems. *PNAS*, 99:7280–7287, 2002.

[9] J. Daly, S. Gronow, D. Jenkins, and F. Plimmer. Consumer behaviour in the valuation of residential property. *Property Management*, 21(5):295–314, 2003.

[10] H. Coolen and J. Hoekstra. Values as determinants of preferences for housing attributes. *J. Housing and the Built Environment*, 16:285–306, 2001.

[11] J. Yates. A spatial analysis of trends in housing markets and changing patterns of household structure and income. Technical report, Australian Housing and Urban Research Institute, May 2002.

[12] M.J. Kearns and U.V. Nazirani. *An Introduction to Computational Learning Theory*. MIT Press, 1994.

[13] D.M. Mark, C. Freksa, S.C. Hirtle, R. Lloyd, and B. Tversky. Cognitive models of geographic space. *Int. J. Geog. Inform. Sci*, 13(8):747–774, 1999.

[14] A.F. Militino, M.D. Ugarte, and L. García-Reinaldos. Alternative models for describing spatial dependence among dwelling selling prices. *J. Real Estate Finance and Economics*, 29(2):193–209, 2004.

[15] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1967.

[16] S. Dragicevic, D.J. Marceau, and C. Marios. Space, time and dynamics modelling in historical GISS databases: A fuzzy logic approach. *Environment and Planning B: Planning and Design*, 28:545–562, 2001.

[17] N. Ghasem-Aghaee and T.I. Ören. Towards fuzzy agents with dynamic personality for human behaviour simulation. In *Proc. Summer Simulation Conf. Montreal*, 2003.

[18] T.I. Ören and N. Ghasem-Aghaee. Personality representation processable in fuzzy logic for human behaviour simulation. In *Proc. Summer Simulation Conf. Montreal*, 2003.

[19] P. Rossini. Improving the results of artificial neural network models for residential valuation. In *Proc. 4th Annual Pacific-Rim Real Estate Society Conference*, 1998.

[20] L. Li and P. Revesz. Interpolation methods for spatio-temporal geographic data. *Computers, Environment and Urban Systems*, 28:201–227, 2004.

[21] A.E. Gelfand, M.D. Ecker, J.R. Knight, and C.F. Sirmancs. The dynmaics of location in house price. *J. Real Estate Finance and Economics*, 29(2):149–166, 2004.

[22] B. Case, J. Clapp, R. Dubin, and M. Rodriguez. Modelling spatial and temporal house price patterns: A comparison of four models. *J. Real Estate Finance and Economics*, 29(2):167–191, 2004.

[23] R. Florida. *The Rise of the Creative Class*. Basic Books, 2002.

[24] S.Z. Amri and T.R.J. Bossomaier. The influence of spatial attributes on house price. In *Proc. 7th. Asia-Pacific Conference on Complex Systems*, 2004.

[25] S. Chiu. Fuzzy model identification based on cluster estimation. *J. Intelligent and Fuzzy Systems*, 2(3), 1994.

[26] N.C. Goldstein. Brain vs. brawn – compoarative strategies for the calibration of a cellular automata-based urban growth model.

[27] N.C. Goldstein, T.J. Candau, and K.C. Clarke. Approaches to simulating the march of bricks and mortar. *Computers, Environment and Urban Systems*, 28:125–147, 2004.