

©Copyright JASSS

## Alison Heppenstall, Andrew Evans and Mark Birkin (2006)

## Using Hybrid Agent-Based Systems to Model Spatially-Influenced Retail Markets

Journal of Artificial Societies and Social Simulation vol. 9, no. 3 <http://jasss.soc.surrey.ac.uk/9/3/2.html>

For information about citing this article, click here

Received: 23-May-2005 Accepted: 20-May-2006 Published: 30-Jun-2006



## Abstract

One emerging area of agent-based modelling is retail markets; however, there are problems with modelling such systems. The vast size of such markets makes individual-level modelling, for example of customers, difficult and this is particularly true where the markets are spatially complex. There is an emerging recognition that the power of agent-based systems is enhanced when integrated with other AI-based and conventional approaches. The resulting hybrid models are powerful tools that combine the flexibility of the agent-based methodology with the strengths of more traditional modelling. Such combinations allow us to consider agent-based modelling of such large-scale and complex retail markets. In particular, this paper examines the application of a hybrid agent-based model to a retail petrol market. An agent model was constructed and experiments were conducted to determine whether the trends and patterns of the retail petrol market could be replicated. Consumer behaviour was incorporated by the inclusion of a spatial interaction (SI) model and a network component. The model is shown to reproduce the spatial patterns seen in the real market, as well as well known behaviours of the market such as the "rocket and feathers" effect. In addition the model was successful at predicting the long term profitability of individual retailers. The results show that agent-based modelling has the ability to improve on existing approaches to modelling retail markets.

## Keywords:

Agents, Spatial Interaction Model, Retail Markets, Networks

## Introduction

1.1

Many geographical systems can be viewed as complex entities containing numerous non-linear processes which may be interrelated at different spatial and temporal scales. This is particularly evident in retail markets. Here, companies make decisions on pricing whilst consumers make decisions on where to purchase one or more commodities. The competition between retailers to attract customers, and therefore make profits, can result in complex spatial interactions between all participants within the system. This is especially true when companies' product prices are set dependant on competitors' prices within geographically localised markets. Given the high degree of complexity within such geographical systems, it is unsurprising that traditional equation-based techniques have failed to produce realistic representations of the behaviour involved.

1.2

Over the last few years, agent-based modelling has become an increasingly popular and powerful tool for investigating socioeconomic systems (<u>O'Sullivan and Haklay 2000</u>). The ability to model at an individual level has seen the rise of diverse marketcentred applications that generally focus on modelling either supply or demand scenarios. For example, on the supply side, agents are commonly used to investigate the potential for improving performance within supply-chain management (<u>Garcia-Flores et al.</u> <u>2000</u>; <u>Julka et al. 2002</u> and <u>Signorile 2002</u>), whilst on the demand side agents are employed to model consumer behaviour (<u>Ha et al.</u> <u>2002</u>; <u>Constantinides 2004</u>; <u>Pettersen et al. 2005</u>).

1.3

There is an ever increasing literature on the use of agents for modelling retail markets. For example, Kirman and Vriend (2001) developed an agent-based model for examining price dispersion and customer loyalty with fish markets. Bunn and Oliveria (2001) applied a multi-agent evolutionary model to the electricity market in the UK. Further research in this area is ongoing with the Electricity Market Complex Adaptive System (EMCAS 2006), a tool for modelling the entire retail market for building electricity systems in Illinois.

1.4

These applications hint at the possibilities offered by agent-based modeling: the ability to model complex / non-linear market behaviour and processes, ultimately furthering our understanding of such complex systems. However, studies to date have tended to concentrate on temporal complexities, with studies of the spatial complexities of markets being rarer and largely avoided through judicious choices of market and scale (notable exceptions include <u>Abdulai 2000</u>; <u>Castiglione 2000</u>; and <u>Lombardo et al. 2004</u>).

#### 1.5

The spatially-influenced market that will be used as a case-study in this paper is the retail petrol market. In this market petrol retailers ("stations") compete for customers within localised, but overlapping, areas. This market has the advantage of selling a single homogeneous product. This means that it does not possess any of the complex trade-offs in the pricing of products to attract

consumers that are found in many retail markets (for example the grocery market). In addition, the retail petrol market is an important system in its own right. Petrol is a ubiquitous commodity valued by retailers and customers alike. Consumer sensitivity to petrol prices was clearly demonstrated in the UK during August and September 2000 when there was a "Petrol Crisis", consisting of consumer blockages of refineries and protests in reaction to soaring fuel taxes. Over the summer of 2005, the issue of petrol prices has again been in the news, this time due to rises in the price of crude oil resulting from a combination of political instabilities (the Iraq war; Nigerian insurgencies; Venezuelan labour issues) and natural disasters (hurricanes Katrina and Rita). Although there are many possible components of the petrol market that might be modelled, given its importance as a social influence, the modelling of petrol prices is of particular interest. In addition, price provides a key indicator that any market model is working appropriately. This study therefore sets out to model petrol price variation on the regional scale.

### 1.6

Previous research on retail petrol markets has focused on the cause of asymmetries within the response of petrol prices to external shocks. Attention has focused on the transmission of positive and negative changes in a variable, for example crude oil prices or exchange rates, to the price of petrol. According to Galoetti et al. (2003), these studies generally differ in one or more of the following aspects: the country under scrutiny; the time, frequency and period of data used; the focus on wholesale and retail gas prices or oil and petrol prices, and finally the dynamic model employed in the empirical investigation. Examples of work within this area include Bacon (1991), Manning (1991), Shin (1994), Reilly & Witt (1998) and Mitchell et al. (2000).

### 1.7

A critical limitation with each of these studies is their inability to detail which retail forecourts consumers will patronise and why. Simply, this is because such approaches are unable to account for important spatial behaviour in the market (for example, that people choose to buy petrol from a spatially limited subset of all the possible retailers - see Section 2 below). As a result of this, these models are equally unable to account for the specific prices of particular firms and daily pricing dynamics.

#### 1.8

In this paper, an alternative methodology to modelling the retail petrol market is presented. The aim is to model the behaviour of the petrol stations, the rules they employ and strategies implemented. Instead of using aggregate data within an empirical model, interactions between petrol stations are modelled at an individual level by use of agents. Consumer behaviour is incorporated by use of a spatial interaction model and a network component which are linked to the agent system. The development of the complete model is detailed along with sensitivity testing and validation. Initial findings are presented and conclusions made about the applicability of such a methodology for such retail markets.

## 🐬 The Retail Petrol Market

## 2.1

Petrol is one of the most valuable oil derived commodities. At the end of 2002, there were 11,707 sites retailing over 36 billion litres of motor fuel in the UK. This equates to an average of approximately 1,350 litres of fuel consumed by each car/van per annum. While consumers tend to buy fuel as part of a multi-purpose journey (Plummer et al. 1998, Ning and Haining 2003) and demand is the epitome of inelasticity, consumers are becoming ever more aware of local variations in petrol prices and search out the lowest prices along their specific task-centred pathways. This has created both a highly competitive and sensitive market, with organisations employing price-centred strategies to maximise profits. This sensitivity to petrol prices was fully borne out in the UK during August - September 2000 with the "Petrol Crisis" and the associated fuel protests that precipitated in reaction to soaring fuel taxes.

#### 2.2

The UK petrol market is comprised of a variety of different types of petrol retailers: large multinationals; nationals; small independents and, increasingly, supermarkets. Since their entry into the UK market in the mid-90s, supermarkets have become a significant force in petrol and diesel retailing. Initially offering established-brand forecourts, the development of "own-label" forecourts was accompanied by a marketing strategy that focused on substantially undercutting the competition. This ploy was a dual success, both attracting customers to purchase fuel and helping to expand sales for the supermarkets. This resulted in the major players such as Esso, BP and Shell responding with enhanced loyalty schemes and price-matching policies. For example, Esso's "Price Watch" policy states they will match the petrol prices of any station within three miles.

#### 2.3

One of the common problems encountered by researchers studying the retail petrol market is the lack of published information on strategies operated by different companies because of commercial sensitivity. However, a study by Ning and Haining (2003) revealed several factors that retailers considered important when setting their prices. These were: the local competitive environment (who are competitors and where are they located), the price of competitors, and the sensitivity of the consumer (to price and loyalty schemes). In addition to these factors, global influences such as the price of crude oil and geographical location of the station (whether rural or urban) have also been found to play a part in the final price.

### **Petrol Price Data**

#### 2.4

The data set used in this study for model testing consists of daily petrol price readings for the UK taken throughout the months of July, August and September 1999. Geographically, the data set covers the UK and many of the main petrol retailers, for example: internationals such as Esso, BP, Texaco and Shell; supermarket garages such as Sainsbury's, ASDA, TESCO and Morrisons, and numerous "independents". The prices of the four main petrol types of unleaded, super-unleaded, diesel, and four-star are recorded; in total there are over 16,000 petrol stations in the dataset.

#### 2.5

The experiments presented here use the unleaded petrol prices (as these are recorded at every station) within the region of West Yorkshire (Figure 1). Figure 1 shows the location and spatial variability of prices within the region for a typical day in the dataset (by means of an interpolated price surface; only areas with information available are shown). In general, rural areas sustain higher prices than urban areas.



Figure 1. Map showing the location and spatial variability of retail petrol prices within West Yorkshire. Urban areas are indicated by station density and by name

However, this rural-urban price differential is not clear cut. Figure 2 shows that within urban areas, there also exists price differences. Although both cities shown are more expensive than the surrounding areas, the prices can range between 3-4 pence ("p"). In Bradford (Figure 2 (a)), this appears to be due to high priced stations on one of the arterial roads. In the case of Wakefield (Figure 2 (b)), higher prices can be attributed to the presence of the motorway (stations located on motorways are typically priced several pence higher than other stations).



Figure 2. Intra-urban price variation for two typical cities

## Modelling the Retail Petrol Market

3.1

It is clear from the discussion of the data presented above that the traditional use of 'global' regression models to examine the relationship between petrol prices and other variables is not sufficient. The recent work by Ning and Haining (2003) highlights the importance of the local environment and thus the influence of geography, while the snap shot of the data above confirms that modelling this type of system requires a method that is able to take account of geographical variations, the local competitive environment and consumer sensitivity. To attempt to take such factors into account, the system was modelled using Geographically Weighted Regression (for an overview of the methodology see Brunsdon et al. 1998a; 1998b). The results from this approach were not promising (see Appendix 1 for brief details), for this reason an alternative method was sought.

## **Petrol Agents**

## 3.2

Agents were chosen because of their ability to model individual behaviour both spatially and temporally. Additionally, it is relatively easy to link agent-models to other systems that represent other processes or behaviour. With any given market there are plainly two sets of behaviours that may be modelled to drive prices: the retailers and the customers; in this model the agents represent the

retailers.

### 3.3

Individual petrol stations were represented as agent-objects using Java and supplied with knowledge of their initial starting price, production costs, and the prices of those stations within their neighbourhood (defined as a fixed-radius circular-distance around the petrol station). The initial prices used varied from real data from day zero of the dataset to specifically generated price surfaces, depending on the behaviour under investigation. Each agent was able to view the prices of neighbouring stations (thus sharing information) and each applied a series of rules to adjust its own prices. In the main, these rules involve the implementation of behaviour based on the price set by the neighbouring stations and their distance. A typical rule might be "if my price is X pence cheaper or more expensive than my nearest neighbour, set my price to my neighbours". During each iteration of the model the individual stations found out their competitors' prices and adjusted their own. Each iteration represented a day, and iterations were run either until a predefined day for comparison with the real data, or until equilibrium. For the purposes of this research, the development of a hybrid agent system, the agents were kept simple. However, they could easily be developed for application to other complex scenarios.

### 3.4

The agents possess many of the characteristics apparent in other agent applications; these are detailed in Table 1.

Table 1: List of agent	characteristics
------------------------	-----------------

Characteristic	Explanation
Autonomous	React to other petrol agents on basis of individual and local information.
Heterogeneous	Stations can have individual rule sets, price and profit levels.
Communicative and Cooperative	Pricing and location information shared between agents for competition.
Pro-active	Monitor their own profit/price levels and take action to maintain or maximise their profit.
Reactive	Decision making on pricing rules and updating of prices based on information supplied to them.

#### 3.5

Rules can be assigned in the model to stations on the basis of the company running them, geographically, or simply given identically to all the stations. In the cases presented here, the rules were universally distributed and matched general rules derived from industry knowledge. Parameter values were set after experimentation and the choices confirmed by comparison with a genetic algorithm calibration (see <u>Heppenstall et al. 2006</u>) (the parameters used along with a definition can be found in Appendix 2). Figure 3 details the rules used and the steps the model took during each iteration.



Figure 3. Flowchart illustrating the operation of rules within the petrol agent during one iteration. For the values of parameters like "maximum price change allowed, see Appendix 2

3.6

Initially the system was run just on the price-centred rules as outlined in Figure 3 with no consumer input, under the purposefully

naïve assumption that the stations work to just undercut nearby competitors while maintaining a minimum price (and are not concerned with distant competitors). The study was run to examine whether the price dynamics in the real system could be replicated simply by stations altering their prices in local markets competitively, with non-equilibrium responses being driven by external shocks. This is actually not an unreasonable scenario as for example, Esso claim that they set prices at least in part, by simply matching the price of local competitors (their "Price Watch" policy), intimating at a system detached at one step from actual customers and profits.

## 3.7

Initial experimentation detailed in Heppenstall et al. (2005) showed that using these rules alone was ineffective at recreating the patterns and trends in the real data: the petrol agents were not competitive enough. Even though petrol consumption is relatively insensitive to price changes and consumers tend to buy petrol on the way to complete other tasks, the model still needs consumer choice to accurately reflect real profit-driven trends.

## Modelling Consumer Behaviour and Profit

### 4.1

A full agent-based customer model including task-orientated travel and petrol consumption is a significant computational undertaking, even if the appropriate data existed (which it does not in the UK). The work presented here represents the first step towards such a larger model. The solution that would both represent consumer behaviour and introduce a profit-based driver to the agent-model was the incorporation of a spatial interaction (SI) model, coupled to the agent-system to create a "hybrid" model. As the vast majority of journeys in the UK are to and from work and this is likely to be when most people buy their petrol, the SI-model was coupled to a travel to work network to supply consumer pathways through the market.

#### 4.2

Traditionally, SI-models are used for "what-if?" scenarios (<u>Birkin et al. 2002</u>). For example, what will happen to the sales of supermarket X if supermarket Y locates nearby? Within this research, the SI-model was used to calculate customer numbers attracted to each station, and by assuming a constant customer fuel requirement, volume of sales information was generated. This information was then used to calculate the amount of profit that was being made. Each petrol station implements profit maximising strategies based on this information,

- 1. If the profit is rising, continue implementing the last price change;
- 2. If the profit is dropping, increase the price. If this does not work, decrease the price;
- 3. If the profit is constant across two iterations (within a margin of tolerance) keep the price constant.

Profits are calculated using appropriate figures for production and retailing costs (see <u>Heppenstall et al. 2005</u> for further details).

### 4.3

The SI-model is in two parts, which are represented by Equations (1) and (2) below. Equation (1) calculates the *relative* amounts of fuel sold by each station to the population of each ward (UK Census area of ~100 to 30,000 people). The amount sold to a given ward decays exponentially with the distance between the ward centroid and the station, and also with rise in price. The exponential fall off is traditional in SI-models and well matched by many retailing examples (Birkin et al. 1996). Equation (2) ensures that the total volume of each fuel sold to each Ward equals the demand  $H_i F^m$  (the amount of fuel type *m* required by car owners in ward *i* per day — in the case presented here *m* was limited to one fuel, unleaded petrol). This gives the actual amount sold in each ward.  $\lambda$  and  $\beta$  are two coefficients which are usually determined by calibrating the model to real data (values used here can be found in Appendix 3; for details of the specific calibration see Heppenstall 2004; Heppenstall et al. 2006).

$$\hat{S}_{ij}^{m} = \delta_{j}^{m} \exp\left[-\beta d_{ij} - \lambda p_{j}^{m}\right] (1)$$

$$S_{ij}^{m} = \frac{\hat{S}_{ij}^{m}}{\sum_{j} \hat{S}_{ij}^{m}} H_{i} F^{m}$$
<sup>(2)</sup>

where:

 $S^{m}_{ii}$  is the amount of fuel *m* sold by garage *j* to ward *i*.

 $\delta^{m_{j}}$  is 1 where garage *j* sells fuel m and 0 otherwise.

 $d_{ij}$  is the distance between ward *i* and garage *j*.

 $p^{m}_{i}$  is the price of fuel *m* at garage *j*.

 $H_i$  is the number of households within the ward *i*.

 $F^m$  is the amount of fuel of type *m* required per car per day.

ε is a stochastic term.

The stochastic term in the calculation of the weighted distribution of sales allows for the representation of imperfect knowledge and non-rational behaviour in customers and for the addition of a random perturbation to the model. For the purposes of this paper  $\epsilon$  is taken as being normally distributed with a mean of 0.0 pence and a standard deviation of 0.05 pence.

The interaction between the agent-model and SI-model can be summarised as:

- 1. The agent-model and population are initialised with the real data.
- 2. The fuel price and station location data is passed to the SI-model.
- 3. The SI-model calculates fuel sales for each station on the basis of potential customers and passes the information to the stations.
- 4. The stations use this information to choose the rules to implement based on their profitability (see above) and new prices are generated.
- 5. The model is iterated from steps 2 to 4 until completion (equilibrium or a set number of days, at one iteration a day).

## **Consumer Movement**

### 4.5

The SI component models the number of consumers that buy fuel at petrol stations based on distance and attractiveness (closeness and price). However, this approach has an over-simplified notion of when people buy fuel. As has been noted, most people tend to buy fuel on the way to somewhere, in most cases on the way to work. The SI-model, however, uses ward populations. If the standard home-location ward population figures available in the UK census were used, the model would be representing everyone as travelling directly from home to the nearest petrol station where fuel is cheap. To rectify this disparity, a network component was added. The network component provided a means by which consumers could be distributed across the study area on the basis of travel to work data (derived from UK Census data). It essentially provides a snapshot of people's locations on the way to work to use as a base population for the SI in preference to the home-location populations.

#### 4.6

To achieve this network component the optimal route between the origin and destination of each consumer needed calculating. This shortest path problem was solved using Dijkstra's Algorithm. In order to re-distribute the population around West Yorkshire using this approach, nodes and edges were constructed. The nodes represented the population weighted centroids of the wards and the edges represented the contiguity between neighbouring wards and were weighted for centroid-to-centroid distances (Figure 4). Once the shortest distance through the intervening network of wards was calculated for each set of ward-to-ward movements, car users were then redistributed to the home, destination and intervening wards with an equal probability for each ward.



Figure 4. A section of the weighted graph network

### 4.7

The network component is run before any of the interactions between the agent and SI-model take place. The redistributed population surface that is created is fed into the SI-model. Figure 5 shows the distribution of consumers before and after the networking was calculated. Figure 5 (a) clearly shows that before the networking was applied, high densities of consumers were found predominately in the centre of the study area. This area covers the West Yorkshire conurbation of Leeds, Bradford, Wakefield, Huddersfield and Halifax. The density of consumers is at its highest in the suburban areas of Leeds and Bradford. After the networking has been applied (Figure 5 (b)), all of the city areas, not just the suburbs, are characterised by having an even higher density of consumers. This increase in population within the city centres is simply a reflection of the number of people who live within the suburbs and beyond, but work in the cities.



Figure 5. Distribution of consumer density before (a) and after (b) application of the networking

## 🐬 Comparison of Model Outputs

## 5.1

The system is assessed initially through comparison with the real data. The model is initialised with real data from day zero, and allowed to run for ten days (the period over which we have a reliable and consistent real dataset in which there are no major external shocks such as an oil price rise). The prices at stations are then compared with the real prices from day ten of our dataset (see Table 2). The system is also assessed for accuracy of predictions against a 'business as usual' (or static price) model generated by assuming the prices at day zero are maintained over the ten days. In addition, to see if the inclusion of the network component significantly affected the results of the model, identical experiments were run with and without the option to redistribute the consumers using the network component (see Appendix 3 for simulation details). For the purposes of the following results, the Hybrid - model is the Agent - model plus the SI-model, while the Network - model is the Hybrid - model plus the network component.

## 5.2

A suite of goodness-of-fit tests were chosen to provide an overall view of how the models were performing. These included mean error, standard deviation of the error (SD), root mean square error (RMSE) and mean absolute error (MAE). The RMSE provides an overall error in pence, but with more importance placed on outliers (this allows assessment of unpredicted or overpredicted extremes in pricing, for example in rural areas). The MAE provides an overall error measure in pence, but considers only the magnitude of the error, not whether it is negative or positive. The Hybrid and Network models, both with and without the redistribution were also run with and without the stochastic element outlined in Section 4 to gauge the variability of the results under a fixed range of perturbations.

 Table 2: Comparison of the Hybrid and Network models running the models in deterministic and stochastic mode. The model was run ten times in stochastic mode to generate the ranges. The first and second rows are the mean and standard deviation of the errors between the real and predicted data

Test		Hybri	d - Model	Netwo	rk-Model
	Business as usual	Deterministic	Stochastic	Deterministic	Stochastic
Mean	0.160	0.138	0.137 — 0.143	0.152	0.149 — 0.155
SD	0.976	1.007	1.004 — 1.012	1.037	1.036 — 1.044
RMSE	0.987	1.018	1.012 — 1.019	1.046	1.044 — 1.054
MAE	0.773	0.783	0.781 — 0.788	0.806	0.805 — 0.814

5.3

Table 2 shows that, on average, both the Hybrid and Network models are accurately predicting the variations within the real data. There are no notable differences between the results for the deterministic and stochastic models for either the Network or Hybrid models, suggesting model stability. When assessing the models using such aggregate statistics, the Hybrid-model results are slightly closer in value to the real data than the Network-model. However, these differences are negligible.

#### 5.4

The difference in the real data between July 27<sup>th</sup> (the model initialisation data) and August 6<sup>th</sup> (the 10<sup>th</sup> day of the data set used for comparison) gives a mean "error" of 0.160: this is our "business as usual" difference. The Hybrid and Network models generated figures close to this with a joint range of values between 0.137 - 0.155. The RMSE and MAE values produced by these models are also close to the "business as usual" values. However, in general these bulk statistics do not show the dynamic models making any improvement on the "business as usual" model. This is not particularly unusual for such models and should not be regarded as rendering the model invalid (Pontius et al. 2004). Generally it indicates a matched stability in both the real system and the model dynamics. A more important indicator of the quality of the model is whether it can reproduce known behaviours from an unrealistic initiation point — that is, whether a realistic distribution of prices is emergent from the behaviours. In this particular case we are interested in whether the model can recreate the spatial distribution of the petrol prices.

To test the ability of the models to reproduce the spatial patterns within the data, identical experiments were run on the stochastic model with and without the option to redistribute the consumers. In this case, each station was assigned the same initial price (71p, the average price within the data set) with the parameter values outlined in Appendix 3 applied. The models were run for 10 days to allow pricing trends to develop and the results mapped.

## **5.6**

Figure 6 shows that re-distributing the population has, in this example, made a negligible difference to the spatial patterns. However, overall both models are reproducing the same spatial trends and patterns that are evident in the real data. Importantly, this shows the ability of the model to recreate spatial patterns from an unrealistic initialisation, indicating the behaviour captured is likely to accurately reflect real-world processes.



Figure 6. Price distributions for the various models ten days after runs started with all stations initialised at 71p: (a) Hybrid-model and (b) Network-model. Data from day ten of the real data set (c) is included for comparison

## Testing Model Robustness and Sensitivity

## **Random Price Fluctuations**

## 6.1

In this section we will examine the robustness and sensitivity of the system through specific tests. We will measure "robustness" through examining whether the reaction of a large but realistic price change at one station can dramatically change the overall price surface. "Sensitivity" will be measured through assessment of whether such a change knocks the model into unrealistic behaviour. The former measure will be as much an effect of the real-world system as of the model, the latter would indicate a problem with the model alone. In addition, these tests will help inform on any emergent properties within the system.

#### 6.2

To test the robustness and sensitivity the model was run as above using day zero data to predict day ten data. However, an error/price variation was added to the price at one station that was chosen randomly. The chosen station was altered at day zero by either 1 pence, -1p, 3p or -3p. These prices were chosen because previous experimentation showed that these values were large enough to force the system to react at least locally (see <u>Heppenstall 2004</u>).

## 6.3

With the exception of the price drop, the parameter values detailed in Appendix 3 were again used. As there was very little difference in the results generated in Section 5, the Network-model was used. The deterministic version was selected so any resulting variation could be explained by the price drop rather than an unknown element. As in Section 5, the results were compared against the real data on the tenth day of the simulation (Table 3).

 Table 3: Aggregated results of running the model with a random price drop and increase of 1p and 3p at random stations. Each experiment was run ten times. The first and second rows are the mean and standard deviation of the errors between the real and predicted data

Test		Price Drop		Price Increase	
	Real Data	1р	3р	1р	Зр
Mean	0.160	0.183 — 0.196	0.199 — 0.777	0.177 — 0.203	0.179 — 0.200
SD	0.976	1.018 — 1.043	1.031 — 1.590	1.012 — 1.038	1.027 — 1.058
RMSE	0.987	1.033 — 1.058	1.048 — 1.767	1.026 — 1.056	1.042 — 1.078
MAE	0.773	0.788 — 0.803	0.798 — 1.271	0.786 — 0.801	0.786 — 0.810
Mean SD RMSE MAE	0.160 0.976 0.987 0.773	0.183 — 0.196 1.018 — 1.043 1.033 — 1.058 0.788 — 0.803	0.199 — 0.777 1.031 — 1.590 1.048 — 1.767 0.798 — 1.271	0.177 — 0.203 1.012 — 1.038 1.026 — 1.056 0.786 — 0.801	0.179 — 0.200 1.027 — 1.058 1.042 — 1.078 0.786 — 0.810

None of the day ten predictions were greatly dissimilar from the real data. For example, the range in mean differences between the real data and the 1 pence drop model prediction was 0.183 - 0.196, while the 3 pence drop model gave a range of 0.199 - 0.777. By comparison, the "business as usual" model gave a mean difference of 0.160. This may indicate that while there is a change in the overall price surface which scales with the initial price change at a single station, the final bulk effect on the surface is comparatively dampened (however, see Section 6.2). The ranges produced by the SD, RMSE and MAE were similarly small indicating that there were no wildly unrealistic fluctuations in the price surface. This indicates that the model is reacting to the price adjustments in a manner consistent with reality.

#### 6.5

In a detailed analysis of the real data, prices were found to vary across cities by 3-4 pence. However changes in daily price were, on average, never found to be more than 1p. The majority of stations would change their price by a smaller increment of 0.1p or 0.5p. It is assumed that the model would respond to lower price changes in a similar manner to higher price changes. Based on the results in Table 3, it can be concluded that the model would also be little affected by these changes, should they be modelled, and able to continue reasonably. Both model and reality seem dynamic and yet stable under externally uninfluenced conditions.

## Fluctuation in Price at a Single Station

## 6.6

Despite the conclusions above about the robustness of the model, the results from Table 3 do not indicate the spatial effects of the price adjustment. How do the surrounding stations react? Do they rapidly adjust their prices? Is there a diffusion of this new price and how does it affect the profit of the station? To investigate this, the Network-model was initialised with the real data (July 27th) and the default parameter values detailed in Appendix 3 used. Station 8537 (the station at the centre of the city of Leeds) was dropped by 5p (from 73p to 68p). This "extreme" value was used to force the system to react to the price change in a manner that could be visualised easily (c.p. the changes above). Experimentation revealed that the system reacted in a broadly similar manner whether the price change was positive or negative.

## 6.7

Figure 7 shows the diffusion of the new price throughout the study area. By day 20, most of the stations have lowered their prices by an average of 3 - 4p (66 - 67p). On day 40, the prices at the majority of stations have settled to approximately 66p, lower than the original price drop, and do not fluctuate again. The stabilisation is due in part to the absence of any other factors to shock the system, for example separate price dropping by another competitor or unprofitable stations going out of business. The spatial diffusion of the prices in this experiment is also typical of the results found for the price adjustments above. The performance statistics shown in Table 3 were calculated after ten days; evidence from Figure 7 shows that after this time the price drop has only affected the stations in the immediate vicinity, perhaps suggesting that the ten day constraint of the real data is not ideal.



Figure 7. Spatial diffusion of prices over time for a price drop of 5 pence. All stations are initialised with real prices from day zero (July 27 <sup>th</sup>)

The behaviour exhibited by the system is a direct result of the profit maximisation rules implemented by each agent. Given this, several stations were selected to examine profit and pricing behaviour over the course of the above simulation (see Figure 8 for locations).



Figure 8. Location of stations examined in the experiments

Between days 10 and 20, the stations closest in proximity to 8537 (stations 9683 (green) and 12370 (dark blue) in Figures 8 & 9) begin to lose profit due to the new aggressive pricing by station 8537 (Figure 9 (b)). Their response is to drop their prices by up to 1p per day (Figure 9 (a)) in an attempt to attract consumers back and thereby maximise profits. Over the course of the simulation, the effect of the price drop diffuses across the area triggering stations located further away from station 8537 to implement the same strategy; for example, stations 9445 (pink) and 398 (light blue) between days 5 to 25.

#### 6.10

While it is difficult to see this behaviour in the real world, as the system is more varied, it can be seen in small-scale examples (see Appendix 4). This is not entirely surprising, as at this spatial and temporal scale the industry rules which were used to construct the model can been seen directly at work.





Figure 9. Change in (a) price and (b) profit over time for the selected stations in the real data diffusion experiment with a price drop of 5 pence

At the same time as the surrounding stations are dropping their prices, station 8537 has large sales but a low profit margin because of its cheap prices, and so increases its price in an attempt to maximise profit levels. This price increase slows down around day 5 and then exhibits minor fluctuations until day 30. At this stage, the prices appear to converge at a point where all the stations are making profits again. This happens around 66-67 pence – any prices lower than this will precipitate negative profits (Figure 9 (a)). With the absence of any external influences (e.g. rise in tax or crude oil prices) to force an upward kick in the prices, the simulation exhibits little fluctuation in prices or profit from day 45 through to day 60.

## 6.12

Experimentation with price fluctuations has shown that the model is able to produce consistent results that are close to the real data. This demonstrates that the model is insensitive to moderate price adjustments. In examining the results spatially, the effects of the price drop can be seen to steadily diffuse outwards over time before reaching a steady state. By day ten, the price drop was seen to have extended to local stations; this is consistent with the conclusions made about the results in Table 3, but indicate that a longer time-series would be ideal.

## Alternative Validation

#### 7.1

The analyses in Section 5 provide validation for the model, but on the basis of price, which is used in model initiation. The analyses in Section 6 suggest the model maintains behaviours that are seen, at the small scale at least, in the real world. However, it is still necessary to confirm that the model's overall emergent behaviour matches overall behavioural trends seen in the real world. This section provides two such alternative validations by showing the model matches real pricing behaviour and can predict station profitability.

## **Rockets and Feathers**

## 7.2

Within the diffusion experiments, it was evident that the system reached a steady solution and remained within this state with minor fluctuations (see Figures 7 and 9). It was concluded that the system did not deviate from this steady state due to the absence of external influences to "shock" the system into different behaviour. Such shocks can be provided in the real system by increases in fuel duty or crude oil prices. One of the common reactions to these influences is a rapid increase in prices followed by a slow fall back to the original price after the external influence has diminished. This phenomenon is commonly referred to as the "rockets and feathers" effect (Bacon 1991). An increase and decrease in crude oil prices will be mimicked within this section and the reaction of the system examined to see whether similar behaviour is produced.

#### 7.3

One of the model parameters, *costToProduce* incorporates transportation costs, fuel duty and crude oil prices. Instigating an increase in this parameter will mimic a rise in cost in one or all of these factors. This will have the effect of reducing the profit that each station is making, hopefully precipitating rapid upwards price adjustments ("rockets") in order to maintain profits. Decreasing the *costToProduce* should allow the stations to act more competitively e.g. undercutting each other to maximise their profit levels. This will hopefully generate a set of slow decreases in price (the "feathers"). The experiment was undertaken for West Yorkshire using the real data from July 27th. The default parameter values outlined in Appendix 3 were used. After 40 days (as the model reached equilibrium) the *costToProduce* variable was increased from 66 pence to 68p. This was then decreased after 80 days from 68p to 66p.

### 7.4

Figure 10 presents the price and profit over time for the same stations used in Section 6. Figure 10 (a) shows the increase in price at the stations between day 40-60. This corresponds to the increase in the *costToProduce* parameter at day 40. At this point, the stations find themselves instantly making less profit and take action via a price increase to regain profit (Figure 10 (b)). Prices and profits rise

universally - this universal price rise is possible because the spatial interaction model has no reason to change the customer allocation - customers essentially have no reason to switch allegiances as the relative price differences between stations have not changed. However, sooner or later the prices rise to an artificially high level where profits can be improved for some relatively lowerprofit stations by dropping prices relative to surrounding competitors. At this point the prices and profits rapidly adjust to a new equilibrium suitable for the new costs (days 60 - 80).



Figure 10. Mean price (a) and profit (b) plotted against time for simulations of the "rockets and feathers" effect using the West Yorkshire data

7.5

On day 80, the *costToProduce* parameter is decreased by 2 pence to 66p. The result is generally a sharp increase in profit (Figure 9 (b)) followed by a decrease in price (days 80 - 120). This tends to start with suburban stations, which rapidly drop their prices for temporarily enlarged profits. The price drop then propagates into the urban areas. Station 9445 exhibits a rapid decrease in price, but collectively the other urban stations decrease in price more gradually than the increase at day 40. The decrease occurs in steps reflecting the manner in which the stations are changing their prices. This decrease in price is a result of the stations employing competitive measures to attract more consumers, i.e. undercutting each other. Interestingly, the rural station (398), which has a relatively captive market, reacts to the increased profits by slightly raising its prices for short-term profit gains. For all stations the profit levels slowly decrease with the price, but the price remains above the *costToProduce* value and so the stations remain in profit. The price finally settles to approximately 68p - 69p. This is roughly the same variation in price that the system showed before the increase in *costToProduce* - on day 39 the price varied between 68 - 69.5p.

#### 7.6

The work within this section has further tested the model by attempting to mimic a well documented phenomenon. Forcing an increase in the *costToProduce*does produce behaviour that mimics the rockets effect (prices increase instantaneously and continue to do so over 20 days). Decreasing the *costToProduce*produces a more gradual drop in price (starts immediately in some stations, but

takes 60 days to stabilise). This suggests that the "rocket and feathers" effect can be accounted for by the internal dynamics of the system. Additional behaviour, such as resistance to all-out competition while profit margins are good for everyone, does not need to be invoked for the real world system.

## **Station Profitability**

## 7.7

Another validation of the model can be achieved by looking at the profitability predictions that the model makes on the basis of the 1999 data used above, and comparing these with a later situation. The model calculates the profit made at each station on each day. On the basis of the amount of profit made, the action of the station is determined. A test of how well the model is predicting these profit levels would be to assess whether stations with low profit levels are still open in the real world in 2004. It is hypothesised that stations with low or negative profit levels are most likely to have been driven out of business.

### 7.8

A survey of stations within Leeds was taken in June 2004 (by simply driving around the area). This provided a list of stations that have closed since the original data was collected in 1999. These stations were overlaid on a profitability map generated by the Network-model initialised with the real data from July 27th 1999 and using the default rules in Appendix 3. Profits on day ten of the model run were plotted (Figure 11).

## 7.9

Of the 43 stations surveyed, 14 have closed since 1999. Figure 11 shows a strong correlation between sites that the model identified as not being profitable and stations that have closed. To formalise this correlation, a Wilcoxon rank sum test was performed. The null hypothesis tested was "there is no difference between the mean profits of stations that are open and closed". The Wilcoxon test gave a result of p=0.000812 indicating that there is a smaller than 0.1% chance of the null hypothesis being true. The null hypothesis can therefore be rejected. As clearly evident in Figure 11, there is a definite difference in the profit in the model between stations that are open and closed in reality.



Figure 11. Map showing the profitability of stations in Leeds using the Network-model with data from 1999. Also shown for comparison are the status (open or closed) of the stations in 2004

7.10

This analysis also allows a confident assertion to be made about supermarket stations. Figure 10 shows that these stations are also located at the high points of the profitability map. The original data pre-dates the appearance of the new supermarket stations at Kirkstall, Seacroft and near Osmondthorpe. Since the appearance of these stations a large number of their closest competitors have closed. This can be particularly seen when examining the sites surrounding the new supermarket stations at Seacroft and near Osmondthorpe. These now closed stations were identified by the model in 1999 as sustaining low profits. With the introduction of the supermarket stations with their aggressive pricing, the marginal profits of these stations would have plummeted resulting in their closure. This effect is not solely reserved for the low profit stations. At Osmondthorpe and Kirkstall, stations that were situated in relatively profitable areas have also closed. This shows that the opening of new stations in an area can significantly affect the local competition dynamics. It is also interesting to note that only petrol stations associated with supermarkets were observed to have opened since 1999, showing the shifting balance in the UK retail petrol market towards supermarkets as important petrol retailers.

The results in this section provide an alternative validation of the performance of the model. They suggest that the profit strategy adopted for the model is useful for predicting profitability as well as predicting fuel prices.

## 😌 Conclusion

#### 8.1

Within this paper, an alternative framework for modelling complex, locally interacting retail markets has been presented. Multi-agent systems provide an exciting framework within which a retail market can be accurately described. They have the potential for individual agents (in this case petrol stations) to be supplied with detailed knowledge of regional-scale markets by the attachment of more specialised models such as spatial interaction and network components. The work presented here has shown the potential for using such an approach. The outcome of local interactions can be easily seen at a global level and behavioural and statistical data models can be easily combined.

#### 8.2

Sensitivity testing and diffusion experiments showed that the model is fairly robust and responded sensibly to changes in the system. It reproduced results similar to the well documented phenomena of "rockets and feathers" and successfully predicted the status of the stations in 2004 providing additional validation of the results of the model.

#### 8.3

The success of this approach in recreating the prices and profitabilities of the retail petrol market has been demonstrated, however, the flexibility of this framework would allow for alternative applications: for example, the model could be used for planning decisions, such as locating the best site for a new petrol station, or deciding whether to close a station because it is no longer commercially viable. It could also be used by companies as a tool for experimenting with different marketing strategies. The ability of the model to predict long-term patterns as well as generalised short term variations makes this a powerful tool.

## 8.4

Even with the recent upsurge in computing power, the hybridisation of agent-based systems with traditional market models, like the spatial interaction model, provides a useful methodology. This is not simple for the reduction of computational effort and the generalisation of processes, but also because it ties agent-based systems to a well developed and mature literature of market studies. This provides a solid basis for expansion of those elements of the modeling effort that are more amenable to an agent-based approach. The model presented here utilises this combination to provide an individual-based behavioural model of retailers on a regional scale, and ties rule-based behaviour to real world price and profitability patterns.

# Scknowledgements

The work reported within this paper was funded by the Economic and Social Research Council and GMap. All digitised boundary data are UK Crown and ED-Line Copyright. The network component was developed during an exchange visit to The GeoVista Center funded by the Worldwide Universities Network. Particular thanks to the two anonymous referees who provided invaluable comments.

## 🐬 Appendix 1: Regression Modelling

### A1.1

Geographically weighted regression (GWR) is an extension to multiple linear regression which adds in a spatial element (see Brunsdon et al., 1998, for further details). Variables were input into the GWR that were identified as being important from analysis of the real data and from available literature. These were population density (a pseudo variable for geography: for example high population densities indicate urban areas and low population densities rural regions) and distance to neighbours (i.e. local competition). Supermarkets, motorways and the Esso brand were incorporated as they were identified as operating distinct pricing rules. Supermarkets were on average 1-2p cheaper than their competitors, motorway stations were 1-2p more expensive, and Esso operates its Price Watch Policy (they will match any competitors' price within 3 miles).

### A1.2

Variables not included were cost of crude oil, fuel tax and duty, and exchange rates. Each of these factors has a temporal impact on the model, for example, an increase in crude oil prices has a slight time lag before impacting on the pump price, however they do not change the relative prices between stations on a daily basis. The analysis with the GWR was aimed at examining how well the regression model would capture the trends and patterns on a particular day.

#### A1.3

The results of the GWR had an  $r^2$  0.323 (with a standard error of 0.903). This was an improvement on the results from standard linear multiple regression modelling (an  $r^2$  of 0.253), but the results were still poor (see <u>Heppenstall 2004</u>). This suggested that alternative methods for modelling this system were required.

## Appendix 2: Definition of Parameters Used

## Table A1: Parameters

Parameter	Definition
В	Coefficient controlling the impact of distance.
λ	Coefficient controlling the impact of price.
fixedCosts (£)	The amount the petrol station has to pay per day to keep running.
costToProduce (p)	e The amount per litre that it costs the station to produce and sell the petrol.
changeInProfi	The level of profit under which the station will not change its strategy.

## (p)

*Overprice (p)* The amount by which petrol stations can be more expensive than their neighbours without becoming uncompetitive and being forced to cut their price.

*Undercut (p)* The amount by which petrol stations can undercut competitors, e.g. 1p, 2p. *Neighbourhood* The distance that the petrol stations will treat as their neighbourhood e.g. 2km, 3km. *(km)* 

## Appendix 3: Experiment Setup

### A3.1

For each simulation, all the petrol stations in West Yorkshire were input. Unless otherwise stated, each station was set to 71p (the mean price of the data set) or initiated with data from day zero. The model was run with the following parameter values for ten days or equilibrium.

Table A2:	Experiment	parameter	values
-----------	------------	-----------	--------

Parameter	Value
В	0.0003
λ	0.7
fixedCosts (£)	80.00
costToProduce (p)	66.00
changeInProfit (p)	40.00
Overprice (p)	5
Undercut (p)	1
Neighbourhood (km)	5.0

#### A3.2

These parameter values were determined on the basis of limited industry information, for example the Competition Commission Report (<u>1990</u>) and research undertaken by Ning and Haining (<u>2003</u>) as well as detailed analysis performed on the real data. The actual parameter values were validated by the use of a genetic algorithm (see <u>Heppenstall et al. 2006</u>).

## Appendix 4: Small Scale Example of Station Behaviour



#### A4.1

The above figure shows the reaction of several stations to a price change in their neighbour over 3 days. On July 27th, station 1 and 3 are more expensive than their mutual neighbour, station 2. On July 28th, station 2 has reacted to this price difference by increasing its price to 73.9p, higher than station 1 or 3. Station 3, however has reacted to the lower price set by station 4 and dropped its price. By July 29th, station 1 and 2 have dropped their prices, but station 3 has increased again. The behaviour exhibited by these stations is interesting. They are all within a few kilometres of each other and clearly aware of their neighbours' prices. The varying prices seen each day could be a result of a time lag, for example, station 2 increases then decreases its price in possible reaction to station 3 (which appears to be reacting in the same manner to station 2). This example of small scale behaviour also shows that some stations can increase and decrease their price by several pence per day to remain competitive, though this tends to be rare.



ABDULAI, A. (2000) Spatial price transmission and asymmetry in the Ghanaian maize market. *Journal of Development Economics*, 63 (2): 327 - 349.

BACON, R.W. (1991). Rockets and Feathers: the Asymmetrical Speed of Adjustment of UK Retail Gasoline Prices to Cost Changes. *Energy Econ*, 1:211 – 218.

BIRKIN, M., CLARKE, G., CLARKE, M. AND WILSON A. (1996). Intelligent GIS: Location decisions and strategic planning. GeoInformation International

BIRKIN, M., CLARKE, G. AND CLARKE, M. P. (2002). Retail Geography and Intelligent Network Planning. Wiley, 1st edn.

BRUNSDON, C., FOTHERINGHAM, A. S. AND CHARLTON, M. E. (1998a). Geographically Weighted Regression - Modelling Spatial Non-Stationarity. *Journal of the Royal Statistical Society - Series D - The Statistician*, 47(3): pp431 - 443.

BRUNSDON, C., FOTHERINGHAM, A. S. AND CHARLTON, M. E. (1998b). Spatial Non-Stationarity and Autoregressive Models. *Environment and Planning A*, 30(6): pp957-993.

BUNN, D. W., and OLIVERIA, F. S. (2001). Agent-based simulation: An application to the New Electricity Trading Arrangements of England and Wales. IEEE Transactions on *Evolutionary Computation* (5): pp493-503.

CASTIGLIONE, F. (2000) Diffusion and aggregation in an agent based model of stock market fluctuations. *International Journal of Modern Physics* C 11 (5): 865 - 879.

COMPETITION COMMISSION (1990). *Petrol Pricing*. URL <u>http://www.competition-</u> commission.org.uk/rep\_pub/reports/1990/fulltext/265c4.pdf.

CONSTANTINIDES, E. (2002). Influencing the online consumer's behaviour: the Web Experience. *Internet Research-Electronic Networking Applications and Policy* 14(2) pp 111 - 126.

EMCAS, Electricity Market Complex Adaptive Systems. Argonne National Laboratory. http://www.dis.anl.gov/ceeesa/programs/emcas.html. Accessed May 2006.

GALOETTI, M., LANZA, A. AND MANERA, M. (2003) "Rockets and Feathers Revisited: An International Comparison on European Gasoline Markets". Energy Economics, 25 pp175-190.

GARCIA-FLORES, R. AND WANG, X.Z. AND GOLTZ, G.E. (2000). Agent-based Information Flow for Process Industries' Supply Chain Modelling. *Computers and Chemical Engineering* 24(2-7) pp1135 - 1141.

HA, S.H., BAE, S.M. and PARK, S.C. (2002). Customer's Time-Variant Purchase Behaviour and Corresponding Marketing Strategies: An Online Retailer's Case. *Computers and Industrial Engineering*. 43(4) pp801 - 820.

HEPPENSTALL, A.J. (2004) Application of Hybrid Intelligent Agents to Modelling a Dynamic, Locally Interacting Retail Market. Unpublished Thesis. School of Geography, University of Leeds.

HEPPENSTALL, A.J., EVANS, A.J. AND BIRKIN, M.H. (2005). A Hybrid Multi-Agent/Spatial Interaction Model System for Petrol Price Setting. *Transactions in GIS* 9(1) pp 35 - 51.

HEPPENSTALL, A.J., EVANS, A.J. AND BIRKIN, M.H. (2006), Genetic Algorithm Optimisation of a Multi-Agent System for Simulating a Retail Market. In press, *Environment and Planning B.* 

JULKA, N. AND SRINIVASEN, R. and KARIMI, I. (2002). Agent-based supply chain Management-1: Framework. *Computers and Chemical Engineering*. 26(12) pp 1755-1769.

KIRMAN, A., AND VRIEND, N. J. (2001). Evolving Market Structure: An ACE Model of Price Dispersion and Loyalty. Journal of Economic Dynamics and Control 25, pp 459-502.

LOMBARDO, S., PETRI, M., and ZOTTA, D. (2004). Intelligent Gis and retail location dynamics: A multi agent system integrated with ArcGis. *Computational Science And Its Applications - Iccsa 2004, Pt 2 Lecture Notes In Computer Science,* 3044: 1046 - 1056.

MANNING, D. N. (1991). Petrol Prices, Oil Price Rises and Oil Price Falls: Evidence for the UK since 1972. Applied Economics, 23:1535 - 1541.

MITCHELL, J. D., Ong, L. L. and Izan, H. Y. (2000). Idiosyncrasies in Australian Petrol Price Behaviour: Evidence of Seasonalities. *Energy Economics*, 28:243 – 258.

NING, X. AND HAINING, R. (2003). Spatial Pricing in Interdependent Markets: A Case Study of Petrol Retailing in Sheffield. *Environment* and *Planning A*, 35:2131 - 2159.

O'SULLIVAN, D. AND HAKLAY, D. (2000). Agent-based Models and Individualism: Is the World Agent-Based? *Environment and Planning* A, 32(8) pp1409 - 1425.

PETTERSEN, E., PHILPOTT, A.B. AND WALLACE, S.W. (2005). An Electricity Market Game Between Consumers, Retailers and Network Operators. *Decision Support Systems*, 40(3-4) pp427 - 438.

PLUMMER, P. S., SHEPPARD, E. AND HAINING, R. P. (1998). Modeling Spatial Price Competition: Marxian versus Neoclassical Approaches. *Annals of the American Association of Geographers*, 88(4): pp575 - 594.

PONTIUS, R.G. Jr, HUFFAKER, D., and Denman, K. (2004). Useful Techniques of Validation for Spatially Explicit Land-Change Models. *Ecological Modelling*, 179: 445 - 461.

REILLY, B. AND WITT, R. (1998), "Petrol Price Asymmetries Revisited", Energy Economics, 20 pp297 - 308.

SHIN, D. (1994). Do Petrol Prices Respond Symmetrically to changes in crude oil prices? OPEC Review, 18:137 - 157.

SIGNORILE, R. (2002). Simulation of a Multi-Agent System for Retail Inventory Control: A Case-Study. Simulation-Transactions of the Society for Modelling and Simulation International 78 (5) pp 304 - 11.

Return to Contents of this issue

© Copyright Journal of Artificial Societies and Social Simulation, [2006]

