Proceedings of the 2009 INFORMS Simulation Society Research Workshop
L.H. Lee, M.E. Kuhl, J.W. Fowler, and S. Robinson, eds.

DETERMINING THE RANGE OF PREDICTIONS FOR CALIBRATED AGENT-BASED SIMULATION MODELS

DongFang Shi
Roger J. Brooks

Department of Management Science
Lancaster University Management School
Lancaster. LA1 4YX. U.K.

ABSTRACT

Agent-based simulation is increasingly used to study systems in many areas of business and science. Using agent-based simulation for prediction could be very valuable. However, these models usually have many parameters which are difficult to measure directly leading to uncertainty as to the best values to use. It is very important to consider the alternative feasible sets of parameter values, otherwise the modeling results could be misleading. A method is described to take account of this when making predictions and was applied to an agent-based word of mouth consumer model. The result was a wide prediction range indicating the limitations of using such a model for predictions. The method could also be applied in other situations.

1 INTRODUCTION

In recent years agent-based (or individual-based) simulation has received a lot of attention. Agent-based simulation refers to simulations of systems that contain agent entities whose behavior depends dynamically on the state of the system. This enables the agents to adapt their behavior to changing conditions.

There is no standard definition of an agent. Some definitions list a set of properties but a better approach is perhaps simply to say that an agent is an entity for which some cognitive process is modeled (Edmonds and Mohring 2005). The internal rules that represent the cognitive decision process can vary from a simple function of the inputs received to very complex rules incorporating various internal state parameters including a representation of the agent's worldview of some part of the environment. In some cases the rules can change during the simulation to represent learning. The number of agents modeled can also vary from an individual agent through to a large population. Populations are usually heterogeneous with individual agents having different parameters or even quite different rules (e.g. different trading strategies in a stock market simulation). Interactions between the agents are usually a key part of the behavior of the system. A very wide variety of applications have been studied using agent-based simulation including stock markets, auctions, the spread of disease, ecosystems, military battles, crowd dynamics, sports games, transport, social behavior, social networks, the development of technology, and consumer market behavior (such as fads).

Elements of cognitive processes have often been included in discrete event simulation (such as the decisions of customers as to which queue they join) and therefore agent-based simulation is not a completely new type of modeling. Novelty in agent-based simulations can lie in the complexity of the rules of behavior, in the size of the population, or in the application of simulation to new situations. One of the causes of the greater use of agent-based simulation is that increasing computing power now makes such simulations feasible. There is also an appreciation that for some systems an agent-based approach may be necessary in order to capture the dynamics of the system.

Much of the agent-based simulation work has had the aim of increasing the understanding of the type of system rather than trying to reproduce a specific situation. Such an approach can potentially be valuable in producing important new insights and improving understanding. For some applications, using agent-based simulation for prediction (rather than just better understanding) could be very powerful. For example, a company might wish to use a model of the population of their customers with word of mouth interactions to predict the sales of their product or the effect of an advertising campaign. However, the problem is that agent-based models typically have a very large number of parameters and many of these cannot be measured directly or estimated with sufficient precision. The only other information available may be historical output data from the real system. Such data can be used to calibrate the model by finding parameter values that produce a good fit with the data. This is an inverse problem since it consists of using the outputs to determine the inputs. The problem is that there will usually be many solutions because there are many parameters and few historical data values and because any model that produces a good fit should be considered acceptable. A perfect fit is not expected because any simulation is a simplification of the real system and also there may be measurement errors in the historical data. The result is that a wide range of sets of parameter values may give an acceptable fit and are therefore feasible values. However, they may give quite different predictions.

The inverse problem has been studied in other areas of science including groundwater modeling (Yeh 1986). Brooks et al. (1994) set out an approach to find the range of predictions of a groundwater model from alternative feasible calibrations that was applied to a model of the Birmingham aquifer in the U.K.. The method involves setting a fitness measure, a critical value of the fitness measure that defines an acceptable fit, and a prediction measure. Searches are then made across the parameter space that aim to find the highest and lowest predictions for those parameter values giving an acceptable fit. The results for the Birmingham aquifer model were quite a wide range of predictions and since this was a risk assessment exercise the most pessimistic results were likely to be the most relevant. The groundwater model was a deterministic model and so the method needs to be adapted to be applied to stochastic agent-based models.

The aims of the study described here were to develop a plausible agent-based model of word of mouth marketing, to extend the Brooks et al. (1994) methodology to stochastic models and to apply the methodology to the agent-based model to investigate the seriousness of the calibration problem in this case.

2 METHODOLOGY

The approach used in the study was to develop an agent-based model and to treat this model as the real system. Output data from this model was then taken as measured values from the real world and, in a pseudo-modeling exercise, used to calibrate an agent-based model of the system. The advantage of such a pseudo-modeling exercise is that the "real system" is completely known. Consequently, the models' predictions can be compared with the "true" future values, and the precise differences between the model and the real system are also known. In this case, the same model design was used for the "real system" and the model. In other words, the model structure here exactly matches that of the real system. This is clearly an idealized situation since in a real study the simulation model would be a simplification of the real world and the structure of the real world would not be fully known. However, the advantage of this approach is that the uncertainty in the predictions is then entirely due to the nature of the calibration process rather than any differences between the way the model and the real system work. This allows the study to isolate this effect which would not be possible if an actual real world system was used.

3 SIMULATION MODEL DESIGN

The application chosen for the research was a consumer word of mouth model. The type of product we had in mind when constructing the model was one with a fairly short life cycle, with a high likelihood of the passing of information and opinions between consumers by word of mouth, and that is purchased as a one-off item (rather than a repeated purchase). Examples would include a computer game, a music album or a cinema ticket for a particular film. The population represented might be school or university students.

Other studies that have investigated word of mouth consumer behavior include Baxter et al. (2003), who investigated the impacts of customer relationship management (CRM) strategies, and Kijima and Hirata (2004) who looked at the effect of different network structures. The agents in the Baxter et al. (2003) model have perception values for the price and quality of the product, although details of the way that perceptions were allocated and how they changed during the model run are not provided in the paper. In the Kijima and Hirata (2004) model the agents have an enthusiasm parameter (which equals the probability of making a purchase) being a weighted sum of various attributes of the product adjusted by the reliability of the information obtained from other agents and the attitude to risk of the agent. Agents are only able to pass information to other agents for a limited period after finding out about the product, based on the SIR (susceptible / infected / removed) approach used in modeling disease transmission.

The literature therefore provides very limited information and no consensus as to the best way to model a consumer market with word of mouth interactions. Instead, the attributes of the agents and the interactions between agents in our model are based on our subjective views of the important factors in the real world, whilst trying to keep the model structure as simple as possible and the number of variables as small as possible.

In our model, each agent has state variables for both knowledge and preference, whose values change during the simulation. As explained later, preference is partly a function of knowledge. The state variables are:

- Knowledge (K): How much information the agent knows about the product (value between 0 and 100).
- Preference (P): How much the agent likes or dislikes the product (value between -100 and 100 with positive values indicating like).

Therefore, the state of the model at any point in time during the simulation run is the K and P values of all the agents. The agents also have three fixed attributes assigned at the start of the simulation, with the values selected at random for each agent from probability distributions (normal distribution denoted by normal(mean, standard deviation):

- Influence (*I*): Represents the agent's social standing within the population (normal(10,3) restricted to the interval [0, 20]).
- Unbiased true preference (*U*): Represents the preference that the agent would have about the product with complete knowledge in the absence of peer pressure (*U*+: 90% probability of selection from normal(75,15) restricted to [35, 95], *U*-: 10% probability of a negative *U* selected from normal(-75,15), restricted to [-95, -35]).
- Buying criterion (*B*): The preference value at which the agent buys the product. This represents different attitudes regarding purchasing behavior from cautious to free spending (normal(65,10), restricted to the interval [0, 100]).

The general willingness to adopt new products is represented by the *B* parameter, whereas *U* is the inherent attractiveness of this particular product to the agent, and *I* enables agents to have different levels of influence. The agent buys the product when its *P* value is greater than or equal to its *B* value. In modeling knowledge (*K*) and preference (both *P* and *U*) it was not considered necessary to divide up these variables according to separate attributes of the product.

The population of agents simulated was 500 and agents were allocated to social groups of between 2 and 8 members representing close friends. The probability of talking to each other member of the group each day, *T*, was set at 0.1. Agents also lose a percentage of their knowledge each day with the value of this parameter, *L*, being set at 1%. There were a total of 5 interactions between agents each day outside the social groups.

The interactions simulated in the model are conversations between agents about the product, interactions between agents and the environment (with independent sources and company sources treated separately) regarding the product (representing seeing the product in the shops, seeing adverts or reading media articles), and agents buying the product. The model also has fixed parameters for the knowledge, preference and influence of the company and independent sources of information as well as probabilities for the agents receiving information from these sources. The initial probability of each agent receiving information from the environment is a random value from the uniform(0,E) distribution (with the value of E set at 0.25), with 80% of these messages being from the company and 20% from independent sources.

For simplicity, it is assumed in the model that only pairwise interactions between two parties take place. The model structure is based on the assumption that exchanges regarding the product are a combination of raw information (i.e. facts about the product) and an influence on preference through peer pressure.

Fixed equations are used for the interactions (both interactions between agents and between agents and the environment), and the forms of the equations for agent *a* interacting with agent / environment *b* are as follows:

Change in knowledge due to the interaction:

$$K_a^{new} = K_a^{old} + \alpha\left(100 - K_a^{old}\right)\frac{K_b}{100} \quad (1)$$

Effect on preference of the change in knowledge:

$$P_a^{new} = P_a^{old} + U_a \frac{\left(K_a^{new} - K_a^{old}\right)}{100} f_U \quad (2)$$

Effect on preference of peer pressure:

$$P_a^{new} = P_a^{old} + \alpha\left(P_b - P_a^{old}\right) f_K f_I \quad (3)$$

where α is a uniform(0.01, 0.15) random variable, and

$$f_U = \frac{20}{20 + \left|P_a^{old} - U_a K_a^{old}/100\right|}$$

$$f_K = \frac{\left(100 + K_b - K_a\right)}{200}$$

$$f_I = \frac{\left(10 + I_b - I_a\right)}{20}$$

Increase in knowledge when product is bought:

$$K_a^{new} = 0.5\left(K_a^{old} + 100\right) \quad (4)$$

In each of these interactions the agents may increase their knowledge *K* value, representing gaining information about the product (Equation 1). The change in knowledge depends on the existing knowledge of both parties. The gain in knowledge is the proportion of knowledge not known by the agent multiplied by the knowledge of the other party multiplied by α. For example if the agents knowledge is 70 then they will gain 0.3α of the other party's knowledge. This reflects that even if the other party knows less than the agent they will still probably have some different knowledge.

Whenever an agent's *K* value changes (whether gaining or losing knowledge) this changes its *P* value as a function of the *U* value (Equation 2). The underlying assumption is that, in the absence of peer pressure, $P = U \times K/100$ (i.e. a linear relationship between *P* and *K*). Therefore, an increase in knowledge increases the preference by a proportion of the *U* value (with an adjustment to take account of existing peer pressure).

The preference *P* value will also change due to the influence of the preference of the other agent (peer pressure) or the environment (e.g. the opinion in a magazine review) (Equation 3). The strength of both of these interactions depends on the relative knowledge and influence (the *I* values) of the two parties.

The agent buys the product when its preference *P* value reaches its buying criterion *B* value. This increases the knowledge of the product by half the current lack of knowledge (Equation 4). As is the case with any change in knowledge, Equation 2 is then used to change the agent's preference.

At the start of the simulation all the agents have no knowledge and no preference about the product since it is a new product. However, the company conducts an initial marketing campaign and the agents may also see the product in the shops or read about it in the media. These interactions enable the agents to gain knowledge and change their preference in the initial stages of the simulation. The limited time of the campaign was modeled by the probability of receiving outside information being reduced linearly down to 25% of the initial value for each agent over a period of 75 days.

The above values for the parameters were set so that the model gave plausible results as the "real system". The model was run 1000 times and this was assumed to represent the total population (e.g. 1000 schools). Figure 1 shows the product life cycle of average sales per day for the 1000 runs. The model behavior was also investigated through sensitivity analysis on the parameter values.

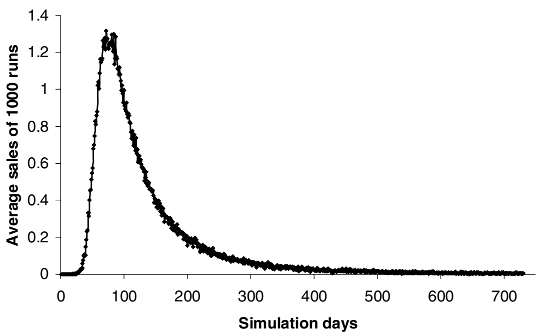

Figure 1: Product life cycle for "real system"

4 CALIBRATION AND PREDICTION

The scenario investigated in the research was that the current time is several weeks after the product was launched. Data for the total sales to date have been collected and provide the best available data to use for calibration. The company wishes to use an agent-based model to predict the total sales of the product and so the model should match the sales to date and needs to forecast the final sales. The sales values are expressed throughout as the mean number of items sold per population of 500 agents. The real system was assumed to consist of 1000 such populations and so the "observed" values were obtained by an average of 1000 runs of the "real system" model. This gave a true final sales value of 124 (i.e. 24.8% of the population).

As explained in section 2, the model is the same as the model used for the real system. However, six of the parameters were assumed to be unknown and unmeasurable and so had to be determined through calibration. These parameters were the mean values of the normal distributions for *U*+, *U*-, and *B*, the upper limit of the uniform distribution for receiving information from the environment *E*, and the fixed values for *T* and *L* (section 3). All other parameter values were the same as for the "real system" model.

A calibration method similar to that of Brooks et al. (1994) was used (section 1). This required setting a fitness measure and a critical value that defines an acceptable fit. In this case a single data value is available (total sales to date) and the measure needs to take account of the stochastic nature of the model. The choice of fitness measure is subjective and needs to reflect the desired accuracy of the model. The measure chosen here was the difference between the 95% confidence interval from 100 replications of the model for average sales to date per population and the actual value, with an acceptable fit being that the distance is 0 (i.e. the actual value lies within the interval). Using a large number of replications makes this quite a strict measure since the confidence interval is likely to be quite narrow. A fitness function *F* was defined to implement this which takes the value of 0 if the true value is in the interval and 10 + the absolute difference between the true value and the interval if it is not in the interval. The prediction measure is simply total sales per population.

The aim of the calibration process is to find the highest and lowest total sales amongst the parameter values that meet the fitness criterion. There is no method that guarantees to find a global optimum for a complicated function. Instead, the approach followed was to run the model for a grid of three values for the six parameters (i.e., $3^6 = 729$ points in total). 100 replications were done for each point and the average sales for the initial period and the average total sales calculated. Then several local searches were carried out from different starting points on the grid and the extreme values from the local searches gave the prediction range. The starting points for the local searches were the points with the highest sales, lowest sales, three highest sales with fitness 0, three lowest sales with fitness 0 and the default parameters. The Nelder-Mead Downhill Simplex (Nelder and Mead 1965) was used as the local search method with an objective function to minimize $e^F + S$, or $e^F - S$ where *S* is the total sales. The local searches only used 10 replications so as to reduce the run time required (which was still considerable even on a high performance cluster). However, once the search appeared to have converged, the search was continued with 100 replications until a point was found that met the fitness criterion. The final prediction range was the extreme values from the 95% confidence intervals for total sales from the 100 replications.

The calibration method was applied to four scenarios that differed only by the length of the initial period with the lengths being 70, 105, 140 and 175 days. The total sales to date used for calibration for these four initial periods were 24.4, 63.8, 86.5 and 99.0 respectively. For each scenario a contour map of *F* was plotted by combining *U*+, *U*- and *B* into one variable and *E*, *T* and *L* into another variable. These appeared to indicate that in each case there are several disjoint regions of acceptable fit ($F = 0$) in quite different regions of the parameter space. The results of applying the calibration method were very wide prediction ranges for all four scenarios, being [58, 376], [79, 319], [91, 277], [109, 187] respectively, and these are shown in Figure 2. Therefore, even though the model structure was correct, only six parameters were treated as unknown and a fairly strict fitness function was used, the prediction range is so wide for each scenario that it would provide very little useful information. This is due to the uncertainty of the values of the parameters obtained by calibration.

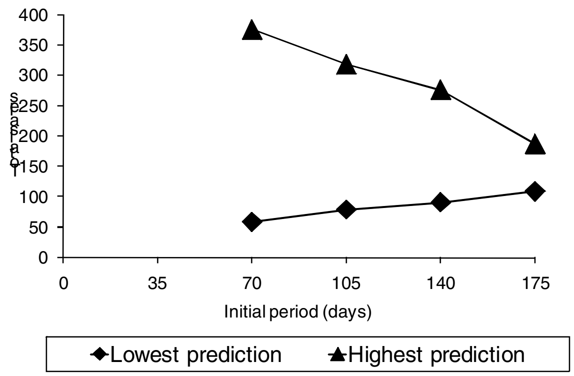

Figure 2: Product life cycle for “real system”

5 CONCLUSIONS

When some of the parameter values of a simulation model are obtained by calibration it is possible for quite different sets of parameter values to provide an acceptable model fit. However, the predictions of these different sets may vary considerably. It is therefore important not to use a single set of parameter values as the predictions could be inaccurate and misleading. A methodology for obtaining a prediction range in this situation has been set out and applied to a agent-based word of mouth consumer simulation model. The result was a very wide prediction range.

Future experiments could examine many other scenarios. This could include alternative fitness measures such as using weekly sales to date rather than total sales (although this may not make much difference as the values would be correlated). It could also include scenarios where the model structure is different to the “real system” model. The methodology could also be applied to modeling studies of an actual real system

The issue examined here applies to any simulation model in which the parameter values cannot be measured directly and have to be determined by calibration. However, it is particularly relevant for agent-based simulations because in many cases they model aspects of human characteristics and behavior on which it is very difficult to obtain data. This may therefore limit the usefulness of such models for prediction.

ACKNOWLEDGMENTS

Many thanks to Richard Meek of the Marketing Department at Lancaster University for helpful discussions on consumer behavior. This paper is partly based on Shi, D. and R. J. Brooks 2007. The range of predictions for calibrated agent-based simulation models. *Proceedings of the 2007 Winter Simulation Conference (eds. S. G. Henderson et al.)*: 1198-1206.

REFERENCES

Baxter, N., D. Collings and I. Adjali. 2003. Agent-Based Modelling – Intelligent Customer Relationship Management. *BT Technology Journal* 21(2): 126-132.

Brooks R. J, D. N. Lerner, and A. M. Tobias. 1994. Determining the Range of Predictions of a Groundwater Model which arise from Alternative Calibrations. *Water Resources Research* 30(11): 2993 – 3000.

Edmonds, B and M. Mohring. 2005. Agent-based simulation modeling in social and organizational domains. *Simulation – Transactions of the Society For Modeling And Simulation International* 81(3): 173-174.

Kijima, K. and H. Hirata. 2004. Diffusion of Word-of-Mouth in Segmented Society: Agent-Based Simulation Approach. In: *Artificial Intelligence and Simulation (ed. T. G. Kim), selected papers from 13th International Conference on AI, Simulation, and Planning in High Autonomy Systems*, Lecture Notes in Computer Science (Subseries: Lecture Notes in Artificial Intelligence), 3397: 198-206.

Nelder, J.A. and R. Mead 1965. A simplex method for function minimisation. *Computer Journal*. 7: 308-313.

Yeh, W. W.-G. 1986. Review of Parameter Identification Procedures in Groundwater Hydrology: The Inverse Problem. *Water Resources Review* 22(2): 95-108.

AUTHOR BIOGRAPHIES

DONGFANG SHI Received a BSc (Hons) Economics from Beijing Normal University, China in 2000, and a MSc Management from Lancaster University in 2002. She is currently a PhD student at Lancaster University.

ROGER J. BROOKS is a lecturer in the Management Science department at Lancaster University. He received a PhD and MSc in Operational Research from Birmingham University and a B.A. (Hons) degree in mathematics from Oxford University. He is co-author of a textbook on simulation. His areas of interest include conceptual modeling, Boolean networks and agent-based simulation. His website can be found via <`www.lums.lancs.ac.uk`> and his e-mail is <`roger.brooks@lancaster.ac.uk`>