Interaction between scale and scheduling choices in simulations of spatial agents

Tyler R. Bonnell, Colin A. Chapman & Raja Sengupta


To link to this article: http://dx.doi.org/10.1080/13658816.2016.1158822

Published online: 21 Mar 2016.

Article views: 147

View related articles

View Crossmark data

Citing articles: 1 View citing articles
Interaction between scale and scheduling choices in simulations of spatial agents

Tyler R. Bonnell\textsuperscript{a,b}, Colin A. Chapman\textsuperscript{c,d} and Raja Sengupta\textsuperscript{a}

\textsuperscript{a}Department of Geography, McGill University, Montreal, QC, Canada; \textsuperscript{b}Department of Psychology, University of Lethbridge, Lethbridge, Canada; \textsuperscript{c}Department of Anthropology & McGill School of Environment, McGill University, Montreal, QC, Canada; \textsuperscript{d}Wildlife Conservation Society, Bronx, NY, USA

ABSTRACT
Spatial simulations are a valuable tool in understanding dynamic spatial processes. In developing these simulations, it is often required to make decisions about how to represent features in the environment and how events unfold in time. These spatial and temporal choices have been shown to significantly alter model outcomes, yet their interaction is less well understood. In this paper, we make use of a simple group foraging model and systematically vary how features are represented (cell size of the landscape) as well as how events unfold in time (order in which foragers take action) to better understand their interaction. Our results show similar nonlinear responses to changes in spatial representation found in the literature, and an effect of the order in which agents were processed. There was also a clear interaction between how features are represented and how events unfold in time, where, under certain environmental representations results were found to be more sensitive to the order in which individuals were processed. Furthermore, the effects of feature representation, scheduling of agents, and their interaction were all found to be influenced by the heterogeneity of the spatial surface (food), suggesting that the statistical properties of the underlying spatial variable will additionally play a role. We suggest that navigating these interactions can be facilitated through a better understanding of how these choices affect the decision landscape(s) on which agents operate. Specifically, how changes to representation affect aggregation and resolution of the decision surface, and thereby the degree to which agents interact directly or indirectly. We suggest that the challenges of dealing with spatial representation, scheduling, and their interaction, while building models could also present an opportunity. As explicitly including alternate representations and scheduling choices during model selection can aid in identifying optimal agent–environment representations. Potentially leading to improved insights into the relationships between spatial processes and the environments in which they occur.

1. Introduction

Many problems of interest are inherently spatial, such as: how a parasite spreads through a population of hosts (Kramer-Schadt \textit{et al.} 2009, Linard \textit{et al.} 2009, Bonnell...
et al. 2010), how urban environments develop over time (Barredo et al. 2003, Batty 2005), or collective movement of wildlife/human groups (Turner and Penn 2002, Bennett and Tang 2006, Sellers et al. 2007). These types of problems are often characterized by disaggregated elements interacting with their local surrounding (Epstein 2006), and can show complex adaptive behaviours resulting in significant challenges for management (Grimm et al. 2005, An 2012, Parrott et al. 2012, Parrott and Meyer 2012). One approach at developing insight into the operation of these spatially-disaggregated systems is a simulation approach, where the system is described by spatially distributed agents representing individuals or processes (Railsback and Grimm 2011). In developing spatially explicit simulations, it is often necessary to make choices about how to represent spatial features in the environment (e.g., forest patches represented as 10-, 20-, or 1000-m grid cells) and how to represent feature’s behaviour in time (e.g., schedule each forest patch to update every month in a random order, or update all forest patches simultaneously each day). These choices generally reflect the scale of questions being asked, though the specific choices can be arbitrary, and there is evidence that these choices can have uncertain effects on results (Grimm and Railsback 2005, O’sullivan and Perry 2013).

The implication of altering environmental representation in spatial simulations has received some attention. Zonal issues, such as the effects of using hexagon and square grids in simulations, has identified the importance of cell shape in studies looking into connectivity or questions of mobility (Birch et al. 2007). Similarly, scale choices such as patch size have been shown to alter the outcomes of land cover change simulations (Evans and Kelley 2004). It is not likely that scaling and zoning choices will affect spatial simulations in the same way in every model. In the case of simulations using dynamic agents (Sengupta and Sieber 2007), scale and zonal choices will ultimately interact with an agent’s specific behavioural algorithm, suggesting that the focus should be on how algorithm–environment interactions drive alterations in model outcomes. Moreover, the interaction between the chosen landscape representation and agent behavioural algorithms are mediated though an event scheduler controlling how time is represented in the model and how agents are processed, e.g., agents are scheduled for processing one at a time (asynchronous) or all at once (synchronous), the choice of which has been shown to have a significant impact on model results (Page 1997, Cornforth et al. 2005, Caron-Lormier et al. 2008, Miller et al. 2008, Fatès and Chevrier 2010). Yet the interaction between spatial representation and scheduling choices has not received much attention.

To quantify the interactions between environmental representation and agent scheduling choices we make use of a group movement model, which employs a discreet representation of both time and space. Individual agents in this model make decisions based on simple foraging and safety goals, resulting in group-level movement. This model was chosen as agents both interact with the landscape (food) and thus indirectly with each other, as well as directly with each other due to safety goals. This incorporates indirect and direct agent interactions that are typical of socio-environmental models (Le et al. 2012, Filatova et al. 2013), which are expected to be influenced by both environmental representation as well as scheduling choices, due to their representation of environmental and social factors. To this model, we alter: (1) spatial representation by varying the cell size of the food landscape, (2) scheduling of agent behaviour by varying
agent processing from one at a time to all at once, and (3) overall heterogeneity of the food landscape. We then record the effects of these scale and scheduling choices on simulation outcomes.

2. Model

A model of group foraging was developed in java using Repast Simphony 2.0, and was verified by means of stepping through the code in a debug process line by line, graphical verification, and tracing the behaviour of individuals through time (Xiang et al. 2005, Crooks et al. 2008). In this model, all agents followed a simple adaptive strategy: vary their preference for nearby neighbours based on foraging success (see supplementary data for ODD description, (Grimm et al. 2010)). This strategy allowed agents to balance a simple trade-off between: (1) increased safety benefits from nearby group members against (2) increased foraging competition from nearby group members. Agents accomplished this by altering their ideal number of nearby group members ($X_{\text{ideal}}$) based on their feeding success, where $X_i$ is the number of group members within a safe radius $D_{\text{safe}}$. By increasing $X_{\text{ideal}} + X_{\Delta}$ when feeding is successful, the agent can maximize safety benefits, whereas when feeding is not successful, the agent can minimize foraging competition by lowering $X_{\text{ideal}} - X_{\Delta}$. Feeding success is determined by whether or not a hungry agent (energy levels, $E$, below $E_{\text{ideal}}$) can successfully remove $E_{\text{bitemax}}$ energy from its current resource cell (i.e., does $E_{\text{bite}} = E_{\text{bitemax}}$?). Thus, when $X_{\text{ideal}}$ is high, the individual attempts to maintain proximity to many group members, reducing the foraging options available to it. Whereas, as $X_{\text{ideal}}$ decreases, the agent is less constrained by the location of others within its group, resulting in greater foraging options.

Agents are represented as points in continuous space, and forage on a resource grid following a simple foraging algorithm (Figure 1). During each model time step, representing 30 min, agents decide to move towards food, safety, or not to move. When an agent decides to move towards safety, it moved to the center of the nearest $X_{\text{ideal}}$ group member, where distances are measured as straight line Euclidean distances. When the agent decides to move towards food, it chose the best food site within $D_{\text{search}}$ radius and moved to the cell closest to the food site that still meets its safety requirements $X_{\text{ideal}}$. If safety requirements are met at the best site, the agent moved directly to the food site (Figure 1 and Table 1).

In this model, the main form of communication between agents is their spatial positioning. To manage this communication (their locations in time) and take advantage of a parallelized approach, we built our model schedule as a series of separate agent operations. In our group’s foraging model, this approach translated to a two-step process: agents acquire information (sensory input), and agents perform a behavioural response (action) based on this information. This allowed us to run model events in series to ensure efficient communication between agents and take advantage of running agents in parallel. We implemented this scheduling design, varying the number of agents processed synchronously ($n$), allowing us to go from a purely asynchronous to fully synchronous scheduling (Figure 2). For example, in a simulation of three agents, an $n$ value of 1, would result in one agent performing a sensory update then perform an action, followed by the next agent, then the last agent. Thus, a model with an $n$ of 1
would be fully asynchronous. An $n$ value of 2 would allow 2 agents to perform a sensory update at the same time, followed by an action at the same time. The remaining agent would then perform a sensory update, followed by an action. An $n$ value of 3 in this example would then be considered as fully synchronous, as all agents update sensory inputs (snapshot of current conditions) at the same time, followed by performing an action at the same time.

The resource grid on which the agents foraged was initialized with food values taken from a lognormal distribution (location = 100, shape = $\theta$ parameters), and standardized to have the same resource total for each model run. Environmental heterogeneity was then altered in the model by varying the shape parameter from the lognormal probability distribution (i.e., $\theta$). By varying the shape parameter of this distribution, we were able to alter the landscape from homogenous to heterogeneous (Figure 3).

**Table 1.** Values and descriptions of parameters used in the group foraging model. Values defined as variable vary by individual during the simulation.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_t$</td>
<td>Variable</td>
<td>Number of group mates within safe radius</td>
</tr>
<tr>
<td>$X_{ideal}$</td>
<td>Variable</td>
<td>Ideal number of group mates within safe radius</td>
</tr>
<tr>
<td>$X_\Delta$</td>
<td>One group mate</td>
<td>Change in ideal nearby group mates</td>
</tr>
<tr>
<td>$E_t$</td>
<td>Variable</td>
<td>Amount of energy at time $t$ in an agent</td>
</tr>
<tr>
<td>$E_{ideal}$</td>
<td>100 e</td>
<td>Ideal internal energy for agents</td>
</tr>
<tr>
<td>$E_{loss}$</td>
<td>11/26 e/step</td>
<td>Energy loss per step for agents</td>
</tr>
<tr>
<td>$D_{safe}$</td>
<td>50 m</td>
<td>Safe radius</td>
</tr>
<tr>
<td>$D_{search}$</td>
<td>100 m</td>
<td>Food search radius</td>
</tr>
<tr>
<td>$E_{density}$</td>
<td>400 e/cell</td>
<td>Average density of energy for the resource grid</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.25–2.00</td>
<td>Shape parameter for a lognormal distribution (location = 100)</td>
</tr>
</tbody>
</table>

**Figure 1.** Algorithm used by individual agents balancing safety benefits and competition cost of group foraging.
Figure 2. Visual representation of event scheduling for one step in the model: e.g., five agents gather sensory information at the same time, and then perform actions at the same time. This procedure is then repeated until all agents have been processed (dashed line).

Figure 3. Environmental configurations used in the model, varying the size of grid cells, and spatial heterogeneity of resources.
3. Analysis

To measure behavioural differences between model runs, we monitor scramble competition among agents, and home range size of the group. Scramble competition was recorded as the number of times individuals attempted to feed from their chosen location, but failed to meet their feeding target (i.e., \( E_{\text{bite}} < E_{\text{bitemax}} \)). To estimate home range, the position of the group was recorded every time step in the model, and minimum convex hulls were used to estimate the area explored by the group.

We ran the model simulating three months of foraging 1000 times, where for each run a group of 50 agents were placed randomly within a 30-m radius of the center of a 2 \( \times \) 2 km landscape. For each of these runs a random seed was used, and parameter values were chosen using a Latin Hypercube Sampling (LHS) regime. LHS is a statistical method of choosing stratified-random parameter values from a range of possible values (Blower and Dowlatabadi 1994). Parameter space consisted of environmental heterogeneity (min = 0.25, max = 2), number of agents processed at once (min = 1, max = 50), and size of grid cells (min = 10 m, max = 60 m). To facilitate the interpretation of outcomes of varying parameter values, we make use of general additive models (GAM) to extract the relationship between these parameters and resulting scramble competition among agents. Using a GAM approach relaxes the assumptions about the relationship between dependent and independent variables, allowing for fitting nonlinear trends. We use the mgcv package (Wood 2006) in the R programming environment (R CORE TEAM 2015).

4. Results

We found highly nonlinear and heteroscedastic results when varying scale, scheduling and the heterogeneity of spatial variable (i.e., food). When a separate GAM was fit for each cell size category (10–60 m), we observed that the relationships between environmental heterogeneity and level of synchrony on scramble competition between agents were not consistent (Figure 4). In the case of environmental heterogeneity, the overall trend was reversed as cell size was increased, from a negative effect on scramble competition at low cell sizes to a positive effect at higher cell sizes. The relative deviation explained by each parameter suggests that at lower cell sizes, the effects of heterogeneity were more prominent than those of synchronization (Figure 4, Table 2). The altering trends in Figure 4 show an interaction between cell size and environmental heterogeneity on model results, and to a lesser extent cell size and synchronization. Whereas, the three-way interaction between cell size, environmental heterogeneity, and synchronization was only found to be significant at lower cell sizes (Table 2). The effects of cell size on its own showed large negative effects on scramble competition between agents (Figure 5).

The resulting effect of the amount of scramble competition between agents and group home range size (home range \( \sim \) scramble competition) was not found to be a strictly positive relationship, as might be predicted: e.g., more group motion as a result
of individuals not meeting their feeding requirements. We see a response shape in which increasing scramble competition initially lead to a rapid increase in home range size, followed by a tapered decrease in home range size as scramble competition increases (Figure 6).
Table 2. General additive models explaining scramble competition controlling for cell size of food patches.

<table>
<thead>
<tr>
<th>Cell size</th>
<th>Environmental heterogeneity Deviation explained</th>
<th>Synchronization Deviation explained</th>
<th>Heterogeneity × synchronization Deviation explained</th>
<th>Full model Adj r-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.74</td>
<td>0.01</td>
<td>0.01</td>
<td>0.97</td>
</tr>
<tr>
<td>20</td>
<td>0.23</td>
<td>0.04</td>
<td>0.28</td>
<td>0.55</td>
</tr>
<tr>
<td>30</td>
<td>0.38</td>
<td>0.06</td>
<td>0.00</td>
<td>0.37</td>
</tr>
<tr>
<td>40</td>
<td>0.13</td>
<td>0.24</td>
<td>0.01</td>
<td>0.32</td>
</tr>
<tr>
<td>50</td>
<td>0.16</td>
<td>0.11</td>
<td>&lt;0.01</td>
<td>0.26</td>
</tr>
<tr>
<td>60</td>
<td>0.28</td>
<td>0.14</td>
<td>0.08</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Deviation explained by each term is presented as well as the adjusted r-square value for the full model at each cell size value.

Bold values indicate smoothed terms with p-values <0.01.

Figure 5. Boxplot of total scramble competition among agents per cell size.

Figure 6. Plot of the relationship between scramble competition among simulated agents and home range size of the group. The fitted line was produced using a general additive model, the shaded areas represent 95% confidence intervals.
5. Discussion

5.1. Interaction between agent’s decision rules and environmental representation

Our results support and present another example of how changes to environmental representation can alter subsequent model outcomes. Changes to environmental representation (cell size) in our model showed the possibility of reversing the relationships between dependent and independent variables, e.g., the shifting trend of environmental heterogeneity on scramble competition (Figure 4). The resulting changes to scramble competition between agents due to cell size changes were found to lead to subsequent changes at the group level. Where the group’s home range size was found to have modal functional response to scramble competition, with increases in lower levels of scramble competition leading to increasing home range sizes as individuals within the group move to satisfy their feeding requirements. At higher levels of scramble competition, the increase in numbers of individuals moving for foraging purposes increased, resulting in lowering overall group movement, as the number of agents moving for safety was reduced. This resulted in the breakdown of group movement and smaller home range sizes in the model. This break down in group movement only occurred at cell sizes of 10 m, and in conditions of low heterogeneity (Figure 4 and Table 2).

These results show that the spatial agents used in our simple example were sensitive to, and in a sense adapted to, changes in environmental representation. As cell size varied so too did the mean and variance of the spatial units (e.g., food patches). The alteration of mean and variation of food cells had a direct influence on the decision landscape on which agents operated, e.g., changes to cell size resulted in mean food patch containing 10 units with 10-m cells, 100 units with 30-m cells, and 400 units with 60-m cells. Similarly, the variance among all patches increased as cell size increased: 0.03 with 10-m cells, 3.29 with 30-m cells, and 773 with 60-m cells, with this trend more pronounced in heterogeneous landscapes. Therefore, as cell sizes were altered from 10 to 60 m, agents interacted with an environment in which food was increasingly aggregated into increasingly larger food sites, and where variation among sites increased. The effects of increasing aggregation and variability in food sites on agent foraging algorithm (i.e., choose the closest site with the most food) lead to clear winners in terms of the best food site in which to forage. Cell size also interacted with the physical limits set on the visible range for each agent, resulting in fewer sites to compare (e.g., 60-m cells ≈ 3 sites vs. 10-m cells ≈ 100 cells to compare). The combination of aggregation/variability of food and a reduction of comparable visible sites culminated in a loss of diversity in choices and information, resulting in more conformity between agent decision making within the group. One result of this loss in resolution of the decision landscape was a decreasing level of scramble competition as cell sizes increased (Figure 5), as the group as a whole moved from high value patch to high value patch with lower dissention among agents as to where to move.

5.2. Effects of scheduling under different environmental configurations

With advances in parallel computing methods and hardware, which allow many agents to be processed at the same time (i.e., in parallel), there are large advantages in reducing
runtimes when synchronizing agent scheduling (Tang and Wang 2009). Long runtimes for simulations can limit the ability to explore and analyze a model (Tang and Jia 2014), often resulting in a trade-off between complexity of representation and computational time. However, the choice of synchronous or asynchronous scheduling has been shown in some cases to impact model outcomes (Caron-Lormier et al. 2008). This suggests that apart from runtime efficiency, there is a conceptual aspect to choosing scheduling methods, e.g., what best represents the timing of agent behaviours and their interactions (Cornforth et al. 2005)? Our results similarly suggest that scheduling choices can significantly alter model outcomes (Table 2), and that the ‘best’ agent scheduling might not necessarily be the one in which runtimes are reduced the most. Our results add to these findings by highlighting that the choices of representation can alter the effects of scheduling.

Our results show significant effects of synchronization and the interaction between synchronization and environmental heterogeneity at small cell sizes (10–20 m), with most of the variation explained by environmental heterogeneity (Table 2). Visually striking under these cases is the large increase in scramble competition at low heterogeneity (Figure 4). Inspecting group movement patterns under these conditions reveals a fluctuation pattern, where the group contracts and expands in successive steps. This pattern, not present under other simulated conditions, can be explained by the interaction between agent algorithm and environmental representation. Under homogeneous conditions, all individuals within the group tend to choose to move towards the best food site based on their current locations, resulting in a radial expansion of the group away from central depleted patches. In the following time step, individual safety requirements are largely not met, resulting in group contraction. This fluctuating pattern resulted in a breakdown in group movement as a successive pattern of collective movement and depletion events, resulting in a dramatic increase in scramble competition and lower home range sizes (Figures 5 and 6). As heterogeneity in the environment is increased, more consensus is achieved within the group about which areas to move towards as a simple result of the more patchy nature of the environment, ultimately reducing the fluctuating pattern. Similarly, a decrease in the synchronization of agent processing resulted in lower levels of scramble competition (Figure 4). As in more asynchronous cases, individuals processed first could influence the decisions of those processed after, again resulting in higher consensus as to which direction to move. As group movement patterns do not necessarily show a fluctuating pattern of expansion and contraction, a fully synchronized approach might not be the best representation of decision making within our model, even if it is the most efficient way to take advantage of parallelization of the model scheduling.

5.3. Navigating spatial representation and scheduling choices

Our results suggest that focusing on the properties of the decision surface influencing agent behaviour can help elucidate how representation and scheduling choices influence model outcomes. We suggest that spatial representation and scheduling choices
will influence spatial agents in two ways: through (1) changing the aggregation of spatial variables (i.e., food in our case) and (2) a reduction in resolution of the decision surface (i.e., large cell size choices leads to fewer spatial locations to compare/inform decision making).

Our results further highlight the interaction between spatial representation and scheduling choices, where the degree to which changes resulting in aggregation patterns and surface resolution which increase direct or indirect agent–agent interactions will increase the effect of scheduling choices. The nature of that effect seems likely to be highly dependent on the specific behavioural algorithm of the agents. As in our case highly synchronized scheduling lead to increased scramble competition in both small and large cell size representation due to different mechanisms: in large cells, occupied by many agents, high synchronization led to many agents attempting to feed from recently depleted patches, whereas in small cell representations synchronization led to a decrease in consensus making as to where to move (i.e., fluctuating pattern). It is a possibility that with a more general comparisons of behavioural algorithms it might be possible to identify certain classes of algorithms as sensitive or not to changes in aggregation or resolution of decision surface(s). For example: algorithms with fixed visual radius vs. Von Newman neighbourhoods to compare food sites. Furthermore, it is likely that the effects of synchronization on model outcomes will be more pronounced in models which abstract time into comparatively larger discreet time blocks, such as in our model (i.e., time step 30 min), where individual decision making is not considered an instantaneous reaction to its current context.

We find little mention of the effect of the underlying statistical properties of spatial variables on spatial representation choices. Although these are quite apparent when comparing extreme examples, such as a purely homogenous surface, where cell size changes will not alter the relative aggregation patterns on the landscape. We show here that heterogeneity of the surface will affect the outcomes of spatial representation and scheduling choices, but other spatial statistical properties, such as spatial autocorrelation could similarly play a role. Indeed, using our understanding of aggregation and resolution we would expect that on surfaces where there is high spatial autocorrelation, effects of aggregation due to cell size choices would be predicted to decrease in importance as nearby cells will be of similar values. Through a similar argument, changes in resolution would be expected to also be lower in the presence of high spatial autocorrelation as nearby cells will remain of relatively similar values.

In developing spatial simulation models, varying spatial representation and scheduling parameters and comparing model results to observed patterns is one approach to dealing with uncertainty in representational choices. This approach integrates well with the pattern oriented modeling method of model selection, and the model validation process in general (Brown et al. 2005, Grimm et al. 2005). Using this approach, alternate choices of spatial units, as well as scheduling choices, can be compared based on their success at reproducing observable patterns during a validation process, offering a way to quantitatively search for an agent–environment representation which best fits observed data. Similarly, for already built models, performing a sensitivity analysis (Fonoberova et al. 2013, Parry et al. 2013) on the model based on changes in representation choices could provide a degree of understanding of the importance these choices on model outcomes. However,
both these approaches add additional parameters to the validation and sensitivity analysis of models which are, in general, already heavily parameterized.

The awareness that many spatial model can be explained as combinations of simpler base processes (O’sullivan and Perry 2013) offers a way to better communicate and understand spatial simulation models. Furthermore, the ability to share and create models from already built models is becoming more common (e.g., OpenABM.org). These factors suggest that mash up models are likely to become more common. A better understanding of the mechanisms by which spatial representation and scheduling changes will be important when combining and altering existing models. Focusing on the changes to aggregation, resolution and their interactions with both scheduling and statistical properties of spatial variables could provide essential insights into potential problems for specific mashups.

6. Conclusions

Our results provide further evidence of the sensitivity of model outcomes to changes in spatial and temporal representation, and highlight their synergistic effects on model outcomes. We suggest mechanisms by which alterations to spatial representation affect agent behaviour: namely, through changing in aggregation patterns and resolution of spatial variables. Similarly, given that synchronization of agents’ scheduling was found to have significant impacts on model outcomes, we suggest that when choosing a scheduling design to process spatial agents, conceptual factors should play a role, as well as runtime improvements. Both the effects of spatial representation and scheduling changes were found to be significantly affected by the underlying statistical properties of the spatial variable (e.g., mean spatial heterogeneity). We therefore caution that the results of spatial simulations will be sensitive to issues of representation choices (space and time), and that inferring behavioural processes under a given representation is not necessarily transferable to other representations. However, this challenge also presents an opportunity. When incorporating spatial representation and scheduling choices implicitly in the model selection process, important insights can be gained into the interactions between spatial processes and the environments in which they occur.

Acknowledgements

We would like to thank three anonymous reviewers and Mel Lefebvre for helpful comments on the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by funds from Natural Science and Engineering Research Council of Canada (C.A.C.) and Fonds de Recherche Nature et Technologies (T.R.B.).
References


