Encouraging Safe, Green, and Economically Sustainable Use of our Natural Environment

A. Berryman [1] and E. Donlon [2]†

[1] *University of Oxford, Oxford, UK*
[2] *Technological University Dublin, Dublin, Ireland*

(*Communicated to* MIIR *on 7 March 2022*)

Study Group: ESGI 165, Durham University.

Communicated by: Bernard Piette.

Industrial Partner: Natural Resources Wales.

Presenter: Luke Maggs.

Team Members: Michele Aleandri (Group mentor), Luiss University; Sophie Abrahams, University of Oxford; Anna Berryman, University of Oxford; Georgia Brennan, University of Oxford; Chris Budd OBE, University of Bath; Joaquim Correia, University of Évora; Edward Donlon, Technological University Dublin; Sam Kamperis, Oxford Brookes University; Daniel Mckinnell, University of Manchester; Brady Metherall, University of Oxford; Frank Millward, University of Manchester; Joe Roberts, University of Oxford; Keming Yu, Brunel University; Fanqi Zeng, University of Bristol.

Industrial Sector: Social.

Tools: Data analysis, agent-based modelling, regression modelling, game theory.

Key Words: Environmental statistics, agent-based modelling, regression modelling.

MSC2020 Codes: 62J05, 62P12, 91B76, 93A16.

Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 764850 (SAGEX). This is paper SAGEX-21-28.

† Corresponding Author: `edward.donlon@TUDublin.ie`

Summary

This report addresses anti-social behaviour at green spaces in Wales. Using the data available, this report investigates classifying sites and site users. This classification is used to understand specific anti-social behaviours with agent-based modelling. Regression modelling is also used to calculate a site user's impact on anti-social behaviour, and the scaling behaviour of associated quantities of interest as the number of site visitors increased was examined. Each area of research provides a different lens to understand what is happening at sites across Wales.

1 Introduction

In recent years, use of Wales' green spaces have increased and problem behaviours have been exacerbated by the COVID-19 pandemic. Problems rising with the increased use include littering, over-crowding, clashes between park users, and wildlife disturbances. Understanding how users affect different sites during their visits is vital to solving some of these problems.

National Resources Wales (NRW) want to facilitate responsible and sustainable outdoor experiences at sites ranging from sprawling national parks to local green spaces. However, once a user has decided to visit a site, it is difficult to influence their decision and any reactionary solutions are often too late to prevent issues. Volume of users alone does not account for everything; if visitors have different wants, this can cause friction even with low numbers of visitors.

The Study Group was asked: How do users affect sites? How do we better redistribute resources? Where is best to intervene to disrupt bad behaviour? How do we balance the negative experience of users against negative effects to sites? These questions all need to be answered with limited resources for sites, and limited data.

This report is structured as follows. In Section 1.1, we introduce the data used throughout the analysis. In Section 2, we categorise sites and site users to understand the differences in behaviour between them; and Section 3 uses this categorisation to investigate specific anti-social behaviours using agent-based modelling. A data-driven approach is explored in Section 4 with conclusions and further work in the final Section 5.

1.1 Data

NRW provided a variety of data to help predict bad behaviour. In the following section we use three data sets in particular:

- **The Wales Outdoor Recreation Survey (WORS) 2014**—WORS contained data about the exercise habits of a 6000 person sample of the population of Wales. Participants were asked about how often they visited the outdoors, how far they travelled to their destination and which activities they had participated in within the last month, amongst other questions.
- **Car Park, Walkers, Mountain Bike, and Visitor Centre data from a number of Forest Parks**—Monthly footfall data from a variety of sites across Wales. Data on

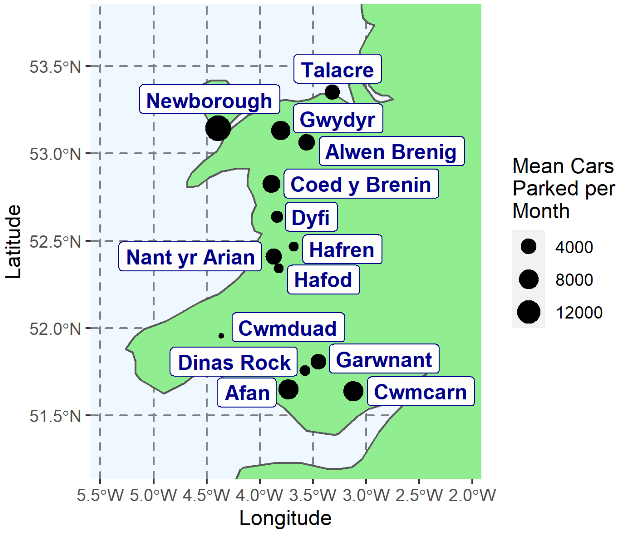


Figure 1. Mean number of cars parked at each site per month.

walkers and mountain bikers was collected using counters placed at the side of paths. NRW use twice the car park capacity as the best estimate for the total number of visitors to a site.

- **Questionnaire Data from Park Rangers on the Impact of Bad Behaviour**— Results from a qualitative survey of park rangers rating the impacts of anti-social behaviour on a scale from 1–5. The rangers were asked about the impact of behaviour on their time, park budgets and the environment.

We illustrate the data on monthly car park users from October 2003 to December 2017 in Figure 1. We see that Newborough is the busiest site for cars, and Afan, Cwncarn and Gwydyr are the others with greater than 700 cars on average.

In Figure 2, it is shown that the number of visitors using the car parks peaks during the summer months, as expected. It is therefore important that resource allocation for the sites is weighted around the summer period. This information could be used to build a model to classify different sites based on car park peak months. Further work could also involve a model incorporating pedestrian to cyclist ratios.

Figure 3 shows that there is strong evidence to suggest that car park behaviour can be predicted from other locations. Strong correlations are present in data from the Afan and Alwen Brenig car parks.

2 Categorisation of users and sites

To enable NRW to understand what is happening at different sites around Wales, we will first classify different users and different sites. We will leverage the data available to determine this categorisation. Since different users interact with sites in different ways, when we investigate the effects of anti-social behaviour at different sites this categorisation will be important.

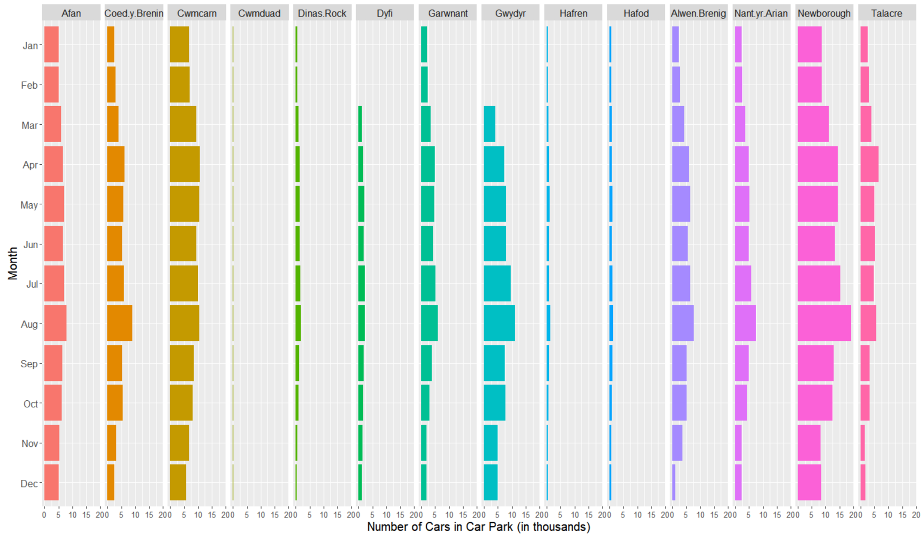


Figure 2. Distribution of the number of cars parked (in thousands) at each site during the year. Periods with no available data are indicated by the absence of bars.

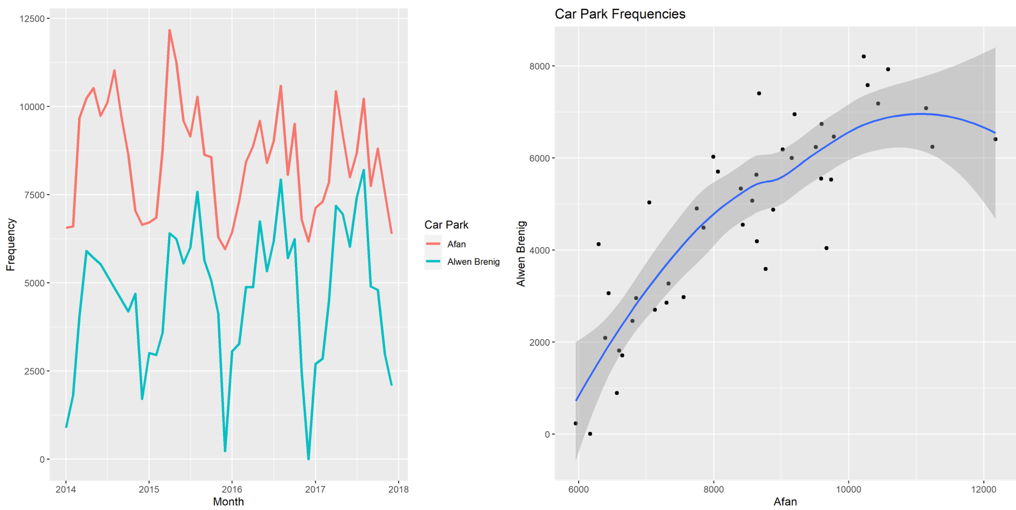


Figure 3. Afan and Alwen Brenig car park visitor data.

2.1 Users

We base our categorisation of site users on a ‘purism scale’ defined by Tverijonaite *et al.* [4]. In the study by Tverijonaite *et al.*, visitors to five natural sites in Iceland answered multiple choice survey questions to give them a score between 0 and 70, from which they were categorised into a scale from ‘purist’ to ‘urbanist’ [4]. ‘Purists’ are likely to be frequent site users who are more willing to travel farther to enjoy green spaces and prefer unspoiled natural sites. At the other end of the scale are ‘urbanists’; these people are likely to be infrequent site users who choose to visit nearby sites with conveniences such as toilets and cafes. In this study, we choose to allocate users to three distinct categories: urbanists; purists; and neutralists who sit at the centre of the purism scale.

We use the Wales Outdoor Recreation Survey (WORS) to examine the characteristics (behaviours) of different types of users. For each respondent, we first place them on the purism scale, then analyse their behavior based on their answers to the WORS. To categorise users, we find a number, x, by adding their answers to the following questions: Q3a *How many visits to the outdoors have you made in the last 4 weeks?* and Q8 *Approximately how far, in miles, did you travel to reach this place?*. We use these questions as distance travelled and frequency of visits is indicative of a persons purism. x is re-scaled by the maximum x of all respondents to give a number between 0 and 1. The user's purism score, p, is then $p = 1 - x$, where 0 is most purist and 1 is most urbanist. We define $p \leq 0.4$ to be purists, $0.4 < p < 0.65$ to be neutralists, and $p \geq 0.65$ to be urbanists.

For each category of users (purists, neutralists and urbanists), we analyse their WORS answers (removing Q8 and Q3b which were used for categorisation) to establish the likelihood of them displaying different attributes or behaviors. The results of this analysis for some chosen attributes is displayed in Table 1. For a given category of user, questions Q15a *During this visit, did you personally spend any money on any of the following items?* and Q14 *Were you accompanied by a dog on your most recent visit to the outdoors?* gave us direct information on the proportion of users that: buy food and drink; buy gifts and souvenirs; and visited with a dog. From Q5 *During this visit, how long did you spend in the outdoors?*, we can directly calculate the mean time spent at a site. We use the answers to question Q24 *Which of the following activities you have done at least once in the last 12 months to help protect the environment and nature?-1. Recycled*, as a proxy for the chance of littering, where we make the assumption that users that don't recycle will leave litter. We estimated how keen the different groups are to move away from other people by how many answered Q15 *What reasons, if any, best describe why you made this visit to the outdoors* with A1 *For peace and quiet*.

In Table 1 we can see that, using our method, urbanists are more likely to litter, are more social, and buy more food and drink, whereas purists move away from others more, buy more gifts and souvenirs, and are more likely to have a dog. We can also see how the time spent in the sites varies between the different groups; urbanists generally spent shorter amounts of time at a site, while purists stayed longer. In Figure 4 we see the data available on time spent. We report the average time spent by the different groups in Table 1, but the key result is the relative times between the groups.

We also calculate a responsibility factor using the survey data to understand how concerned users are about wildlife and the environment in Wales. This factor was taken as a proxy for how likely park users would be to disturb wildlife. Question 24 of the WORS asked users which types of action they has taken to help the environment. We used the questions asking were whether they had gardened for wildlife, been part of a conservation group or volunteered at a project to protect the environment. Each user was given a score from 0-3, based on the number of actions that they had taken. The responses were averaged and the means were normalised. Combining these responses yields a responsibility factor of 0.32 for urbanists, 0.38 for neutralists and 0.44 for purists.

Table 1. Characteristics of site users.

Attribute	Urbanist	Neutralist	Purist
Mean time spent in site	2h 4m	2h 23m	3h 10m
Responsibility factor	0.32	0.38	0.44
Fraction who litter	0.027	0.019	0.018
Fraction who buy food and drink	0.09	0.05	0.05
Fraction who buy gifts and souvenirs	0.08	0.14	0.18
Fraction who have dog	0.19	0.27	0.32

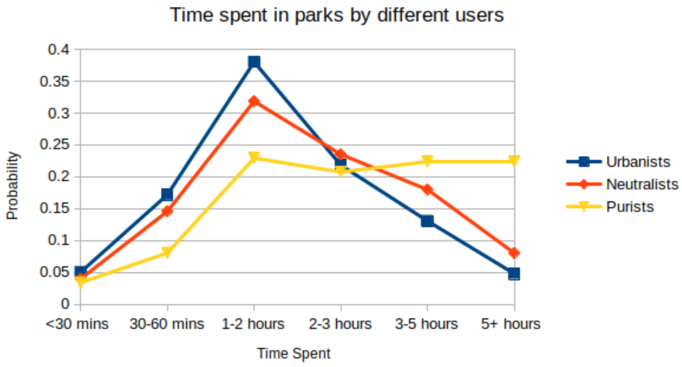


Figure 4. Time spent at a site, by user group.

One limitation of the survey data is that respondents were from a variety of different sites across Wales. Site specific questions such as buying food and drink or gifts and souvenirs will be affected by site differences. Some users would not have been able to buy food and drink as there is no cafe at the site they visited, however, might have purchased something had there been a cafe.

2.2 Sites

To give each site a 'purism score', we use data for 2004-2018 that includes: total monthly site visitor numbers; total monthly shop visitor numbers (if a shop is present); total monthly cafe visitor numbers (if a cafe is present); and total monthly car park numbers. We also calculate the distance from the park to the nearest city.

Using the data, we give each site a score, S, between 0 and 1 for a given day. This score will later be used to determine the distribution of site visitors in the three categories, purists, neutralists, and urbanists for that month. We propose the following simple formula to classify each site with a site score S:

$$\mathrm{S} = \alpha_1 d + \alpha_2 v + \alpha_3 c, \tag{2.1}$$

where d, v, and c are all between 0 and 1, and represent respectively: the relative distance to the nearest city; estimated daily number of visitors for the given month; and facilities such as number of toilets, whether a shop/visitor centre exists, and the number of parking

Table 2. Site scores for Cwmcarn and Coed-y-Brenin in August and December.

Site	Site score (August)	Site score (December)
Cwmcarn	0.77	0.73
Coed-y-Brenin	0.57	0.49

spaces. The coefficients α_n can be adjusted, ensuring $\sum_n \alpha_n = 1$, to account for the relative importance of d, v and c.

For a given site, in a given day, v is calculated using the monthly total visitor number in the site data for the same month in the previous year. This figure is divided by the number of days in the month to give an average predicted daily value for that month. v is normalised by max daily visitors for any site in any month (1200 visitors). d is calculated using the distance between the site and the nearest city, normalised using the maximum distance a site is from a city (i.e the most "remote" site). The value of c could be calculated in a number of different ways, with the aim of quantifying the facilities a site has. For example, we could focus on visitor centres only by assigning $c = 0$ if there is no visitor centre, and $c = 1$ if there is a visitor centres. More terms could be added to include things such as a gift shop, cafe etc. One could also introduce the number of toilets on the site by having a scale from 0 to 1, normalised by the site with the most toilets.

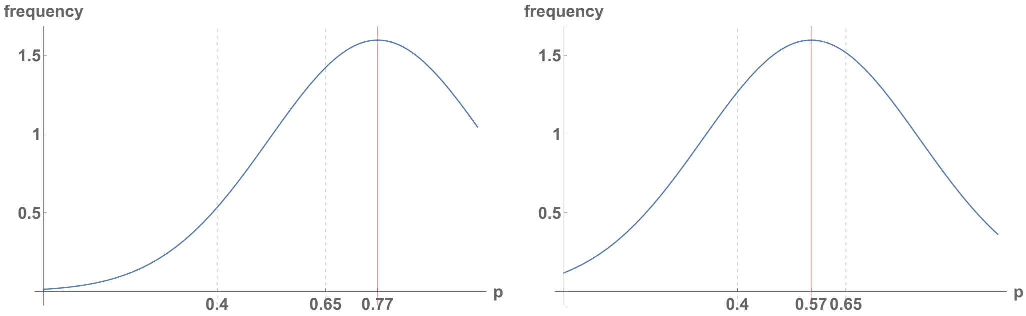
In Table 2, we see the site scores for two sites in August and December. Cwmcarn is near Cardiff and Newport and tends to have more visitors than Coed-y-Brenin in Snowdonia. We will use these scores in Section 2.1 to understand the agents likely to be visiting each site. We see from Table 2 that the site score changes depending on the season due to the fact there are more visitors in the summer than in the winter, where as the location and the facilities remain the same. This suggests that as well as different numbers of users visiting the sites in different seasons, there will be a different distribution of types of users.

2.3 Site score to agent type

We use the site score, S, to determine the expected proportion of purists, neutralists and urbanists at the site on the chosen day. To determine the proportions of the three groups of site users, we propose a normal distribution for user purism score p, bounded between 0 and 1, with mode S, as in Figure 5. We integrate this distribution over three regions where purists have $0 \leq p \leq 0.4$, neutralists have $0.4 < p < 0.65$ and urbanists have $0.65 \leq p \leq 1$. The integrals are normalised by the total area under the distribution, to give the the proportions of agents that are purists (P), neutralists (N), and urbanists (U). For now, we take the variance to be always 0.25 however, with more data this parameter can be fitted.

Using the two site scores in Table 2 we calculate the ratio of users expected at the two sites in August and December in Table 3. Cwmcarn is expected to have more urbanists than Coed-y-Brenin, and in summer there are more urbanists than in winter.

Site and user purism scores give NRW a new tool to understand the visitors at a site



(a) Cwmcarn in August, $S = 0.77$ (b) Coed-y-Brenin in August, $S = 0.57$

Figure 5. Distribution of site users with user purism score, p. The distribution is assumed to be a normal distribution with mode S.

Table 3. Proportion of site users that are purists (P), neutralists (N), and urbanists (U), for two sample sites in August and December.

Site	P:N:U (August)	P:N:U (December)
Cwmcarn	0.08:0.30:0.63	0.11:0.33:0.56
Coed-y-Brenin	0.25:0.40:0.35	0.35:0.40:0.25

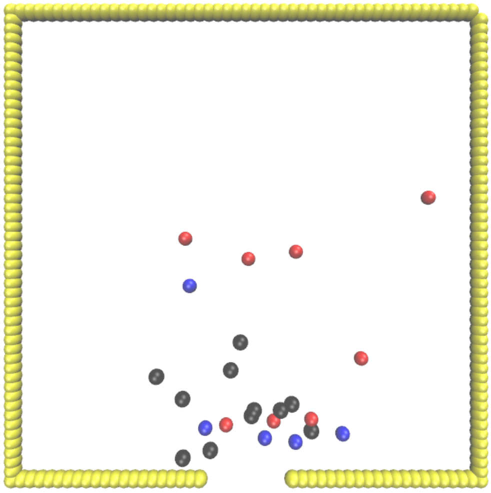
on a given day. The predicted ratio P:N:U gives an understanding of which types of users are predominantly expected at a site, while the characteristics in Table 1 indicate how we expect that type of user to behave. This information can be used as an input into an agent based model (ABM) as will be discussed in the following section; the proportions of purists, neutralists, and urbanists and their typical behaviour can inform the ABM and enable us to simulate and investigate different anti-social behaviours at a site.

3 Agent-based modelling

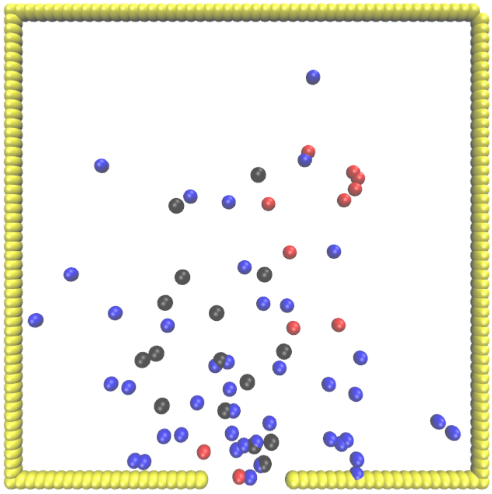
To tackle the question of how the public use Wales' natural environment, and how we might model anti-social behaviour, we develop an agent-based model (ABM) [1, 2] to simulate users moving around and interacting within a site. We begin by writing our own agent-based model[1] in `C`.

The underlying dynamics of the agents are given by a Langevin equation with drift, inside a two dimensional, square domain, with an entrance / exit. Each agent has an associated preferred activity, and each activity has a corresponding location within the site. As a concrete example, we assume our particular site contains a picnic area in the lower right, mountain biking trails in the upper right, and hiking trails throughout the whole site. Agents' preferred activity and their proximity to the activity's location determine the drift term in their motion. An agent will drift toward their activity location until they are reasonably close, at which point the bias is turned off. In the case of hikers

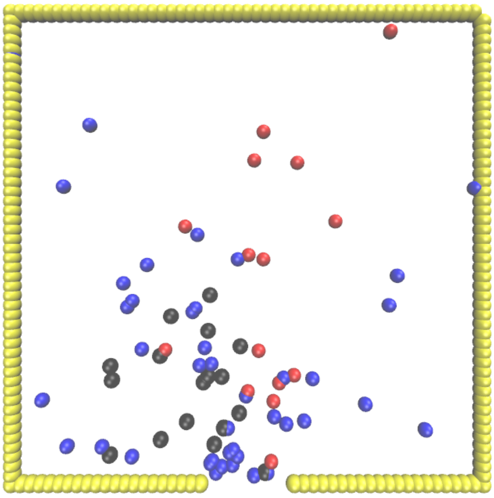
[1] Code available at `https://drive.google.com/file/d/16qtg8Y2y8c-zw2vTYkjonDvCYku8luyD/view?usp=sharing`.



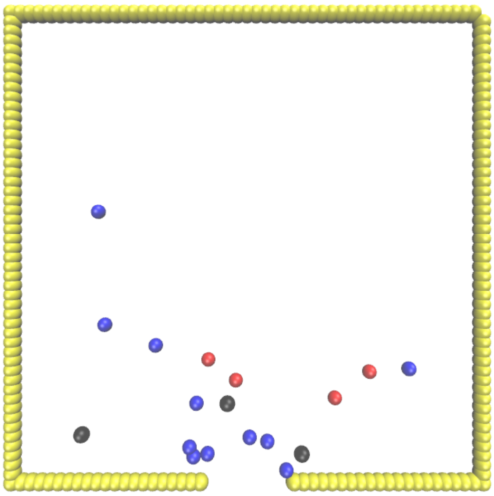
(a) Site opens and agents begin to diffuse in.



(b) Agents loosely cluster based on their preferred activity.



(c) Agents cluster in a more defined manner based on their preferred activity.



(d) Site will be closing soon and agents leave.

Figure 6. Sample run of our agent-based model. Different agents are drawn to different regions of the site based on their preferred activity (red agents: upper right, black agents: lower left, blue agents: no preferred region). Visualisations created with VMD [3]. Full video at https://drive.google.com/file/d/1f1Xyt_d-lj2cWCLqmDDHw7OQLSPsj8DV/view?usp=sharing.

in our example, they have no drift term in their dynamics since the hiking trails are throughout the site. Agents also enter and exit the site by a random decision variable. When an agent enters the site they are simply placed at the entrance. However, once an agent decides to leave, they drift to the exit, and then are removed from the site.

An example run of this agent-based model can be seen in Figure 6. In Figure 6a we

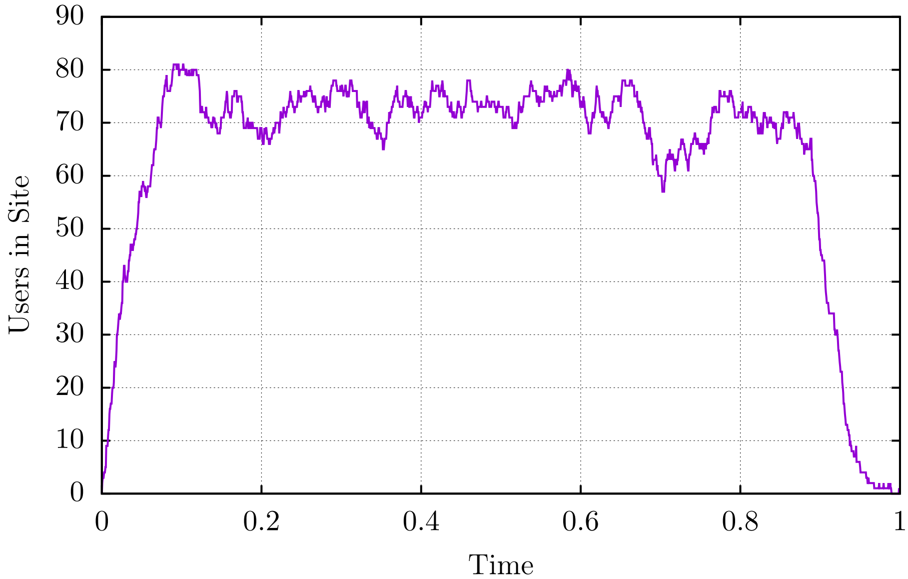


Figure 7. Number of agents in the site over the course of a day.

see agents begin to enter the site shortly after the simulation begins. Closer to midday, as shown in Figures 6b and 6c, generally agents are in their preferred activity region. The picnickers (black) are in the lower left, the mountain bikers (red) are in the upper right, and the hikers (blue) are fairly uniformly distributed throughout the site. The main exception to this is agents that have recently entered the site and are near the entrance, or agents that are leaving and are close to the exit. Finally, as the simulation nears the end all agents head to the exit as the site is closing (Figure 6d). Our simple model allows us to easily measure properties of the system such as how many users are within the site during the course of a day. We show this in Figure 7. We find that agents quickly enter the site once the simulation begins. Generally, there are 70–80 agents within the site, out of a possible 100, with one short dip below 60. As the simulation nears the end, the agents leave the site and the number quickly drops, mirroring the opening of the site.

To make more efficient use of our time, and to more easily incorporate complex interactions, we transition to using NetLogo [6]. NetLogo is a multi-agent modelling environment used in biology, social science, and mathematics. NetLogo also allows us to provide a more complete tool that NRW can use and develop further themselves, since analysis and visualisations can be done directly within the NetLogo interface.

3.1 Littering

A key anti-social behaviour that affects Wales' green spaces on a daily basis is littering. To illustrate the potential power of the NetLogo software for NRW, we investigate how bin placement affects littering among different populations of users. We will use the user ratios and characteristics found above to reflect different sites at different times of the year.

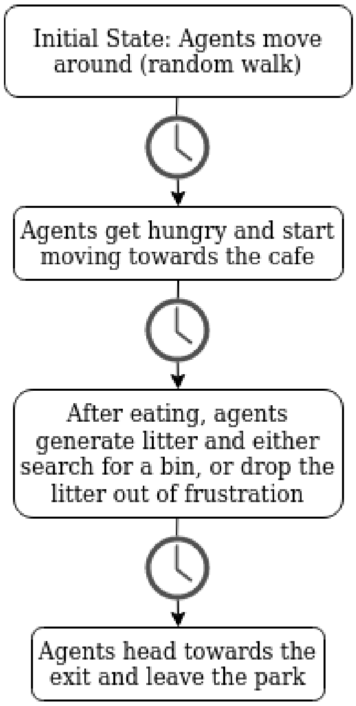


Figure 8. Littering state chart.

The agents in the model arrive from the entrance, and follow a random walk around a square domain. Each agent will start to feel hungry after a random time, drawn uniformly between 60 and 180. At this time, the agent begins to head towards the cafe in the centre. After consuming their food (while continuing to move around the domain) for 10 time steps, the agent now holds litter. Once an agent holds litter, they will feel increasingly frustrated. After some time, drawn randomly between 15 and 23, an agent will have no more patience to hold onto their litter and will therefore drop it, regardless of whether they are at a bin. Not all litter will be dropped around the domain however.

Agents in the three different groups will be attracted to the bins within the domain with different strength. If an agent holding litter gets within a certain radius of a bin, they will then head towards that bin to drop their litter in. The radius at which agents head towards the bin is 2, 5, and 8 for urbanists, neutralists, and purists respectively. When an agent shares the same square as a bin, and they are holding litter, they will use the bin to throw away their litter.

After some time, the agent will head towards the exit and leave the site. This time is drawn uniformly from $[120, 180]$ for urbanists, $[150, 225]$ for neutralists, and $[180, 270]$ for purists, based on the typical duration of visits in Table 1. The agent's state chart is illustrated in Figure 8.

Using the user ratios from Table 3, we run three experiments for each site in August and December. We consider one bin, either north or south of the cafe, and two bins both north and south of the cafe. The results of these simulations are in Table 4. Consistently

Table 4. Littering for 100 agents with the ratios given by Table 3 (litter in bins / litter on ground).

Site	Month	North bin	South bin	Two bins
Coed-y-Brenin	August	31/51	27/45	49/31
Coed-y-Brenin	December	34/47	36/45	55/27
Cwmcarn	August	26/50	15/51	32/45
Cwmcarn	December	26/52	14/53	37/39

two bins capture more litter than just one. When there is only one bin, the north bin collects more litter than the south, since the exit is north of the cafe.[2]. Videos of the simulation for Coed-y-Brenin in the summer, and Cwmcarn in the winter are available at https://drive.google.com/drive/folders/1so1ZnEg9_utrBTPSamZc3JAEGIYcQV51?usp=sharing.

To further illustrate how NRW can use this tool in the future, we model the total litter build up over the course of a 10 days to visualise the distribution of litter. These simulations use the ratio of users expected at Cwmcarn in August.

In Figure 9, we can see that with no bin, Figure 9a, litter accumulates mostly near the cafe and exit. In Figures 9b, 9c, and 9d, the volume of litter dropped is reduced—the north bin enables agents to dispose of their litter while leaving, greatly reducing the litter around the exit. In Figure 9e, bins are placed at the cafe and exit, where we found the most littering in Figure 9a, which greatly reduces the amount of litter dropped. Placing six random bins (Figure 9f), the total litter is reduced further and most litter accumulates sort of north-east of the cafe where there are no bins.

We hope that this simple example will provide NRW with a tool that can be used in future discussion around distribution of resources. We have incorporated differences between sites through the distribution of different site users. The 'landscape' of the simulation can be further tailored to different sites to enable discussions at the site level. In the next subsection we will consider a site with a protected area.

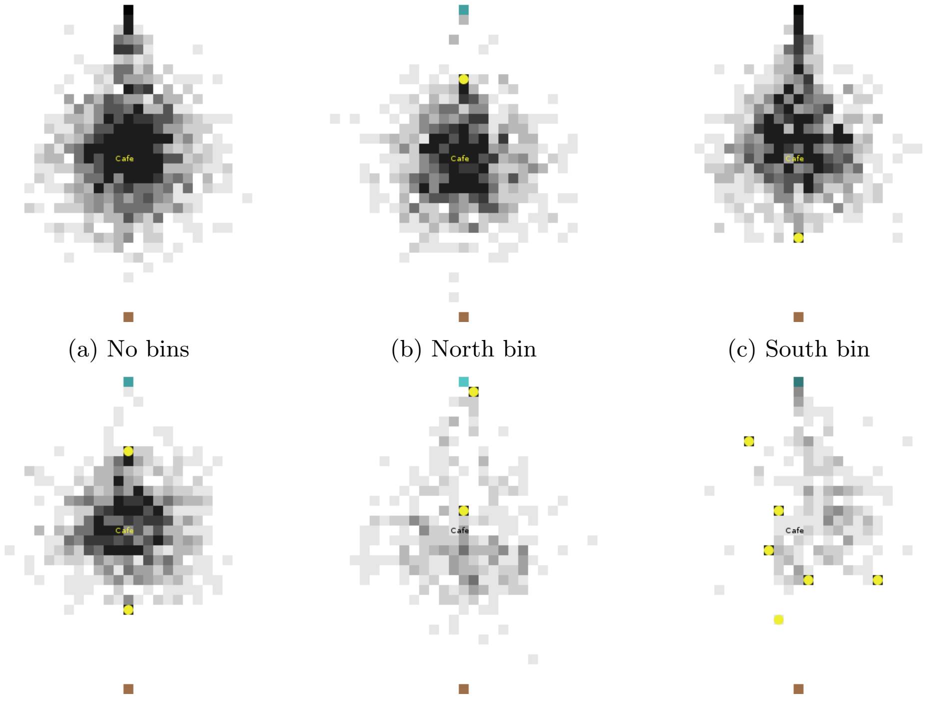
3.2 Protected areas

Another anti-social behaviour that NRW try to discourage is users damaging protected habitats and species. To avoid protected animals from mixing with users, especially those with dogs, various strategies such as signage and fencing can be used. In this section we set up a simulation using NetLogo[3] in `C.` of behaviour around a protected area which can then be developed by NRW to look at the impact of such strategies.

The simulation begins with agents from each of the three classes being randomly assigned one of three particular starting locations within the site. The initial populations for each class of agent are approximated using the method outlined in Section 2.3, which

[2] Code available at https://drive.google.com/file/d/1TQa2NfDYilWHfnPHAfy29r3jfq6RX4-4/view?usp=sharing

[3] Code available at https://drive.google.com/file/d/1PevUHxrsyl9UeuATXSpfpmf0ZEwQ7dZb/view?usp=sharing.



(a) No bins (b) North bin (c) South bin

(d) Two bins (e) Cafe and exit bins (f) Six random bins

Figure 9. Litter distribution at Cwmcarn in August over a week (with no litter picked up) with different bin configurations. The entrance is the southern, brown square and the exit it the northern, cyan square. Bins are yellow, the cafe is in the middle of the site, and litter build-up turns a square from white to black.

take into account the type of sits, its location and the season. Agents then embark on a random walk within a certain fixed radius of their starting location, emulating a real park user's wander. After a certain time t_p, each agent then randomly chooses a desired destination within the park (i.e. one of the two other locations) with equal probability. The agents then decide whether to take a short route, or a more responsible route around the protected area. This decision is made based on a park user's characteristics: there is a presumption made here that a purist is more likely to take the responsible route than an urbanist. The probability an agent takes the responsible path is

$$P(\text{responsible path taken}) = \frac{d_v}{d_r}R, \tag{3.1}$$

where d_v is the euclidean distance from the agent's current location to the new destination as the crow flies, d_r is the euclidean distance if the agent avoids the protected area, and R is the responsibility factor of the agent given in Table 1. If the agent opts to avoid the protected area, then they choose a predetermined alternate route to their destination. In this simulation, agents traversed the corners of the protected area in such a way that they avoided walking through it. An alternative (and perhaps more realistic) approach for future work is to implement a boundary detection process that would guide agents along the perimeter of the protected area.

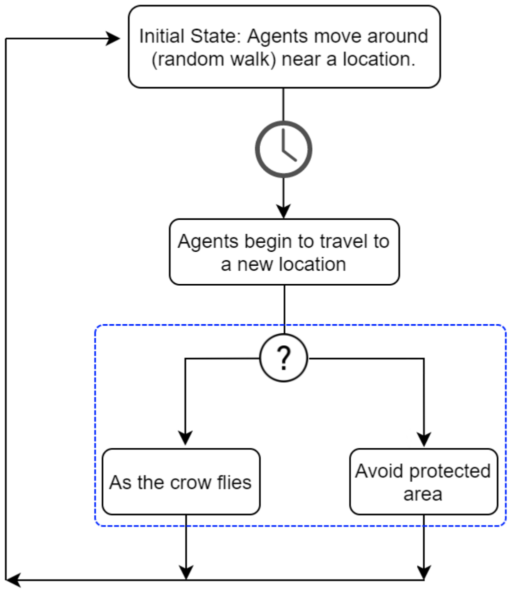


Figure 10. Protected areas state chart.

If the agent decides to take the direct route, damage to the protected area (per unit area) is computed as

$$D(t) = \sum_i \lambda_i n_i(t), \tag{3.2}$$

where $i \in \{\text{urbanist, neutralist, purist}\}$, λ_i is the probability of owning a dog given agent type i (given in Table 1, and $n_i(t)$ is the number of agents of type i in the area per time step. When an agent arrives at their destination, they begin the process again. This process is illustrated in Figure 10.

In Figure 11, we see the resulting damage for Coed-y-Brenin (in Snowdonia) site users during a typical summer season. Interestingly in this simulation, purists cause the most damage to the protected area. Although they are the most likely to take the long route around the area, they are also most likely to have a dog and therefore have the ability to cause more damage, were they to cross the protected area. Further extensions of this set up could be to include signage to discourage users from entering the protected area. Videos of this simulation and an analogous simulation for Cwncarn are available at https://drive.google.com/drive/folders/14hVPo82rO_jFNVGXg6DjCoxIVXZBLUuT?usp=sharing.

3.3 Further work

These two examples, of littering and the protected area, give insights into the power of agent-based modelling using NetLogo for NRW. There are many extensions that could be added to these simulations, starting with combining the two to start building a more complex model of a real site. Different geographies, features, and anti-social behaviours

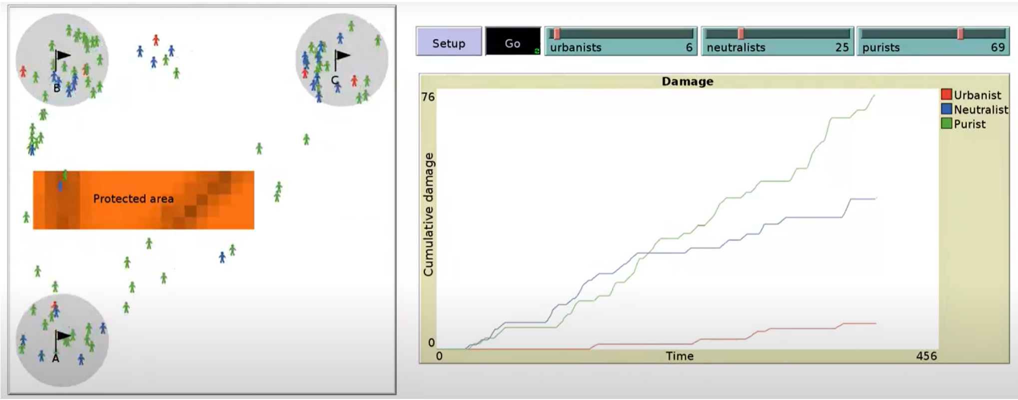


Figure 11. Screen shot of the NetLogo simulation of the protected area.

can also be included to make the analysis more site specific. For example, agents that prefer to be alone in nature could modelled with the Weeks–Chandler–Andersen (WCA) potential to be repelled by other agents. Similarly, families or groups of friends could be modelled using the Lennard-Jones potential so they stay clustered together in a group. Finally, for very large and complex systems, NRW may wish to look into ESPResSo[4] [5]. While ESPResSo is primarily used for molecular dynamics simulations, it was designed and optimised for simulating systems with expensive interactions.

4 Data-driven approach

In this section, data-driven models are considered as proposed frameworks in order to gain a better understanding of the factors that may influence anti-social behaviours.

4.1 Regression modelling

One could use the car park data and the rangers survey data (introduced in section 1.1) to produce a linear regression model in order to predict the impact that a park user has on overall anti-social behaviour. A single location (Afan Forest Park) was considered, and data regarding three types of undesired behaviour were included (littering, off-road parking, and dog fouling). By using the park rangers' estimations of the impacts of this behaviour on their time, their relative impact with differing numbers of people in the park can be predicted. The model is defined as:

$$Y_{ij} = \alpha_i + \beta x_j + \epsilon_{ij} \tag{4.1}$$

where Y_{ij} is the impact score to bad behaviour i from visitor j, α_i is the mean impact of bad behaviour i, and ϵ_{ij} is the error. Calculating the mean occupancy of the car park and variations from that, due to seasonality, will allow us to predict impact of bad behaviours over time.

In Figure 12, we see that when it is crowded, off-road parking tends to have the higher

[4] Extensible Simulation Package for Research on Soft Matter

impact on the park rangers' time than littering and dog fouling. Dog fouling has lowest impact, which may mean that visitors at Afan park have travelled longer distances and so are less likely to have brought their dog with them. On the other hand, the low score of Afan park may mean that, in off-peak seasons, local people are more likely to visit. They are then presumably more likely to walk their dogs, and then dog fouling has the highest impact on park rangers' time. Then off-road parking is not a main issue in the low car score seasons.

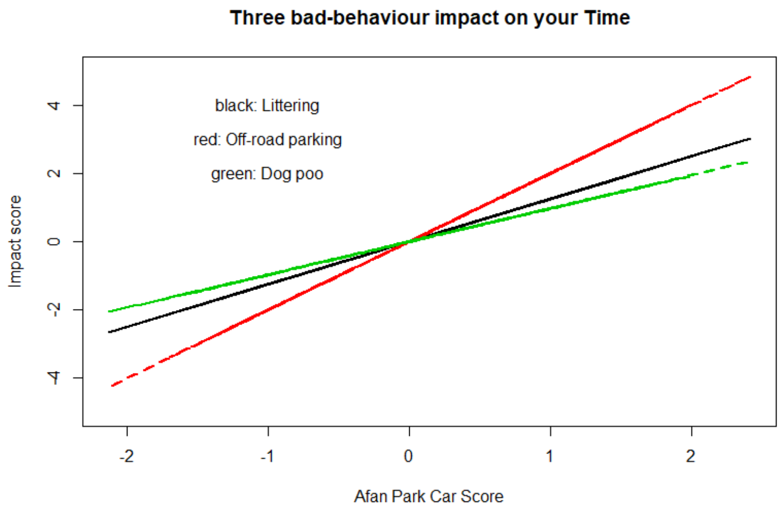


Figure 12. Anti-social behaviours plotted as a function of time for Afan car park.

Figure 13 shows that off- road parking increases steeply with car park occupancy. Dog fouling has a lower mean impact, but will be a more significant problem at low car park occupancy. Littering is most clearly tied to the occupancy of the car park. It increases at a similar rate to the impact of dog fouling, but can be more time consuming at higher car park occupancy.

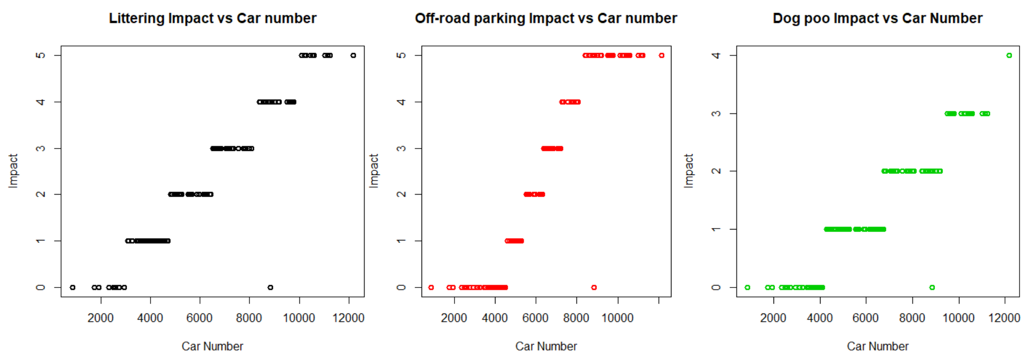


Figure 13. The impact of three undesired behaviours as car park occupancy increases.

4.2 Utility functions

In the previous section, it was demonstrated that there is a correlation between the number of cars in a car park and the impact of bad behaviours in the parks, in particular off-road parking, littering, and dog fouling. In this section we want to focus on the relationship between the number of agents visiting a park and the costs/benefits for parks.

Moreover, we want to define a set of utility functions for the agents and the park that can be used to find what is the best number of agents in the parks to maximize the benefits and minimize the costs. This analysis is possible after the parameters of the functions are fitted with real data, then this approach has the goal to indicate what parameters you can look at for future data collection.

Let us define N as then number of park users per day. The possible *costs* that affect a park are: financial, environmental (direct) eg. destruction of area, environmental (indirect) eg. larger numbers of tourists, cars, visual impact, and the concentration of tourists in one park can lead to a loss of income of the other parks.

On the hand end as the number of park visitors increases the *benefits* are: more tourism leads to more income, focused park users in only one area avoid general damage, and more control and more information to maintain the parks.

From this, one could define an *happiness* function for an agent that is a measure of the satisfaction of an agent of visiting a park. One proposed model for the happiness of park user i is

$$H_i(N) = K(\alpha(1 - \omega t_i(N))) - \beta f_l(N) - \gamma \sum_{j \neq i} \omega(\nu_i, \nu_j) - C_i \tag{4.2}$$

where

- ωt_i is the proportion of time wasted in the park, due to queuing for facilities,
- f_l is a function representing the impact of littering,
- $\omega(\cdot, \cdot)$ measures the impact of the interaction,
- C_i is the fixed cost to visit a park,
- K, α, β, γ are the model parameters taken to normalized the contributions of each term in the function.

From the point of view of a park authorities, the main goal is to maintain the nature of the park and to offer a service to the citizens. Then it makes sense to define a function of *happiness* of a park in order to measure the health of a park. Let H_P be the happiness of the park rangers, such that

$$H_P(N) = K_P \sum_i H_i(N) - C_e(N) - C_{\text{int}} + I(N) \tag{4.3}$$

where

- C_e represents the cost of the environment that is affected by the bad behaviour of the agents and relative consumption,
- C_{int} represents the cost of the intervention by the authorities to provide facilities to the tourists,

- I measures the impact of incomes from the park users, they can be used to keep the park cleaned, safe etc,
- K_P is a normalization constant.

The goal is to maximize the functions (4.2) and (4.3), but they need to be justified with real data. This suggests what are the kind of data we need to collect.

5 Conclusions and future work

The principal aspects that NRW wanted to investigate during the study group were how best to facilitate responsible and sustainable use of their sites, and how to gain a better understanding of how users affect sites. In this report, various methods for exploring these aforementioned issues were presented. First, a method for classifying park users was outlined and then examples were given for predicting what type of park users visit different parks. This was then used to inform various agent-based models, which sought to model anti-social behaviours such as littering and disruption to protected areas. Despite there being myriad factors which influence why an individual may commit anti-social acts, these simplified models of human behavior can exhibit some useful insights into these phenomena and will hopefully help curtail these undesired actions in future. Data was then presented which illustrated the correlation between the number of users visiting a site and increased levels of littering, off-road parking and dog fouling. The various costs and benefits relating to the number of users visiting a site were then discussed.

During the course of the study group, several potential directions for future work were identified. One option is to further refine the park scores allocated in Section 2.2 using information such as social media attention and the amount of facilities available at the sites. A greater number of categories of users could also be used, that would more accurately reflect the types of visitors to each site. A user-friendly simulation software that could predict the occurrences of various anti-social behaviours, based on factors such as the number of visitors in each national park and the time of year, could be developed. To do this, it would be beneficial to incorporate more data into the agent-based models about the relative impacts of each anti-social behaviour at different times of year. This data could be collected possibly by surveying park rangers. More realistic agent-based models could then be developed, with domains that accurately reflect real national park topographies.

References

[1] Billari, F. C., Fent, T., Prskawetz, A., and Scheffran, J., eds. *Agent-Based Computational Modelling: Applications in Demography, Social, Economic and Environmental Sciences*. Physica–Verlag, 2006.

[2] Grow, A., and Van Bavel, J., eds. *Agent-Based Modelling in Population Studies: Concepts, Methods, and Applications*. Springer International Publishing, 2017.

[3] Humphrey, W., Dalke, A. and Schulten, K., "VMD: Visual Molecular Dynamics" *Journal of Molecular Graphics* 1996, 14.1, 33–38.

[4] Tverijonaite, E., Ólafsdóttir, R., and Thorsteinsson, T. 2018. Accessibility of pro-

tected areas and visitor behaviour: A case study from Iceland. *Journal of outdoor recreation and tourism*, **24**, 1–10.

[5] Weik, F., Weeber, R., Szuttor, K., Breitsprecher, K., de Graaf, J., Kuron, M., Landsgesell, J., Menke, H., Sean, D., and Holm, C. “ESPResSo 4.0 – an extensible software package for simulating soft matter systems”. *The European Physical Journal Special Topics* 227(14), (1789—1816), 2019.

[6] Wilensky, U. 1999. NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.