This file describes the six datasets, three R scripts, and the Netlogo agent-based model (ABM) code used to run the analysis for “Nonrandom foraging and resource distributions affect the relationships between host density, contact rates, and parasite transmission”. Netlogo was run using version 6.1.1 and R version 4.0.0 was used to analyze all the data.

**Authors**: Zachary Gajewski1, Philip McEmurray2,3, Jeremy Wojdak4, Cari McGregor4, Lily Zeller1, Hannah Cooper1, Lisa K. Belden2, Skylar Hopkins1

1. Department of Applied Ecology, North Carolina State University, Raleigh, NC
2. Department of Biological Sciences, Virginia Tech, Blacksburg, VA
3. Department of Anthropology, Washington University in St. Louis, St. Louis, MO
4. Department of Biology, Radford University, Radford, VA

**Corresponding Author:** Zachary Gajewski, [gajewskizach@gmail.com](mailto:gajewskizach@gmail.com)

**Experimental Study Design**

To understand how resource distributions influenced host contacts were performed an 8x7 factorial microcosm experiment that varied snail density (2,4,6,8,10,12,14, or 16 snails) and periphyton resource distributions (ranging from uniformly distributed to completely clustered). Periphyton was grown on tiles and arranging these tiles along with empty tiles allowed for the creation of different resource distributions in the microcosm. In each of the microcosm experiments, snails were uniquely colored and visually observed for 45 minutes. During that 45 minutes, all snail contacts, durations of those contacts, number of patches visited, and durations of those visits were recorded. Each treatment was replicated twice and all treatments were videoed to be able to check observations. An agent-based model (ABM), described below, was built based on the microcosm experiments and used to examine how host foraging behavior and resource distribution influence host contact rates, and thus, disease dynamics. Using Bayesian methods, we fit Holling type II functions to the ABM and experimental contact data. All data files, R scripts used for analyzes, and the Netlogo file for the ABM used in this study are described below.

**Netlogo File**

**SnailModel.nlogo**

This is an agent-based model based on microcosm experiments conducted with an aquatic snail that links individual-level behavior with population-level processes, focusing on how foraging behaviors and resource distribution affect contact rates between individuals in a population.

The agent-based model consists of patches arranged in a 12 by 12 grid, and each patch has a state variable representing the biomass of the algae on the patch. The model allows for seven different algae distributions, ranging from “0” or completely uniform to increasingly clustered to completely clustered at “6”. Agents, representing snails, move around during the agent-based model simulations and are allowed to forage completely randomly, forage optimally, or somewhere in between those two settings. Foraging behavior is controlled by stochasticity, where 0 indicates optimal foraging and 100 indicates random foraging. Optimal foraging is defined by agents staying on a resource patch until it dropped below a certain biomass and then they would randomly search for a new patch.

Before starting the simulation, parameters specific to patches and agents can be adjusted. The number of agents included in the simulation can be adjusted by using the number-of-snails slider. The snail-feeding-rate is the amount of algae a snail consumes per time step and can be adjusted by typing a new number in the text box. GUD represents giving up density or the amount of resources that are left on a tile when a snail will move off the tile when foraging optimally, and can be adjusted with the slider. Lastly, the agents’ foraging behavior can be set using the stochasticity slider. A stochasticity of 0 indicates that the agents will forage completely optimally, while a stochasticity of 100 means agents will forage completely randomly.

The patch parameters that can be adjusted include initial-algae-coverage and algae-distribution. The algae-distribution box allows the user to enter the aggregation index of 0,1,2,3,4,5, or 6. Resources begin uniformly spaced at an aggregation index of 0 and begin to cluster as aggregation index increases, until they are completely clustered at an aggregation index of 6. The initial-algae-coverage determines the total amount of algae at the beginning of the experiment, and can be adjusted with the slider.

Lastly, simple disease dynamics in the simulation can be adjusted with the disease parameters. The number of starting infected agents can be adjusted with the starting-infections slider. Additionally, how infectious the pathogen is can be increased by increasing the infectiousness slider. Lastly, the recover-time, or time that is takes for a snail to recover from the infection can be adjusted with the slider.

The model outputs the average number of contacts that agents made, the duration of those contacts, the average number of patches visited, the average patch time, the average proportion of resources left, and the average edge time.

When all parameters are set, the simulation is started by using the SETUP button to set up the agents and patches and the GO button starts the simulation.

**R Script Files**

**FittingTypeIIModel\_SnailCounts\_25Apr22.R**

Uses the following data files and R function:

* ExpSnailDurationContactData.csv, this file contains the contact duration data from the microcosm experiments and is used to find the average contact duration used in the experimental Holling Type II model fits.
* SnailExpData\_23Feb22.csv, this file contains the experimental count data used in the analyze and fit with a Holling Type II function.
* OutputABMSimFinal\_15May23.csv, this file contains all of the agent-based model simulation data and settings.

In this script file we use Bayesian methods to fit a Hollings type II functions to the experimental and ABM contact data across host density. One Hollings type II function is fit to the experimental data, while 6 models are fit to the ABM data, one per stochasticity level. This script also uses the ABM file to calculate R0 values. Lastly, this R script creates Figures 2 and 3.

**Exp\_ABM\_PatchComparison.R**

Uses the following data files:

* OutputABMSimFinal\_15May23.csv, this file contains all of the agent-based model simulation data and settings.
* ExpPatchData\_23Jun22.csv, this file contains all of the experimental patch information data, including number of patches visited and patch time.

In this script file we compare experimental and ABM data to ensure that the ABM is producing realistic results. We compare the patches visited and patch duration time. Additionally, this file makes figure 4.

**CleaningGUD\_Data.R**

Uses the following data files:

* GUD\_Data\_7Feb22.csv, this file contains experimental data on the amount of algae left on a tile when a snail left a tile.
* TankInfo\_Data.csv, this file contains general information about the microcosm experiment set ups.

In this script we read in data collected from ImageJ and microcosm experiment videos, which have the area of algae left on a tile when a snail leaves. This script uses that data and calculates giving up density and feeding rates, both are used in the agent-based model. Additionally, this script is used to make supplemental figure 2.

**Data Files**

**OutputABMSimFinal\_15May23.csv**

File Description:

This data file is cleaned output from the agent-based model. We ran an 8x7x6 factorial experiment, each with 10 replicates (N = 3360 simulations). There were eight agent densities (8, 16, 24, 32, 40, 48, 56, and 64), seven resource distributions, and six levels of movement stochasticity ranging from optimal foraging to random foraging (0, 8, 18, 30, 50, and 100). This data is originally given in wide format and was rearranged into long format in this data file. This file contains columns for all the settings used in the ABM and observed variables in the simulation. Each row is a timestep in one of the simulations.

Column Description:

*run:* This column specifies the number of simulation the row is from.There was a total of 3360 simulations run each with 50 rows.

*start\_infection:* This column shows the number of agents that started the simulation as infected. All of our simulations started with 1 agents as infected.

*Feeding:* This shows the feeding rate of the agents and is set at the start of the simulation.

*algae\_conc:* This column shows the starting concentration of algae on an algae patch. This is set at the start of the simulation.

*Snails:* This column shows the agent density in the simulation. This column was 8,16,24,32,40,48,56, or 64. These were four times the amount of densities used in the microcosm experiments because the ABM used four times the amount of patches.

*stoch:* This column indicates the stochasticity of the agents in the simulation. A lower stochasticity represents more random foraging, while, a higher stochasticity represents optimal foraging. In our simulation we used 0, 8, 18, 30, 50 or 100.

*Infect:* This column represents the set infectiousness level, the higher the value the more likely uninfected agents are to become infected when making contact with an infected agent.

*Algae:* This column shows the aggregation index. The aggregation index is a numeric variable describing the aggregation of resources. An aggregation index of 0 indicates that the resources are uniformly spaces and start clustering with increase aggregation index, until they are completely clustered at 6.

*recover:* The recover time is set at the start of the simulation and is the amount of time it takes infected agents to become healthy again.

*GUD:* This column shows the set giving up density, or the amount of resources left on a patch when an agent leaves, when the agent is optimally foraging. This value is set at the start of the simulation.

*resources:* This column shows the average amount of resources left on resource patches.

*edge:* This column shows the average amount of time or ticks that agents spent on edge patches.

*patch\_time:* This column show the average amount of time or ticks that agents spent on a patch.

*Ss:* The number of susceptible agents in the simulation. This column starts as starting density – 1 at the beginning of the simulation.

*Is:* This column shows the number of infected agents. For our simulations this always starts at 1.

*Rs:* This column shows the number of recovered agents and for our simulations starts at 0.

*contacts:* This column gives the average number of contacts in the simulation at the give tick. A contact in the simulation is defined as two agents being on the same patch at the same time.

*total\_contacts:* This column shows the total number of contacts made in the simulation.

*patches:* This column shows the average number of patches that agents visited.

*unique\_contacts:* This column represents the number of unique contacts made in the simulation. If two agents make contact twice this column will only show 1, because the contacts is only unqiue the first time.

*contact\_dur:* This column shows the average contact duration between all the contacts in the simulation.

*ticks:* This column shows the ticks of the simulation, or the time measure. For our purposes, ticks represented minutes. This column always went from 0, start of the simulation, to 50.

**ExpSnailDurationContactData.csv**

File Description:

This data file contains all the contact observations made between snails in the microcosm experiments. Given in this file is also the contact start time, end time, and total duration. Lastly, information about the microcosm that the contact took place in is given. Each row represents a contact made with in one of the microcosm experiments.

Column Descriptions:

*Date:*  Date of that microcosm study took place on

*Trial:* All microcosm experiment were repeated again in another trial. This column is either 1 or 2, for the first round of microcosm experiments or second.

*Tank:* Tank number that the contact took place in.

*Density:* Host density or the number of snails in that tank/trial.

*AggregationIndex*: A numeric variable describing the aggregation of resources. An aggregation index of 0 indicates that the resources are uniformly spaces and start clustering with increase aggregation index, until they are completely clustered at 6.

*Snail1*: One of the snails making contacts. The column is filled with the unique identifier for this snail.

*Snail2:* The second snail making contact. The column is filled with the unique identifier for this snail.

*StartTime:* The time that the contact first started between the snails. The format is give as H:MM:SS (Hour:Minute:Second).

*EndTime:* The time that the contact ended between the two snails. The format is give as H:MM:SS (Hour:Minute:Second).

*Duration:* The total contact duration between the two snails (*Snail1* and *Snail2*), given in seconds.

*Notes:* A column for any notes on about the contact between the two snails.

**SnailExpData\_23Feb22.csv**

File Description:

This data file contains information on each snail in the microcosm experiments, including the number of contacts each snail made. Additionally, information about the microcosm set up, such as *Snaildensity* and *AggregationIndex* are given. Each row represents an individual snail.

Column Description:

*Trial:* All microcosm experiment were repeated again in another trial. This column is either 1 or 2, for the first round of microcosm experiments or second.

*Tank:* Tank number

*ColorID*: A unique identifier for each snail in the tank/trial. Each row is a different snail.

*SnailDensity:* The snail density for the tank/trial. Snail density was either 2,4,6,8,10,12,14, and 16. The density was adjusted down based on dead snail observations in the tank/trial.

*AggregationIndex:* A numeric variable describing the aggregation of resources. An aggregation index of 0 indicates that the resources are uniformly spaces and start clustering with increase aggregation index, until they are completely clustered at 6.

*Length:* The length of the snail

*Count:* The number of contacts that were made by the listed snail (identified in *ColorID*).

**ExpPatchData\_23Jun22.csv**

File Description:

This data file contains the movement information of every snails in the microcosm experiments, with videos. Each snail was watched for the 45minutes and what tile they were on, the duration of their visit, whether the tile was on the edge or contain algae were all reported. Additionally, we noted when snails were floating, as they tended to move around in the microcosm faster. Each row represents a movement made by a snail in the microcosms experiments.

Column Description:

*date:* A column containing the date that the microcosm experiment took place on. Given as Month/Day/Year.

*replicate:* All microcosm experiment were repeated again in another trial. This column is either 1 or 2, for the first round of microcosm experiments or second.

*tank:* Tank number

*snail.density:* The snail density for the tank/trial. Snail density was either 2,4,6,8,10,12,14, and 16. The density was adjusted down based on dead snail observations in the tank/trial.

*snail:* The unique identifier for the snail being observed.

*tile.x:* The x coordinates for the tile that the snail (given in *snail* column) was currently on.

*tile.y:* The y coordinates for the tile that the snail (given in *snail* column) was currently on.

*edge.y.n.:* A yes, y, if the tile that the snail was currently on was on the edge of the microcosm and no, n, if it was not.

*algae.y.n.:* A yes, y, if the tile that the snail was currently had algae on it and no, n, if the tile had no algae on it.

*start.time:* The time when the snail (indicated in *snail* column) first moved on to the current tile.

*end.ime:* The time when the snail (indicated in *snail* column) moved off of the tile it was on to a new tile.

*duration:* The duration that the snail was on the listed tile (in *tile.x* and *tile.y* columns), given in seconds.

*aggregation:* A numeric variable describing the aggregation of resources. An aggregation index of 0 indicates that the resources are uniformly spaces and start clustering with increase aggregation index, until they are completely clustered at 6.

*floating:* A column indicating whether the snail was floating over the indicated tile, marked as a 1, or was on the tile, marked as a 0.

*notes:* Any notes from the observers about the snails movement or tile location.

**GUD\_Data\_7Feb22.csv**

File Description:

This data file contains information about snails moving off algae tiles and the amount of green area or area covered by algae they left behind. This data was collected from the videos recorded of the microcosm experiments. Images from the videos were taken, cropped, and loaded into ImageJ were the total area of the tile the snail moved off of was found and the amount of algae area left was found. Each row represents a movement by a snail. Note that empty tiles or tiles that are not resource tiles contain 0’s and NA’s. Additionally, snails moving off empty times stopped being included in this data set because there was not green area to find in imageJ.

Column Description:

*tank:* Tank number

*replicate:* All microcosm experiment were repeated again in another trial. This column is either 1 or 2, for the first round of microcosm experiments or second.

*snail:* The unique identifier for each snail in the microcosm experiment. The identifier is based off of color that the snail was marked with in the experiment.

*NeverMoved:* This column indicates if the snail moved after the time given in *MoveTime*. A yes is indicated with a 0 and no is a 1. Since the experiment ended at 45minutes, all times of 0:45:00 are given a 1, or they did not move after this time.

*MoveTime*: The time that the snail moved off the tile that it was on. Give as Hours:Minutes:Seconds.

*TotalArea:* The total area of the tile that they snail was leaving. The area was found using ImageJ and a clipped and cropped image from a video of the microcosms.

*GreenArea..in.2.*: The total green area or area covered by algae left on the tile when the snail moved off the tile. The green area was found using a cropped image take from the video of the microcosm experiment and ImageJ.

*Notes:* Any notes or comments about the cropped image, snail movement, total area, or green area were recorded here.

*snail.density:* The snail density for the tank/trial. Snail density was either 2,4,6,8,10,12,14, and 16.

*aggregation.index:* A numeric variable describing the aggregation of resources. An aggregation index of 0 indicates that the resources are uniformly spaces and start clustering with increase aggregation index, until they are completely clustered at 6.

**TankInfo\_Data.csv**

File Description:

This data file contains information about the snail density, resource distribution, and snails in each of the microcosm experiments. Each row is a different microcosm set up.

Column Description:

*replicate:* All microcosm experiment were repeated again in another trial. This column is either 1 or 2, for the first round of microcosm experiments or second.

*tank:* Tank number

*snail density:* The snail density for the tank/trial. Snail density was either 2,4,6,8,10,12,14, and 16.

*aggregation:* A numeric variable describing the aggregation of resources. An aggregation index of 0 indicates that the resources are uniformly spaces and start clustering with increase aggregation index, until they are completely clustered at 6.

*snail color:* The unique identifier for each snail in the microcosm experiment. The identifier is based off of color that the snail was marked with in the experiment.