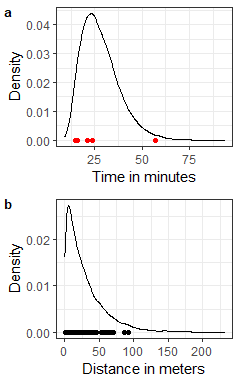
Figures and Tables

Table of Contents

## Methods

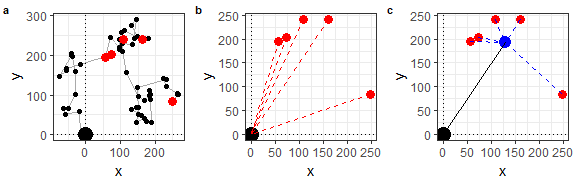
#### Figure 1.

Density distributions from which we sample **A.** gut retention time (GRT) and **B.** movement distance (MD). Red dots represent an example for one simulation run, where we sample five gut retention times. The maximum GRT for that simulation run determines the number of movement distances sampled in that simulation run, represented in the figure by the black dots.



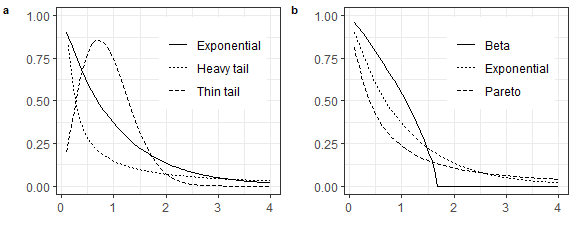
#### Figure 2.

Example of one simulation run and variables measured for each run. **A.** Shows the animal trajectory followed after starting at the origin (big circle at location 0,0). The animal moves every minute following the movement distances (MD) sampled from the distribution as show in example of Figure 1. Every small black dot represents the animal’s location at every minute interval. Seeds get dropped every time the simulation reaches one of the sampled gut retention times (GRT), at the animal’s location in that time (red dots). **B.** Seed dispersal calculations as euclidean distances from the parent plant located at the origin (black circle at 0,0) to each of the five seeds for the simulation run (red dots). **C.** Calculation of seed dispersion measures and average seed dispersal. The average seed location for the simulation run is shown by the blue circle. The distance from each seed to the average location is used to calculate seed dispersion as a measure of seed aggregation, dashed blue lines. Average seed dispersal is calculated as the distance from the parent plant to the average seed location, shown as the black line connecting the black circle to the blue dot.



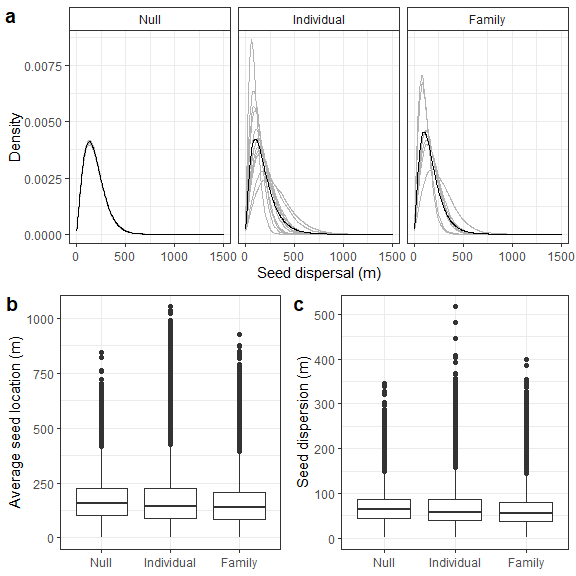
#### Figure 3.

Variation in Weibull and Generalized Pareto distributions determined by the value of the shape parameter. **A.** In the case of the Weibull distribution, the scale parameter is defined as , and the shape parameters are defined as follows: solid line shows a heavy tail, the dashed line shows a thin tail, and describes an exponential tail with the dotted line. This follows the assumption that a shape parameter describes a fat tail, a thin tail, and a tail following an exponential distribution. **B.** In the case of the Generalized Pareto distribution, a shape parameter describes an exponential distribution function, shape describes a heavy tail, and a tail with Beta distribution function, bounded at an upper value as a function of the threshold and scale parameters. In the second panel, the scale parameter is set to the solid line shows a heavy tail, dashed line shows a Beta tail, and describes an exponential decay with the dotted line.



## Results

#### Figure 4. Seed dispersal distances and aggregation metrics for simulated seeds in the three models considering variation in animal movement rates. **A.** Density kernels for frugivore-generated seed dispersal distances for each seed simulated under the three models of variation in animal movement rates. Each individual line represents the dispersal kernel generated by each individual animal. The null model assumes the same movement rate for all animals, whereas individual and family models use a movement rate for each animal id, or family group. **B.** Box plot comparing the distances from the parent tree at the origin to the mean seed location in each simulation run among the three simulation models. **C.** Box plot comparing the seed aggregation metric, seed dispersion, for each simulation run between the three simulation models. The mean seed location is used to calculate dispersion, as the average distance of each seed to the mean seed location, as shown in Figure 1. *Is there more comparison I should mention here? Like, how the larger values are more abundant in the individual model?*



#### Table 1. Summary table of seed dispersal metrics for each of the simulation models. Seed dispersal and seed dispersion metrics represent the average value for each of the models, with standard errors in parentheses. Kurtosis is calculated for each of the three simulation models as described in the main text. Maximum seed dispersal distance represents the longest dispersal distance out of all seeds dispersed for each of the simulation models. Long-distance dispersal events are calculated as the percentage of seed dispersal distances greater than 500m for each simulated individual, with standard deviations calculated between the individuals in each simulation model.

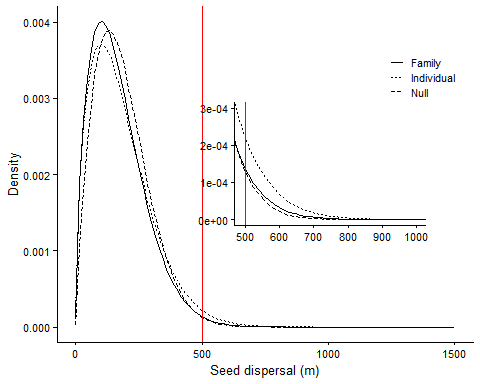
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Mean.dispersal\_sd | Seed.dispersion\_sd | kurtosis | Max\_dispersal | LDD |
| Null | 183.5 (110) | 68.35 (33) | 4.38 | 1248 | 5.19 (0.32)% |
| Individual | 181 (130) | 67.26 (41) | 6.77 | 1501 | 14 (22)% |
| Family | 169.2 (110) | 62.79 (36) | 6.24 | 1371 | 7.68 (13)% |

### Weibull seed dispersal kernel

#### Table 2. Parameter estimates for seed dispersal kernels on each of the simulation models using a Weibull distribution to characterize the kernel. Parameter values are reported with their standard deviations in parentheses.

|  |  |
| --- | --- |
| Weibull\_Shape | Weibull\_Scale |
| 1.801 (0.0018) | 206.7 (0.16) |
| 1.525 (0.0015) | 201.8 (0.18) |
| 1.586 (0.002) | 189.2 (0.21) |

#### Figure 5. Kernels produced from Weibull functions using the estimated parameters for each of the simulated models. Inset shows a zoom to the tail of the distribution emphasizing the variation in long-distance dispersal events described by each of the models. Red line is used for reference as the original 500m threshold to describe long-distance dispersal events in previous studies. We observe fatter tails for the model with individual variation in animal movement, followed by family level variation, and a thinner tail for the null simulation model with no variation in animal movement.

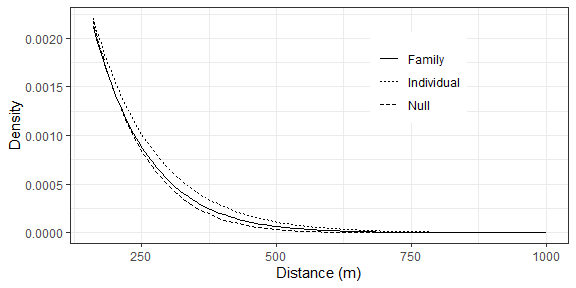


## Generalized Pareto tail distribution for seed dispersal

#### Table 3. Threshold values and parameter estimates for generalized Pareto distribution fits, and associated standard error for parameter estimates.

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Scale | Shape | Threshold |
| Null | 106.6 ± 0.15 | -0.09877 ± 2e-08 | 177 |
| Individual | 118.6 ± 0.38 | -0.02192 ± 0.0022 | 212 |
| Family | 106.7 ± 0.43 | -0.03864 ± 0.0027 | 194 |

#### Figure 6.

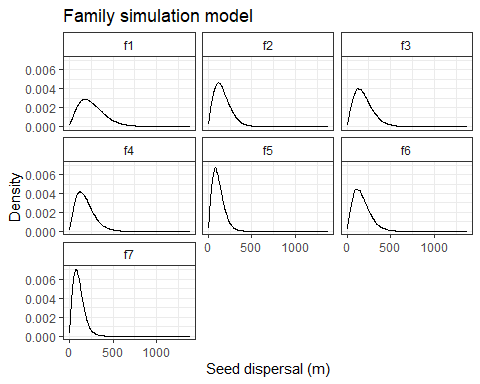
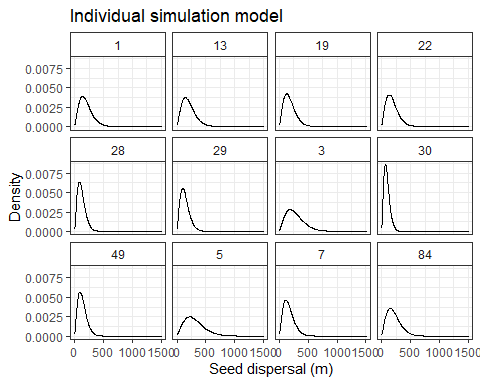
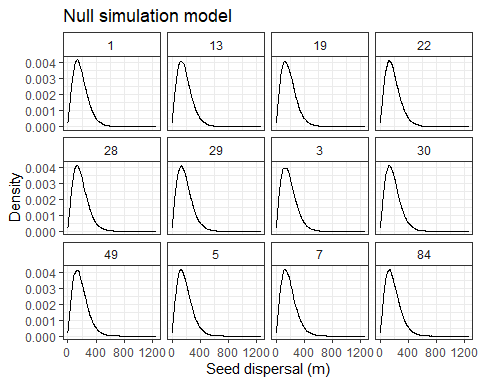


#### Table 4. Probability of long-distance dispersal events for the three models following a generalized Pareto distribution.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 250 | 500 | 1000 | 1250 | 1500 | 1750 | 2000 |
| Null | 0.4921557 | 0.0273207 | 5.00e-07 | 0.0e+00 | 0.0e+00 | 0e+00 | 0 |
| Individual | 0.7249547 | 0.0823946 | 7.59e-04 | 6.0e-05 | 4.1e-06 | 2e-07 | 0 |
| Family | 0.5884545 | 0.0478492 | 1.32e-04 | 3.8e-06 | 1.0e-07 | 0e+00 | 0 |

## Supplementary Figures and tables

### Variation in kernels produced by individual birds in each of the three models



### Threshold diagnostic plots

