Individual Variation in Animal Movement and Effects on Seed Dispersal Distance

Seed dispersal is a fundamental process in plant ecology and due to its high importance in the ecosystem, various modelling approaches have been developed to understand seed movement and dispersal mechanisms. In the case of frugivore-generated seed dispersal patterns, a large influence is observed by animal movement patterns, and intraspecific variation in animal movement and behaviors can result in interesting seed dispersal patterns. In particular, long-distance seed dispersal (LDD) plays a significant role in determining a population’s genetic diversity and range expansion and animal-mediated seed dispersal is a key component of LDD events, as frugivores can travel long distances and transport seeds with them. In this study we seek to understand the implications of individual variation in animal movement and how these can impact estimates of long distance seed dispersal. We apply these methods to empirical data sets that contain locations for an araçari, *Pteroglossus pluricintus*, one of the primary frugivores for the Amazonian canopy tree *Virola flexuosa*. We combine animal movement data and gut retention time to simulate *Virola* seed dispersal with an individual-based model on a homogeneous landscape. This approach showed that there is significant variation in individual araçari movement patterns which directly influence the estimated number of long-distance dispersal events for *Virola* seeds and estimates for seed dispersal kernel functions. Additionally, we include an approach using statistics of extremes to characterize the tail behavior of generated seed dispersal kernels under varying levels of intraspecific variation of animal movement. We found that models that ignore this underlying variation in individual animal movement, underestimate the percentage of long-distance seed dispersal events and maximum seed dispersal distances.

# Introduction

A main goal of ecologists and evolutionary biologists is to document and predict the structure and dynamics of ecological systems. Although most of our current understanding of ecological systems is based on the average patterns observed in nature, understanding variability and its consequences in ecological processes has recently raised the interests of researchers (Benedetti-Cecchi 2003; Inouye 2005). Variation is now considered a property of ecological systems that might contain as much ecological information as do averages and regularities in nature (Benedetti-Cecchi 2003; Violle et al. 2012). Individual variation, i.e. the possibility for individuals within a population to present different ecological strategies or traits (Violle et al. 2012), has been widely discussed in ecology and evolutionary biology. In evolutionary biology the acknowledgement that individuals with different traits might vary in their fitness and survival under natural selection pressures dates back to Darwin’s seminal ideas (Darwin and others 1859). Today, it is widely recognized that individual variation plays a key role in the evolution of species and clades, and the relevance of individual variation in the performance of populations and species facing current threats (e.g. climate change, fragmentation and habitat loss) is a growing field of research (Forsman and Wennersten 2016). In ecology, individual variation is often mentioned as a mechanism for populations to reduce competition pressure (Bolnick et al. 2010; Araújo, Bolnick, and Layman 2011). Based on the niche complementarity principle, phenotypically different individuals will compete less strongly than more similar individuals (Bolnick et al. 2011), thus individual variation would act as a release from intra- and interspecific competition at a local scale. Despite the well recognized importance of individual variation in ecological and evolutionary processes, ecological theory has emphasized interspecific variation (Violle et al. 2012). Not surprisingly, the role of individual variability on the outcome of species interactions has also been poorly studied; most models of species interactions assume that all conspecific individuals are equivalent and interchangeable (Bolnick et al. 2011). Individuals, however, differ in several aspects that will affect their interactions with other species. This variability makes individuals non-interchangeable from the perspective of interacting patterns (Bolnick et al. 2003; González-Varo and Traveset 2016). The consequences of individual variability in interspecies interactions can even have cascading effects, going beyond from the species-species level to the community and ecosystem level (Post et al. 2008; Wolf and Weissing 2012).

Seed dispersal mutualisms may be a particularly attractive system to study the impacts of incorporating individual variation to understanding the importance of such variation to dispersal outcomes. Seed dispersal is one of the most critical stages in plant life history – it results in a spatial pattern of seed deposition which acts as a template that will define the distribution of plants and, consequently, community structure (Howe and Smallwood 1982; Nathan and Muller-Landau 2000; Russo, Portnoy, and Augspurger 2006). In seed dispersal mutualisms, animals provide the dispersal services while rewarding from fruit resources they consume. Plants produce fleshy fruits to attract and reward mutualist animals and rely on the behavior of the disperser to transport their propagules (Nathan and Muller-Landau 2000). The spatial patterns of seed deposited by disperser vectors will determine the probability of seed survival and recruitment, acting upon post-dispersal processes such as density-dependent survival and colonization (Howe and Miriti 2004). Therefore, at a local scale, seed dispersers will determine whether plant species escape from density-dependence processes (Comita et al. 2014), whereas at a larger scale, they can determine how quickly plants can cope with habitat and climate shifts (Ibáñez et al. 2006; Russo, Portnoy, and Augspurger 2006).

From the plant perspective, not all seed dispersers species provide with the same seed dispersal service, and often relatively few disperser species can have disproportionate effects on seed dispersal (Schupp 1993; Jordano et al. 2007; Schupp, Jordano, and Gómez 2010; Loayza and Rios 2014). Frugivores might differ on their selection of plants/fruits for consumption, in their treatment of seed while ingesting and manipulating the fruits, in their processing of seeds if ingested and in the distance where they will discard the seeds (Schupp 1993; Schupp, Jordano, and Gómez 2010; Côrtes and Uriarte 2013). Each one of these stages of the seed dispersal process will have strong consequences on the spatial seed shadow provided by each seed dispersal vector (Jordano et al. 2007; Carlo and Tewksbury 2014). Besides inter-specific differences among dispersal vectors, there is a recent recognition of the potential role of differences among individuals in the resulting seed dispersal they provide (Bolnick et al. 2003; González-Varo and Traveset 2016). Including individual variation in traits (morphological, behavioral and physiological) in seed dispersal studies, can provide new insights into the relative role of different dispersal vectors on seed deposition and ultimately, vegetation structure (González-Varo and Traveset 2016). In a recent review, Zwolak (Zwolak 2018) summarized the types of intraspecific variation in seed-dispersing animals that might affect the resulting seed dispersal service individuals provide, highlighting differences in sex, size or age (ontogenetic shifts), individual specialization and behavioral syndromes as the most important variation types. Incorporating these intrinsic traits into seed dispersal models may lead to more mechanistic understanding of seed dispersal, allow to build more accurate predictive frameworks of dispersal outcomes (Russo, Portnoy, and Augspurger 2006) and to identify links between characteristics of dispersal agents and the seed dispersal they provide (Zwolak 2018).

In this paper, we focus on the implications of intraspecific variation in animal movement for seed dispersal distances and seed aggregation across the landscape. To understand the consequences of this variation in movement, we developed a spatially-explicit individual-based model of the many-banded aracari, *Pteroglossus pluricinctus*, one of the primary frugivores of the Amazonian canopy tree *Virola flexuosa*. Also, previous studies focused on this system have shown that *P. pluricintus* is able to disperse seeds across long distances (K. M. Holbrook and Loiselle 2007; Holbrook and Loiselle 2009; Holbrook 2011), an important for plant population. Long distance dispersal (LDD) events are crucial to reaching suitable sites to germinate and establish and to colonize new habitats (Nathan 2006). Furthermore, rare events (often overlooked) are critical for dispersal and might have a large effect on resulting plant demography (Loayza and Rios 2014). Our first objective in this work was to study the differences of simulated seed shadows between models with and without underlying variability in animal movement. The second objective focused on quantifying the differences of LDD events reflected between seed shadows. Lastly, our third objective focused on the spatial spread between seeds, and how individual variation in animal movement affects seed aggregation. Identifying the implications of individual variation in animal movement over seed dispersal estimates can help us improve future models and understand the effects of frugivore traits on plant population dynamics.

# Methods

## Simulation description

To study the effects of individual variation in animal movement over final dispersal distances of foraged seeds, we developed a spatially-explicit individual-based model in a homogenous landscape consisting of a single source tree. We incorporated animal movement focusing on two characteristics, a movement distance (MD) sampled from a probability density distribution, and a movement angle (MA). In addition to this, we included gut retention time (GRT), the time that ingested seeds stay inside the frugivore until they are dropped. The animal was allowed to move freely within the landscape, and GRT determined when the animal would drop a seed. Once a seed was dropped, its location was recorded and seed dispersal distance was estimated as the distance from the origin to the seed’s location. The average seed dispersal distance for each simulation run was also estimated and used to for seed dispersion, calculated as the mean distance of each seed to the average location of all seeds in the run (Jones et al. 2017). For each simulation run, the simulation started at the source tree, an animal received a specific number of seeds and each seed was assigned a specific GRT sampled from a gamma distribution (Morales and Carlo 2006). The simulation was based on one minute time steps, where at each time step a movement distance (MD) and movement angle (MA) would determine the path for the animal to follow. For each seed, the location where it would get dropped was based on the animal’s location at the time it reached the GRT for each of the seeds. The total simulation time for each run was determined by the largest GRT sampled for that specific simulation run. Each simulation run focused on one individual animal dispersing seeds from one focus tree. Once all seeds in that run were dropped by the animal, the simulation run ended.

## Study design

To understand the consequences of individual variation in animal movement and how these influence seed dispersal distances, we simulated three different scenarios by varying the movement of individual animals. We used our first scenario as our null model, in which all the individuals from the simulation had the same average movement rate (), used as the parameter in the probability density function, and thus their movement distances per unit of time were sampled from the same probability distribution (). We chose this as our null model as it is the approach most commonly used in frugivore-generated seed dispersal kernels, where the movement rate for the frugivore population is determined and used either as a constant movement speed or to sample from a single distribution (EDIT citation, Will Tackenberg, 2008, constant rate, Jones 2017, exponential with different values for theoretical model, Levey bolker 2 papers lognormal). Our second scenario included individual variation in animal movement by incorporating different movement rates for each individual (), which meant that movement distances would be sampled from the same overall probability distribution, but with different parameters (). **Double check if these are the Erlang and hyperexponential distributions, not sure it is worth it to mention here though**. Considering that *Pteroglossus pluricinctus*, the focus frugivore in this paper, by belonging to social groups their movement patterns might be correlated, our third scenario ran simulations at the scale of family group variation in animal movement by assigning a different movement rate to each social group in the simulations, where the movement rate where considered as the average movement in meters per unit of time for all the tracking sessions for all individuals belonging to that social group.

## Parameterization

We focused our study on the many-banded araçari (*Pterglossus pluricinctus*), a small toucan, and its role as a frugivore of the *Virola flexuosa* tree. Previous studies by Holbrook (2011) collected radiotracking information for various dispersers over a period of four years, from 2001 to 2005, in the Ecuadorian Amazon rainforest. The methodology used at the time consisted of capturing and radio-tagging individuals from various toucan species, including *Pterglossus pluricinctus* and two larger Ramphastids, although data on the latter was scarce. Tracking periods lasted between four and six daylight hours, alternating morning and afternoons, attempting to record bird locations every 15 minutes, a time interval shown to represent the minimum seed retention time for *Virola flexuosa* seeds (K. M. Holbrook and Loiselle 2007). Out of the data collected, and following Holbrook (2011), we selected data from individuals with at least 40 recorded point locations. Even though location recordings were attempted every 15 minutes, this was not always possible, due to the individual being too far out from range or due to canopy cover, therefore point locations data is available every 15 minutes or multiples of 15 minutes, with the majority of point location recordings between 15 and 30 minutes (Holbrook 2011). Further details on field methods can be found in Holbrook (2011).

To take advantage of all the locations recorded for series of successive points, despite this variation in duration of time intervals, we calculated rates of movement, as the average number of meters moved per minute over the entire tracking period for that individual bird. These movement rates were later used in our simulation models as the parameter to describe the probability distribution of movement distances at each one minute time step for each simulation run. In the case of our null model, the movement rate was averaged across all individuals for all the tracking periods, and at the social group level we took the average number of meters moved per minute over the tracking periods for all individuals belonging to that group, with a total of 7 independent movement rates.

Parameter values for gut retention time were also based on empirically collected data from previous studies (K. M. Holbrook and Loiselle 2007), were passage trials were carried out with *Pteroglossus pluricinctus* individuals. However, these passage trials were not performed on every individual captured and some of the trials were also performed with captive birds, thus we could not explore individual variation in gut retention times for the individuals in the movement data set. These trials showed that the average gut retention time for *Pterglossus pluricinctus* was 28 minutes (K. M. Holbrook and Loiselle 2007), and presented a maximum retention time of over 100 minutes. The distribution of gut retention times is characteristically fat-tailed, thus we used a gamma distribution (shape = 4, scale = 5) with the appropriate shift to match our average retention time of 28 minutes (Morales and Carlo 2006; Levey et al. 2005), and used this distribution to sample gut retention times for each seed in our simulation runs.

*Include table in the supplement with the movement rates*

## Model processes

Each simulation run in our models consisted of one animal in a homogeneous landscape starting at the focus tree located at coordinates (0,0). We assume the landscape is in meters, as the sampling of movement distances for animal movement comes from estimates of meters moved per minute, the movement rates. Informed from field observations, *Pteroglossus pluricinctus* have a mean visit length of 4.0 minutes for each fruiting tree, consuming between 2 and 5 seeds during each visit, and not visiting another fruiting tree immediately after feeding (Holbrook’s dissertation and (Kimberly Mae Holbrook 2007)). Due to this underlying behavior, we decided to focus our simulations specifically to the events associated to feeding from one focus tree, and the animal movement occurring after foraging and before visiting another fruiting tree. Since one of the objectives for this paper is to explore the occurrence of long-distance dispersal events, we did not set boundaries for the landscape, as using the boundaries associated to the radiotracking sampling would limit animal movement and seed dispersal distances to only the scale sampled by previous researchers. *Pteroglossus pluricinctus* has shown maximum travel distances exceeding 3500m in a single 30 minute tracking interval, thus showing the potential to disperse *Virola* seeds at long ranges (Holbrook 2011).

At the beginning of each simulation run, the animal consumed five seeds at the focus tree location, and each of these seeds was assigned a gut retention time sampled from a gamma distribution (shape=4, scale=5, shift=8, Figure 1a). Once the animal in the simulation consumed the seeds, a movement distance was sampled from the assigned exponential distribution (Figure 1b), which varied depending on the underlying model such as the null, individual or family group model. A random direction from to degrees was used to determine the angle of movement, which with the movement distance would determine the animal’s position for the next time step, repeating this process for every time step and thus following an uncorrelated random walk with no directional tendency (*Edit references* Turchin 1998). Once the simulation run’s time matched the gut retention time for a seed, that seed would get dropped at the animal’s location at that specific time point, thus allowing us to record seed location in the landscape (Figure 2a). For each of our models we ran 10,000 simulation runs per individual or family group, depending on the model, and collected information on animal and seed locations at every time step.

## Seed dispersal and aggregation metrics

We calculated seed dispersal distance as the euclidean distance of each seed to the parent plant. Given that our models only considered one parent plant located at the origin per simulation, we calculated seed dispersal distance (DD) as the distance from each seed’s location to the origin as follows:

where is the dispersal distance for seed in the simulation run and and are its xy coordinates in the landscape. We also used an aggregation metric to determine how evenly seeds would be dispersed across the landscape, and calculated seed dispersion (SD) as the average distance of each seed to the mean seed location in each simulation run (following methods in (Jones et al. 2017)):

where is the number of seeds for each simulation run (5), and and is the mean seed location in the simulation run. We obtained measures of seed dispersion for each simulation run, thus we calculated the average seed dispersion for each model as , where is the total number of simulation runs for the model being considered.The information on seed dispersal distance was used for the toucan-generated seed dispersal kernels for each one of our models (null, individual, and family group), and these seed dispersal distances were also classified as long distance dispersal events if they exceeded 500m (K. M. Holbrook and Loiselle 2007). The proportion of long-distance dispersal events, maximum dispersal distance, and average seed dispersion were used to compare seed dispersal between our three models of individual variation (Table1).

## Seed dispersal kernels

Seed dispersal kernels are functions used to describe the probability of a seed being dispersed or deposited at a specific distance away from its parent plant (Nathan and Muller-Landau 2000). The resulting data from our three simulation models provided a dispersal distance for each simulated seed, and thus we used this information to describe the seed dispersal kernels for each model (null, individual, and family). Dispersal kernels in nature tend to be leptokurtic, with a peak near the origin and long tails, therefore kernel shape can be summarized by its kurtosis (Morales and Carlo 2006). We quantified the dispersal kernels produced in our models by the sample statistics of mean and kurtosis to describe the overall shape and tail of distribution of seed dispersal distances. In addition to this, we also fit a Weibull distribution via maximum likelihood to obtain estimates of the shape and scale parameters in this distribution. This probability distribution is commonly used in dispersal ecology and has been shown to be flexible enough to accommodate variability in the tails and provide informative parameters associated to the mean dispersal distance and fatness of the tail (Morales and Carlo 2006). We fit the data associated to seed dispersal distances produced by each of the three models via Maximum Likelihood using the Weibull density function associated with the package ‘fitdistrplus’(Delignette-Muller and Dutang 2015):

Where the shape parameter determines the tail of the distribution, and is the scale parameter. It is worth mentioning that the standard parameterization of the Weibull distribution in ‘R’ is different from the parameterization used in Morales and Carlo (2006), where the shape parameter remains the same, but the scale parameter is defined as , giving a density distribution function of . Since the shape parameter remains the same, we can note that for the distribution shows a tail with exponential decay, with values of the tail shows fast-decay or thin tail, and when we can see a fat-tailed distribution as shown in (Morales and Carlo 2006) (Fig 3a.) We visually assessed the fit of the distributions via qqplots and used the Kolmogorov-Smirnov statistic to assess goodness of fit.

Given our interest in the tail-end of the seed dispersal kernels and our focus on the occurrence of long-distance events, we decided to use an extreme value theory approach and fit a Generalized Pareto distribution to the dispersal distances above the 500m threshold established. Statistics of extremes are useful to understand the stochastic behavior of rare extreme events (Coles et al. 2001) and have previously been used in ecological scenarios (Gaines and Denny 1993; Katz, Brush, and Parlange 2005) and evolutionary genetics (Beisel et al. 2007; Joyce et al. 2008). More recently, statistics of extremes have been used to model extended dispersal kernels of seeds and pollen (García and Borda-de-Água 2017) and in particular have been proposed to understand long-distance dispersal events (Rogers et al. 2019). In our particular case, we were interested to evaluate how an approach using statistics of extremes could help us fit the tail of the frugivore-generated seed dispersal kernels for each of the models we simulated. We used a peak over threshold (POT) approach, in which we filtered the dispersal distances above a given threshold and fit a Generalized Pareto (GP) distribution to those values. This differs from a block maxima (BM) approach, where the maximum value is selected for each sampling unit and a Generalized Extreme Value (GEV) distribution is used, therefore the choice in the approach is determined by the structure in sampling for a given study (further details can be found in (Coles et al. 2001; García and Borda-de-Água 2017)). The family of generalized Pareto distributions is (Coles et al. 2001):

defined on and , where

An approximation between the GEV and GP distributions occurs where the values over a threshold excess (location, ) have a corresponding approximate GP distribution (Coles et al. 2001). The parameters in the GP distribution of threshold excesses are determined by the values in the corresponding GEV distribution of a random variable (scale , location ). In both distributions, the shape parameter is dominant to determine the tail of the distribution. In the generalized Pareto distribution, a shape parameter the tail has an upper bound at following a Beta distribution function, whereas a shape parameter follows a heavy tail with no upper limit. In the special case of , the GP distribution function approximates to an exponential distribution with a parameter of (Figure 3b.) (Coles et al. 2001). The focus on these shape parameters, both in the case of the Generalized Pareto and the Weibull distribution, is due to the interest in long-distance dispersal events, and using these parameters as a comparisson between dispersal kernels allows us to compare which kernels have higher probabilities for long-distance dispersal events.

Initially, we defined long-distance dispersal events defined as any dispersal event over 500m from the parent tree in order to compare our results to previous studies in this same ecological system (Holbrook and Loiselle 2009; Holbrook 2011; K. M. Holbrook and Loiselle 2007). We used this threshold to assess the percentage of LDD events reported in table 1. However, when selecting a threshold for a POT approach, using a generalized Pareto distribution, we can use diagnostic plots by fitting the data to a sequence of thresholds, with the goal of finding the lowest threshold that provides similar parameter values to any other higher threshold. We performed these analysis and we report our parameter estimates for the generalized Pareto and expected conditional probabilities of long-distance dispersal events using the thresholds from diagnostics plots. We fit Generalized Pareto distributions to the data via maximum likelihood using the package extRemes (Gilleland and Katz 2016) in R (R Core Team 2020).

# Results

## Simulation of seed dispersal distances under the three movement models

We generated seed dispersal distances from combining animal movement and seed gut retention time. The three models we used to simulate the dispersal distances differ in their approach to characterize variation between the animals dispersing seeds. In the case of our null model, where each animal’s step length is sampled from an exponential distribution with the same rate parameter, we would expect to find no differences between the dispersal kernels generated by each of the animals. On the contrary, for the case of simulation models that include individual and family level variation in animal movement, we expected the seed dispersal kernels generated by each individual to show variation. Indeed, Figure 4A shows that individual and family level simulation models show different seed dispersal kernels for each individual bird, whereas in the null model, individual seed dispersal kernels actually overlap with the average dispersal kernel (black line). We generated the average seed dispersal kernels (black lines in Figure 4A) for each of the simulation models by taking all of the seed dispersal distances from the simulation. We did this as in the field we can perform maternity analysis to know seed origin, but it is more complicated to know which specific animal dispersed each seed. Therefore, the average seed dispersal kernels would better reflect the type of data we would collect from field surveys of seed dispersal. From now on, we refer to each of these average seed dispersal kernels as the kernel generated by each of the three simulation models: null kernel, individual kernel, and family kernel.

Average seed locations provided a metric for seed dispersal in each simulation run, and seed dispersion shows how far apart each of the seeds are from each other in every simulation run. We show average seed locations and seed dispersion for the three models (Figure 4B-C) and although the average measures for these are comparable between models, location and dispersion have a greater number of outliers in the indvidual and family models. This suggests that although the average measures are similar between models, the variation between them is different, with the individual and family models showing both a greater maximum dispersal distance, and a higher percentage of long-distance dispersal events (Table 1.) Interestingly, the higher percentage of LDD events in the individual kernels also has a higher standard deviation, a result of the variation in the dispersal kernels produced by each bird (Figure 4A), where some birds or social groups where characteristically long-dispersers versus individuals with shorter dispersal distances and even one individual for which simulated seed dispersal distances were all below 500m (Supplementary information).

## Kernel functions for frugivore-generated dispersal distances

For the seed dispersal data generated from each of the three simulation models, we fit a Weibull distribution function and estimated the parameters for each of the models (Table 2). The shape parameters provided information on the fatness of the tails, as previously mentioned in the methods section above. Although all three models have shape parameters , describing thin tails, we find that there are differences between the three simulation models, with the null model having the largest shape parameter value and thus showing the thinnest tail among the three models (Figure 5). The individual model shows the smallest shape parameter value, and thus describing the dispersal kernel with the heaviest tail among the three models.

Given our interest in long-distance dispersal events and the potential to use statistics of extremes in seed dispersal ecology, we fit the data of the three simulated models to a generalized Pareto distribution following a peak-over-threshold approach. Using this approach, we also performed analysis with threshold diagnostic plots and mean residual life plots (Supplementary information) to estimate the best thresholds for our data. These type of diagnostic approach helps us identify the distances above which we can consider seed dispersal events as long-distance dispersal. We report these threshold values, and parameter estimates (Table 3.) showing that all three dispersal kernels have shape parameter values thus corresponding to thin tails, following a Beta distribution with an upper limit. Although the parameter estimates are close and they show thin tail for all three models, the generalized Pareto fit shows a similar trend as that in the Weibull kernels, with the individual model showing a fatter tail, followed by the family model, and the null model with the thinnest tail (Figure 6). Using these parameter values, we calculated the probabilities of long-distance dispersal events at different distances from the parent tree (Table 4), ranging from 250 to 2000m over the threshold. Although these are very low probabilities, it is worth noting that the parameters associated with the null model reach a zero probability of dispersal for seeds beyond 1250m conditional on the null threshold, whereas the individual level model has probabilites greater than zero for dispersal events of 1750m beyond the threshold (). All three models had a probability of zero for dispersal events greater than 2000 meters beyond the LDD threshold.

# Discussion

Frugivore-generated seed dispersal kernels have been used to estimate seed dispersal distances in multiple studies (Jones et al. 2017; Levey et al. 2005; Levey, Tewksbury, and Bolker 2008; Russo, Portnoy, and Augspurger 2006; Will and Tackenberg 2008), often with the goal of understanding how incorporating animal movement and behavior will impact seed dispersal distances. Usually, the focus tends to be on seed dispersal and how spatial patterns of seed aggregation can change in response to animal behavior or landscape heterogeneity. However, fewer studies have emphasized the role of intraspecific variation in animal movement over generated seed dispersal patterns, or even the relative importance that different disperser species may have in carrying seeds and contributing to seed dispersal (Rehm et al. 2018; Zwolak 2018). Extending on the concept of the total dispersal kernel, where we seek to understand and characterize all the different components contributing to a specific parent tree or plant species dispersal kernel, greater emphasis has been given to the role of disperser species and abiotic processes that may influence a total dispersal kernel (Rogers et al. 2019). However, we show that individual variation in animal movement within a single species can also influence a plant’s seed dispersal kernel, specifically in a system of toucan-generated seed dispersal kernels. In our particular case, the variation introduced in animal movement patterns included variation at the level of individual animals and variation between family groups. Given that *Pteroglossus pluricinctus* maintain cooperative social groups and previous studies showed overlapping home ranges between individuals belonging to the same social group (Holbrook 2011), we wanted to explore how variation in movement between social groups could alter the simulated dispersal kernels. We found that incorporating heterogeneity in animal movement produced seed dispersal kernels with fatter tails, longer dispersal distances, and a higher percentage of long-distance dispersal events compared to null models where no animal movement heterogeneity is included.

Even though long-distance dispersal events are rare or infrequent, they have a disproportionately large effect on gene flow and the genetic pool of populations (Jordano 2017). In the case of well mixed populations, long-distance dispersal can cause random genetic loss through drift or have the opposite effect and maintain high genetic variance in populations that initially drifted (Bohrer, Nathan, and Volis 2005). This is important as in the long-term, these genetic consequences of dispersal can have significant effects for the survival of populations. From a metapopulation perspective, LDD events can have significant consequences for species persistence by enhancing genetic variability and as mechanism for survival in spatially and temporally heterogenous environments. In particular, long-distance dispersal events have a higher probability of reaching isolated populations and therefore establishing a connection and maintaining unrelated populations. It is of special interest to focus in highly heterogenous habitats with rapid change in spatial structure where local extinctions are high, since long-distance dispersal can allow persistence of a metapopulation with immigration and emigration based on long distance dispersal. From an evolutionary perspective, long-distance dispersal can allow for a species to colonize distant habitats and expand its range (Jordano 2017). We have shown that individual differences in frugivore movement can have significant consequences on the number of long-distance seed dispersal events and it remains to be explored how these individual differences can influence population dynamics in the long term and at larger spatial scales. We also show that the approach using statistics of extremes allows us to estimate the probability of dispersal for distances beyond those directly observed in our simulated dispersal kernels, and that simulation models including individual variation in animal movement have higher probabilities for these events. Our simulated dispersal distances had a maximum of over 1500 meters under the model with individual variation, an estimate that is not unreasonable for this system, and that actually underestimates previous estimates (K. M. Holbrook and Loiselle 2007). Parameters estimated for generalized Pareto functions fitted to our frugivore-generated seed dispersal distances showed shape parameters pointing towards thin tails constrained by upper bounds. Although we observed differences between the kernels fitted to the three simulation models, all of them still presented parameter values smaller than zero, and thus showing thin tails. Nonetheless, using an approach of statistics of extremes proved to be valuable in selecting seed dispersal thresholds for considering long-distance dispersal events and carrying out analyses focused on the tail of dispersal kernels, whether that is for seeds or pollen (García and Borda-de-Água 2017).

The thing tails estimated by seed dispersal kernel functions can be in part due to the fact that we used a very simple animal movement model incorporating only step lengths and turning angles. Step lengths for all simulation models we randomly drawn from exponential distributions, with varying rate parameters depending on the model. Using an exponential distribution and a simple movement model such as this one can certainly limit the extent to which an animal can move and thus have an upper bound on seed dispersal distances, something we observe not only on fits with the generalized Pareto, but also with the Weibull dispersal kernels. Previous studies looking at a mechanistic understanding of seed dispersal by frugivores have modeled animal movement with approaches similar to this one, varying step lengths to constants or switching to a lognormal distribution (Jones et al. 2017; Levey et al. 2005; Will and Tackenberg 2008), or other incorporating more complex animal movement models with diffusion processes(Morales and Carlo 2006; Pegman, Perry, and Clout 2017). There has been emphasis on adding different animal behaviors into mechanistic models of seed dispersal, such as perching or sleeping(Jones et al. 2017; Russo, Portnoy, and Augspurger 2006), however a focus on individual animal movement is still not coomon. Incorporating individual variation in animal movement models is an active area in the movement ecology field, with various approaches being developed to incorporate this heterogeneity in hierarchical frameworks (Bastille-Rousseau et al. 2016; Börger and Fryxell 2012). Future directions for this work should focus on taking a step back from seed dispersal, and first characterizing the differences in individual animal movement patterns, since as we observed from our results, individuals generate a wide range of seed dispersal kernels based on their individual movements. In the present study we only focused on using movement rates for each indvidual bird, an average measure of their movement, and thus it is likely that we are also underestimating the frequency of long-distance movements in frugivores. This presents a novel opportunity to also incorporate statistics of extremes for rare long-distance animal movements. Previous work has shown that maximum travel distances for *Pteroglossus pluricinctus* can reach up to 3665 meters in a 30 minute tracking interval (Holbrook 2011), estimates that our movement simulations were not able to capture, and thus underestimate how far animals and seeds can travel. In addition to this, our simulations also assume equal contributions of seed dispersal by each individual, when in reality we understand that not all dispersers are the same, with some distributing larger numbers of seeds depending on their breeding status, age or size [don’t have a citation for this]. Future research focusing on animal movement models and seed dispersal would also greatly benefit from understanding the relative contributions of individual dispersers to overall seed dispersal kernels.

# Acknowledgements

# References

1. Araújo, Márcio S., Daniel I. Bolnick, and Craig A. Layman. 2011. “The Ecological Causes of Individual Specialisation.” *Ecology Letters* 14 (9): 948–58. <https://doi.org/10.1111/j.1461-0248.2011.01662.x>.
2. Bastille-Rousseau, Guillaume, Jonathan R. Potts, Charles B. Yackulic, Jacqueline L. Frair, E. Hance Ellington, and Stephen Blake. 2016. “Flexible Characterization of Animal Movement Pattern Using Net Squared Displacement and a Latent State Model.” *Movement Ecology* 4 (1): 15. <https://doi.org/10.1186/s40462-016-0080-y>.
3. Beisel, Craig J, Darin R Rokyta, Holly A Wichman, and Paul Joyce. 2007. “Testing the Extreme Value Domain of Attraction for Distributions of Beneficial Fitness Effects.” *Genetics* 176 (4): 2441–9.
4. Benedetti-Cecchi, Lisandro. 2003. “The Importance of the Variance Around the Mean Effect Size of Ecological Processes.” *Ecology* 84 (9): 2335–46. <https://doi.org/10.1890/02-8011>.
5. Bohrer, GIL, RAN Nathan, and Sergei Volis. 2005. “Effects of Long-Distance Dispersal for Metapopulation Survival and Genetic Structure at Ecological Time and Spatial Scales.” *Journal of Ecology* 93 (5): 1029–40.
6. Bolnick, Daniel I., Priyanga Amarasekare, Márcio S. Araújo, Reinhard Bürger, Jonathan M. Levine, Mark Novak, Volker H. W. Rudolf, Sebastian J. Schreiber, Mark C. Urban, and David A. Vasseur. 2011. “Why Intraspecific Trait Variation Matters in Community Ecology.” *Trends in Ecology & Evolution* 26 (4): 183–92. <https://doi.org/10.1016/j.tree.2011.01.009>.
7. Bolnick, Daniel I., Travis Ingram, William E. Stutz, Lisa K. Snowberg, On Lee Lau, and Jeff S. Paull. 2010. “Ecological Release from Interspecific Competition Leads to Decoupled Changes in Population and Individual Niche Width.” *Proceedings of the Royal Society B: Biological Sciences* 277 (1689): 1789–97. <https://doi.org/10.1098/rspb.2010.0018>.
8. Bolnick, Daniel I., Richard Svanbäck, James A. Fordyce, Louie H. Yang, Jeremy M. Davis, C. Darrin Hulsey, and Matthew L. Forister. 2003. “The Ecology of Individuals: Incidence and Implications of Individual Specialization.” *The American Naturalist* 161 (1): 1–28. <https://doi.org/10.1086/343878>.
9. Börger, Luca, and John Fryxell. 2012. “Quantifying Individual Differences in Dispersal Using Net Squared Displacement.” *Dispersal Ecology and Evolution* 30: 222–30.
10. Carlo, Tomas A, and Joshua J Tewksbury. 2014. “Directness and Tempo of Avian Seed Dispersal Increases Emergence of Wild Chiltepins in Desert Grasslands.” *Journal of Ecology* 102 (1): 248–55.
11. Coles, Stuart, Joanna Bawa, Lesley Trenner, and Pat Dorazio. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Vol. 208. Springer.
12. Comita, Liza S, Simon A Queenborough, Stephen J Murphy, Jenalle L Eck, Kaiyang Xu, Meghna Krishnadas, Noelle Beckman, and Yan Zhu. 2014. “Testing Predictions of the Janzen–Connell Hypothesis: A Meta-Analysis of Experimental Evidence for Distance-and Density-Dependent Seed and Seedling Survival.” *Journal of Ecology* 102 (4): 845–56.
13. Côrtes, Marina Corrêa, and María Uriarte. 2013. “Integrating Frugivory and Animal Movement: A Review of the Evidence and Implications for Scaling Seed Dispersal.” *Biological Reviews* 88 (2): 255–72. <https://doi.org/10.1111/j.1469-185X.2012.00250.x>.
14. Darwin, Charles, and others. 1859. *The Origin of Species by Means of Natural Selection*. Collin’s Clear-Type Press.
15. Delignette-Muller, Marie Laure, and Christophe Dutang. 2015. “fitdistrplus: An R Package for Fitting Distributions.” *Journal of Statistical Software* 64 (4): 1–34. <http://www.jstatsoft.org/v64/i04/>.
16. Forsman, Anders, and Lena Wennersten. 2016. “Inter-Individual Variation Promotes Ecological Success of Populations and Species: Evidence from Experimental and Comparative Studies.” *Ecography* 39 (7): 630–48. <https://doi.org/10.1111/ecog.01357>.
17. Gaines, Steven D, and Mark W Denny. 1993. “The Largest, Smallest, Highest, Lowest, Longest, and Shortest: Extremes in Ecology.” *Ecology* 74 (6): 1677–92.
18. García, Cristina, and Luís Borda-de-Água. 2017. “Extended Dispersal Kernels in a Changing World: Insights from Statistics of Extremes.” *Journal of Ecology* 105 (1): 63–74. <https://doi.org/10.1111/1365-2745.12685>.
19. Gilleland, Eric, and Richard W. Katz. 2016. “extRemes 2.0: An Extreme Value Analysis Package in R.” *Journal of Statistical Software* 72 (8): 1–39. <https://doi.org/10.18637/jss.v072.i08>.
20. González-Varo, Juan P., and Anna Traveset. 2016. “The Labile Limits of Forbidden Interactions.” *Trends in Ecology & Evolution* 31 (9): 700–710. <https://doi.org/10.1016/j.tree.2016.06.009>.
21. Holbrook, Kimberly M. 2011. “Home Range and Movement Patterns of Toucans: Implications for Seed Dispersal.” *Biotropica* 43 (3): 357–64. <https://doi.org/10.1111/j.1744-7429.2010.00710.x>.
22. Holbrook, Kimberly Mae. 2007. “Seed Dispersal Limitation in a Neotropical Nutmeg, Virola Flexuosa (Myristicaceae): An Ecological and Genetic Approach.”
23. Holbrook, K. M., and B. A. Loiselle. 2007. “Using Toucan-Generated Dispersal Models to Estimate Seed Dispersal in Amazonian Ecuador.” In *Seed Dispersal: Theory and Its Application in a Changing World*, 300–321. CAB International, Wallingford.
24. ———. 2009. “Dispersal in a Neotropical Tree, Virola Flexuosa (Myristicaceae): Does Hunting of Large Vertebrates Limit Seed Removal?” *Ecology* 90 (6): 1449–55. <https://doi.org/10.1890/08-1332.1>.
25. Howe, Henry F, and Maria N Miriti. 2004. “When Seed Dispersal Matters.” *BioScience* 54 (7): 651–60.
26. Howe, Henry F, and Judith Smallwood. 1982. “Ecology of Seed Dispersal.” *Annual Review of Ecology and Systematics* 13 (1): 201–28.
27. Ibáñez, Inés, James S Clark, Michael C Dietze, Ken Feeley, Michelle Hersh, Shannon LaDeau, Allen McBride, Nathan E Welch, and Michael S Wolosin. 2006. “Predicting Biodiversity Change: Outside the Climate Envelope, Beyond the Species–Area Curve.” *Ecology* 87 (8): 1896–1906.
28. Inouye, Brian D. 2005. “The Importance of the Variance Around the Mean Effect Size of Ecological Processes: Comment.” *Ecology* 86 (1): 262–65. <https://www.jstor.org/stable/3451006>.
29. Jones, Landon R., Scott M. Duke‐Sylvester, Paul L. Leberg, and Derek M. Johnson. 2017. “Closing the Gaps for Animal Seed Dispersal: Separating the Effects of Habitat Loss on Dispersal Distances and Seed Aggregation.” *Ecology and Evolution* 7 (14): 5410–25. <https://doi.org/10.1002/ece3.3113>.
30. Jordano, Pedro. 2017. “What Is Long-Distance Dispersal? And a Taxonomy of Dispersal Events.” *Journal of Ecology* 105 (1): 75–84. <https://doi.org/10.1111/1365-2745.12690>.
31. Jordano, Pedro, Coralith Garcı́a, José A Godoy, and Juan Luis Garcı́a-Castaño. 2007. “Differential Contribution of Frugivores to Complex Seed Dispersal Patterns.” *Proceedings of the National Academy of Sciences* 104 (9): 3278–82.
32. Joyce, Paul, Darin R Rokyta, Craig J Beisel, and H Allen Orr. 2008. “A General Extreme Value Theory Model for the Adaptation of Dna Sequences Under Strong Selection and Weak Mutation.” *Genetics* 180 (3): 1627–43.
33. Katz, Richard W, Grace S Brush, and Marc B Parlange. 2005. “Statistics of Extremes: Modeling Ecological Disturbances.” *Ecology* 86 (5): 1124–34.
34. Levey, Douglas J., Benjamin M. Bolker, Joshua J. Tewksbury, Sarah Sargent, and Nick M. Haddad. 2005. “Effects of Landscape Corridors on Seed Dispersal by Birds.” *Science* 309 (5731): 146–48. <https://doi.org/10.1126/science.1111479>.
35. Levey, Douglas J., Joshua J. Tewksbury, and Benjamin M. Bolker. 2008. “Modelling Long-Distance Seed Dispersal in Heterogeneous Landscapes.” *Journal of Ecology* 96 (4): 599–608. <https://doi.org/10.1111/j.1365-2745.2008.01401.x>.
36. Loayza, Andrea P, and Rodrigo S Rios. 2014. “Seed-Swallowing Toucans Are Less Effective Dispersers of Guettarda Viburnoides (Rubiaceae) Than Pulp-Feeding Jays.” *Biotropica* 46 (1): 69–77.
37. Morales, Juan Manuel, and Tomás A. Carlo. 2006. “The Effects of Plant Distribution and Frugivore Density on the Scale and Shape of Dispersal Kernels.” *Ecology* 87 (6): 1489–96. [https://doi.org/10.1890/0012-9658(2006)87[1489:TEOPDA]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87%5B1489:TEOPDA%5D2.0.CO;2).
38. Nathan, Ran. 2006. “Long-Distance Dispersal of Plants.” *Science* 313 (5788): 786–88.
39. Nathan, Ran, and Helene C. Muller-Landau. 2000. “Spatial Patterns of Seed Dispersal, Their Determinants and Consequences for Recruitment.” *Trends in Ecology & Evolution* 15 (7): 278–85. <https://doi.org/10.1016/S0169-5347(00)01874-7>.
40. Pegman, Andrew P. McKenzie, George L. W. Perry, and Mick N. Clout. 2017. “Exploring the Interaction of Avian Frugivory and Plant Spatial Heterogeneity and Its Effect on Seed Dispersal Kernels Using a Simulation Model.” *Ecography* 40 (9): 1098–1109. <https://doi.org/10.1111/ecog.02191>.
41. Post, David M., Eric P. Palkovacs, Erika G. Schielke, and Stanley I. Dodson. 2008. “Intraspecific Variation in a Predator Affects Community Structure and Cascading Trophic Interactions.” *Ecology* 89 (7): 2019–32. <https://doi.org/10.1890/07-1216.1>.
42. R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
43. Rehm, Evan M., Janelle Chojnacki, Haldre S. Rogers, and Julie A. Savidge. 2018. “Differences Among Avian Frugivores in Seed Dispersal to Degraded Habitats.” *Restoration Ecology* 26 (4): 760–66. <https://doi.org/10.1111/rec.12623>.
44. Rogers, Haldre S., Noelle G. Beckman, Florian Hartig, Jeremy S. Johnson, Gesine Pufal, Katriona Shea, Damaris Zurell, et al. 2019. “The Total Dispersal Kernel: A Review and Future Directions.” *AoB PLANTS* 11 (5). <https://doi.org/10.1093/aobpla/plz042>.
45. Russo, Sabrina E., Stephen Portnoy, and Carol K. Augspurger. 2006. “Incorporating Animal Behavior into Seed Dispersal Models: Implications for Seed Shadows.” *Ecology* 87 (12): 3160–74. [https://doi.org/10.1890/0012-9658(2006)87[3160:IABISD]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87%5B3160:IABISD%5D2.0.CO;2).
46. Schupp, Eugene W. 1993. “Quantity, Quality and the Effectiveness of Seed Dispersal by Animals.” *Vegetatio* 107 (1): 15–29.
47. Schupp, Eugene W., Pedro Jordano, and José María Gómez. 2010. “Seed Dispersal Effectiveness Revisited: A Conceptual Review.” *New Phytologist* 188 (2): 333–53. <https://doi.org/10.1111/j.1469-8137.2010.03402.x>.
48. Violle, Cyrille, Brian J. Enquist, Brian J. McGill, Lin Jiang, Cécile H. Albert, Catherine Hulshof, Vincent Jung, and Julie Messier. 2012. “The Return of the Variance: Intraspecific Variability in Community Ecology.” *Trends in Ecology & Evolution* 27 (4): 244–52. <https://doi.org/10.1016/j.tree.2011.11.014>.
49. Will, Heidrun, and Oliver Tackenberg. 2008. “A Mechanistic Simulation Model of Seed Dispersal by Animals.” *Journal of Ecology* 96 (5): 1011–22. <https://doi.org/10.1111/j.1365-2745.2007.01341.x>.
50. Wolf, Max, and Franz J. Weissing. 2012. “Animal Personalities: Consequences for Ecology and Evolution.” *Trends in Ecology & Evolution* 27 (8): 452–61. <https://doi.org/10.1016/j.tree.2012.05.001>.
51. Zwolak, Rafał. 2018. “How Intraspecific Variation in Seed-Dispersing Animals Matters for Plants.” *Biological Reviews* 93 (2): 897–913. <https://doi.org/10.1111/brv.12377>.

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.

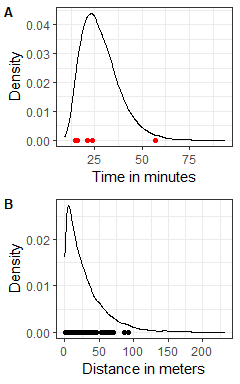


Figure 2.

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.

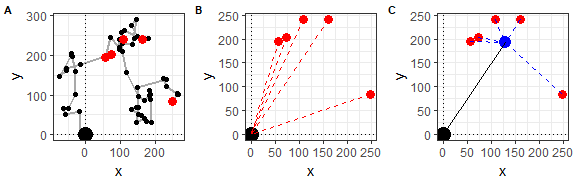


Figure 3.

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.

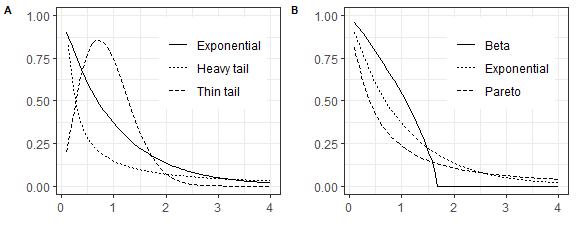


Figure 4. Seed dispersal distances and aggregation metrics for simulated seeds in the three models considering variation in animal movement rates. A. Density kernels for frugivore-generated seed dispersal distances for each seed simulated under the three models of variation in animal movement rates. 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.

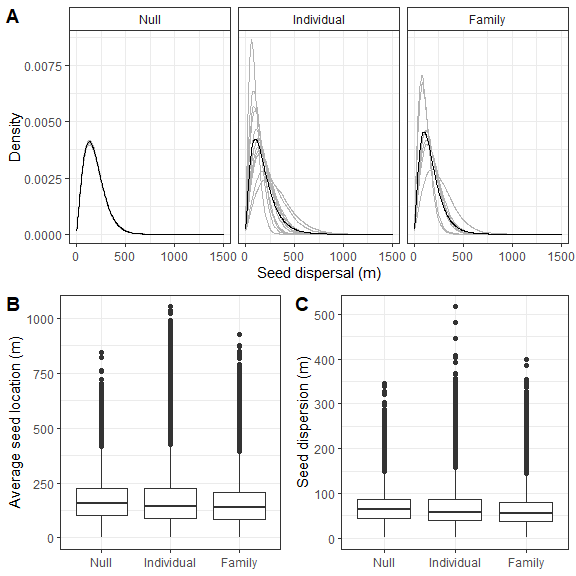


Table 1. Summary table of seed dispersal metrics for each of the simulation models. Seed dispersal and seed dispersion metrics represent the average value for each of the models, with standard errors in parentheses. Kurtosis is calculated for each of the three simulation models as described in the main 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)% |

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.

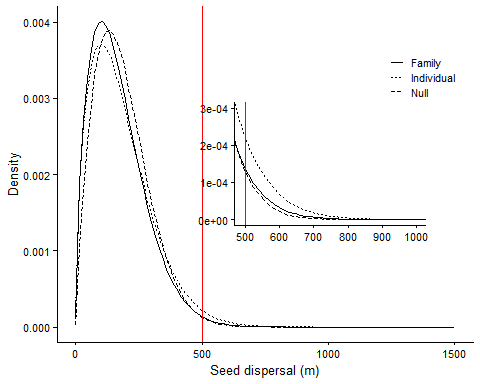


Table 3. Threshold values and parameter estimates for generalized Pareto distribution fits, and associated standard error for parameter estimates.

|  |  |  |  |
| --- | --- | --- | --- |
| 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.

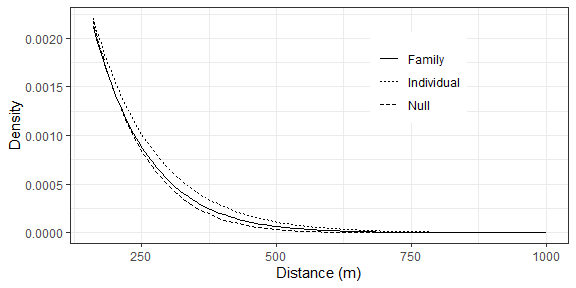
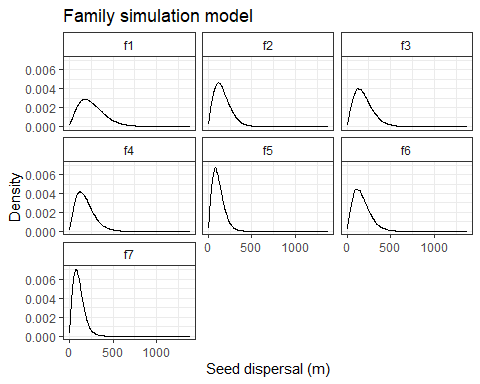
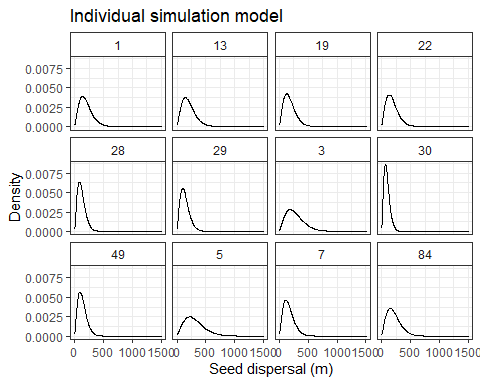
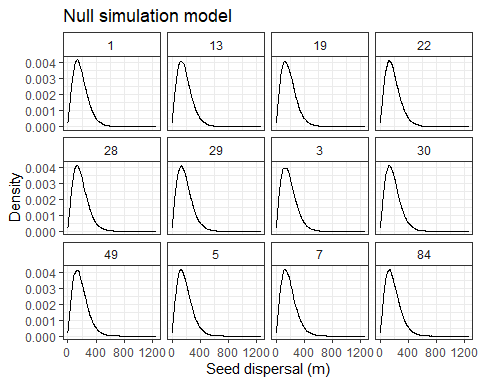


Table 4. Probability of long-distance dispersal events for the three models following a generalized Pareto distribution.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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

