Variation in distance between consecutive animal locations

This should be the abstract

# Introduction

Movement is a fundamental ecological process, critical to an animal’s biology and it’s interactions with the environment. The study of animal movement tends to focus on four basic mechanistic questions: the why, when, where, and how an animal moves, and these questions can be answered at the scale of individuals or populations. A general framework for movement ecology is needed to link movement patterns and processes. Such a framework would start with a focus on movement itself and the individual’s internal state, navigation capacity, motion capacity, and responses to external factors (Nathan *et al.* 2008). The overall goal of studying movement is to find the linkages between the causes and mechanisms of animal movement, that explain spatiotemporal patterns that then feed back into various ecological and evolutionary processes. Some of these approaches involve separating different movement patterns into states, with each state having an associated set of parameters that determine that specific movement pattern (Langrock *et al.* 2012). These approaches segment movement patterns, in order to associate specific types of movement to underlying behaviors or fine-scale environmental influences (Getz citations, Morales, more focus on the random walk models, Patterson2008). Several of these approaches focus on the why and underlying processes that drive animal movement, whether that is an animal’s internal state or responses to their environment.

Random walks have been used broadly in biology to characterize animal movement, and increasing levels of complexity can be incorporated by using multiple random walks, continuous time or different levels of correlation between turning angles (Morales *et al.* 2004; Codling *et al.* 2008). More recent developments have focused on implementing higher order Markov processes through state-space models, which provide a probabilistic framework that allows for prediction of future states of a system based from previous states, by coupling an observation model with an underlying mechanistic model. When extending this framework to movement ecology, we understand that the process model provides the mechanism for movement across space and time, and the observation model is related to the sampling process. In movement data, a state-space process allows us to couple a statistical model for the observation method, accounting for errors in sampling or detection, with a separate model for the movement dynamics which are determined by the effects of behavior or responses to the environment (Patterson *et al.* 2008). In general, state-space models are used to identify different states, or behavioral states, that determine different movement patterns associated to step length and turning angle. And some of the main questions asked with this approach are related to identifying these different states, or behavioral state-switching, from time series data on location tracking.

Although I have mostly described approaches used to understand why movement happens, animal movement also has consequences, which might influence species interactions, ecological function, and vital demographic processes (such as survival, reproduction, or dispersal of other species) and thus result in population or regional level impacts. A clear example of the consequences of animal movement on other species involves animal-mediated seed dispersal or pollination, impacting a plant’s spatial population dynamics.The specific case of seed dispersal is interesting because of the different levels and interactions happening. At the individual level, we focus on an individual bird consuming fruit and dispersing the seeds of the tree. How this particular individual moves across the landscape directly affects the dispersal for that specific tree and other trees from which it consumes fruit. At the same time, that individual has intra-individual variation in movement patterns associated to specific events in that individual’s lifetime, such as breeding and nesting, or even smaller scale variation related to different states, such as foraging or resting. Interspecific interactions come at play as multiple species find themselves together feeding from the same trees, and thus competing for these resources, while at the same time intraspecific interactions due to social behavior broadly determine general movement patterns for animals restricted within family home ranges. From the plant’s perspective, the variation in movement patterns across all dispersers will determine eventual plant population spatial spread.

Specifically in the context of seed dispersal ecology, the general focus has been on understanding how the typical movement patterns, built around average measures at population levels, can influence plant population spread and seed dispersal [sources]. Animal movement can be modeled with a simple diffusion process or a random walk (Levey *et al.* 2005; Jones *et al.* 2017), with the simplest approach assuming model parameters are the same across individuals, something known as ‘complete pooling’ (Langrock *et al.* 2012). This focus on generality has driven to a lack of frameworks for understanding how variation in animal movement is maintained, and the consequences of this variation across populations and communities. From the animal movement perspective, some alternatives have been considered to incorporate individual variation with ‘no pooling’ or ‘partial pooling’ approaches, where each or some individuals have their own set of parameters (Patterson *et al.* 2009), however a high number of parameters quickly limits applications to high levels of individual variation. Recent calls for incorporating individual variation in movement and their consequences to seed dispersal or even pathogen spread (Snell *et al.* 2019) have pushed for the need to develop theory that explores different types of variation in movement patterns to understand its consequences across ecological organizational levels. In particular, when considering species interactions, incorporating individual variation becomes essential to understand how positive or negative feedback loops play into the maintenance of variation in movement as individuals respond to external factors in their environment (Shaw 2020).

We focus here on developing a process to analyze animal movement sequences based on movement lengths, or step lengths, and turning angle probability distributions. Following the methods of Kareiva and Shigesada (1983), we assume the movement is a correlated random walk and derive expected square displacements associated to a number of consecutive moves under various probability distribution models.A simple correlated random walk, which involves a directional correlation between successive step movement (Codling *et al.* 2008). Unlike other studies which focus on deciphering the underlying behaviors from animal movement data, our focus was to uncover potential differences between individuals and how these differences can be obscured when modeling population-level animal movement. We know these simple correlated random walks, or simple movement models might not be enough to describe net displacements under more complex processes, for which developments in higher order markov processes show to be promising. However, our goal in this study was to use a simpler model and focus on potential differences when analyzing complete pooling versus no pooling data. In the context of animal-mediated seed dispersal, understanding these indivdiual level differences can help us describe and characterize the implications that different animals may have over spatial distributions of plant populations, and thus identify potential dispersers with higher or more importance, particularly when associated to long-distance seed dispersal events.

# Methods

We focused our study of individual variation on the many-banded araçari (*Pterglossus pluricinctus*), a small toucan in Yasuní National Park, Ecuador. We used previously collected data from studies using radiotelemetry to estimate home ranges and evaluate the potential seed dispersal distances for the *Virola flexuosa* tree (Holbrook 2011). Over a period of four years, from 2001 to 2005, *P. pluricinctus* individuals were captured and radio-transmitters were attached at the base of the central tail feathers (Holbrook 2011). Toucan locations were estimated by triangulation using receivers and hand-held antennas, with tracking periods lasting between 4-5 hours per day, locating individual birds every 15 minutes. Further details on field methods can be found in Holbrook (**???**). It is important to note that these methods did not tack all birds simultaneously, and that although individual toucan locations were attempted every 15 minutes, this was not always possible, thus some successive locations have time categories associated to multiples of 15. $\color{red}{For our analyses, we focused on toucans with a minimum of \*\*20?\*\* recorded locations, and we calculated turning angles (in radians), and the associated Euclidean step lengths in meters for each successive set of locations and turning directions. We worked under a velocity-based framework, and thus divided step lengths by the time interval during which they were recorded. This gave us a total of \*\*600ish?\*\* data points for turning angles and velocities.}$

**Could I find anything pre HMMs of SSMs? Wondering what people could do before? And wondering this because we have broken up tracking periods, not continuous ones**  
Seems like I need to focus more on random walks

Assume weibull distributions for step length and wrapped cauchy distributions for the turning angles (Langrock *et al.* 2012), but they tried Gamma and von Mises, respectively, but got outperformed by AIC. Used negative binomial for the state dwell times.

Random walks used broadly in biology to characterize animal and cell movement. Describes que commonly used distributions for angles: von Mises, wrapped normal and wrapped Cauchy distributions. (Codling *et al.* 2008) - IDEA: so maybe use these distributions for the angles. We might see that there is not much difference in angle movement across individuals? They all tend to move forward perhaps, or something like that, and where the real differences are is in the distances. - Also might be intriguing to incorporate the time as a set of multiple random walks? or repeats the velocity value for that number of minutes. Maybe just using a correlated random walk, which takes into account short-term correlations in the direction of movement? - Codling describes this as a velocity jump process, since the markov process is the walker’s velocity rather than the location. Details at the end of page 7. - We consider the individuals moving in space at a constant speed , and at each time step each individual changes direction and moves a distance in a new direction (with probability ), or moves a distance in the previous direction (with probability ). Hence, turning events occur as a Poisson process with rate . **This is textual from paper, so modify** Their focus is also on multiple individuals at the same time, and the density of individuals at a given time and location. - In my case, the distance would also be randomly sampled, or this are the velocities I suppose. So, perhaps it is constant speed, and the distance is actually times to account for 15, 30, etc minutes based on the data. - Remember too that this velocity is actually the distance over time - Pg11. A correlated random walk, CRW, consisting of a series of independet draws from a step length PDF and a turning angle PDF for each step, a first order markov process. - using mean square displacement. The effect of step length variability on MSD can be quite significant. - MSD being of interest to ecologists due to its relation with the diffusion coefficient D. equations 2.13 and 2.14 \*\* Should write a conclusion about this paper\*\*

(Kareiva & Shigesada 1983) - approximating displacements by connecting a series of straight lines, and thus summarize animal movement based on parameters for movement length and turning angle. Using square displacement instead of linear displacement, because it’s expectation can be calculated from turning angles and step lengths. Assume that the length of each move and the turning angle are independent random variables, each with its own probability density. Assume that a series of moves can be represented by random draws from these probability densities, and since each random draw is independent from the preceding ones, the random draw process is a first order markov chain. The result is a correlated random walk, and the distribution g(theta) gives a measure of the degree to which the angles of movement are correlated. *Would be interesting to add a figure with what I am actually analyzing. So, visually show how the gps locations get turned into straight lines, and angles, with an associated time interval. Or we could just multiply everything? So have it all for 15 minute intervals, and then repeat the ones that went on longer, perhaps set a limit for 60 minutes?*

probabilistic rules of movement.

As in other papers. Note that we adopt the convention of using upper case letters for random variables and lower case letters for possible numerical values of these random variables.

“Location accuracy using radio telemetry may be reduced in tropical forests… conservative distance categories of 100-m increments to better represent precision. Time categories: 15, 30, 60, 90 minutes. Then calculated probability of movements made within each distance category within each time category, summed across each time category to give a final probability for each distance category.”

“Distances travelled per movement bout ranged from 0 to >2000m. Strong leptokurtic distributions of movements with most being <300m. Longest recorded movement for Pluricinctus is 3665m (recorded in 30 minutes)” \*this is a corresponding velocity of 122.1666667 meters per minute.

Evaluating models with their individual-specific counterparts, using AIC or BIC. In individual-specific models the AIC results from the joint likelihood of the individual-specific models.

Consider incorporating NAs for missed observations so we keep the regular 15 minute time interval.

# Questions

* What is the distribution of distances between consecutive recorded locations?
* Are there differences between individuals or can we use the same distribution to describe these distances between two locations?

The reasoning behind this, is that distances between locations can be used later on to describe variation in step length when simulating animal movement under simple models such as random walk.

### MPM: meters per minute

Due to the variation in time between consecutive locations, we scaled the step lengths to displacement in one minute. **Visualize the variation and distribution of these distances between locations**

# Fit distribution

Useful resource to use the package [fitdistrplus](https://cran.r-project.org/web/packages/fitdistrplus/vignettes/FAQ.html) Also check [this](https://stackoverflow.com/questions/37152482/power-law-fitted-by-fitdistr-function-in-package-fitdistrplus). Should I do the opimization by hand?

## Information criteria

So, for each individual, which is the best fitting distribution, based on AIC or BIC? This is the distribution with the lowest AIC or BIC score for each individual. These are basically the distributions that will be used for the mixed distribution model later on.

However, it is not as straight-forward since there are multiple competing models for certain individuals, which we consider as those within less than 2 units of the delta AIC or BIC. When using AIC, there are competing models for almost every individual, whereas with BIC, only 4 of the individuals have competing models.

Now we can compare what we call *models*, in which we compare a model that does complete pooling and considers all the data together, and the other model that considers individual variation, and so it fits a separate distribution to the data from each individual.

The individual variation models will fit a distribution to each individual. At one level we will use the same distribution for all individuals and only consider variation in parameters. At the next level, we will consider a change in parameters and distributions, where individuals can have different distributions, and this is what we call the mixed distribution model for individual variation.

Considering just focusing on BIC because of JMP’s paper, and also in this specific case, it provides less competing models for each individual

## Probability distribution parameters

We can compare the different parameter values estimated for each distribution and their standard deviations. These parameters can give us some insight into some of the characteristics of each set of distances moved per minute, such as which individuals have higher means or longer tails in the distribution of their distances moved per minute. The x axis in the following plots highlights which individuals have the lowest BIC (yellow) or AIC(magenta) with that model. The blue lines show the value of that parameter for the population level model, and dashed blue lines show the upper and lower limits of that parameter estimate. Overall, at the population level, the lognormal model had the lowest AIC and BIC values.

# References

Codling, E.A., Plank, M.J. & Benhamou, S. (2008). Random walk models in biology. *Journal of the Royal society interface*, 5, 813–834.

Holbrook, K.M. (2011). Home range and movement patterns of toucans: Implications for seed dispersal. *Biotropica*, 43, 357–364.

Jones, L.R., Duke-Sylvester, S.M., Leberg, P.L. & Johnson, D.M. (2017). Closing the gaps for animal seed dispersal: Separating the effects of habitat loss on dispersal distances and seed aggregation. *Ecology and evolution*, 7, 5410–5425.

Jonsen, I.D., Myers, R.A. & James, M.C. (2006). Robust hierarchical state-space models reveal diel variation in travel rates of migrating leatherback turtles. *Journal of Animal Ecology*, 75, 1046–1057.

Kareiva, P. & Shigesada, N. (1983). Analyzing insect movement as a correlated random walk. *Oecologia*, 56, 234–238.

Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D. & Morales, J.M. (2012). Flexible and practical modeling of animal telemetry data: Hidden markov models and extensions. *Ecology*, 93, 2336–2342.

Levey, D.J., Bolker, B.M., Tewksbury, J.J., Sargent, S. & Haddad, N.M. (2005). Effects of landscape corridors on seed dispersal by birds. *Science*, 309, 146–148.

Morales, J.M., Haydon, D.T., Frair, J., Holsinger, K.E. & Fryxell, J.M. (2004). Extracting more out of relocation data: Building movement models as mixtures of random walks. *Ecology*, 85, 2436–2445.

Nathan, R., Getz, W.M., Revilla, E., Holyoak, M., Kadmon, R. & Saltz, D. *et al.* (2008). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences*, 105, 19052–19059.

Patterson, T.A., Basson, M., Bravington, M.V. & Gunn, J.S. (2009). Classifying movement behaviour in relation to environmental conditions using hidden markov models. *Journal of Animal Ecology*, 78, 1113–1123.

Patterson, T.A., Thomas, L., Wilcox, C., Ovaskainen, O. & Matthiopoulos, J. (2008). State-space models of individual animal movement. *Trends in ecology & evolution*, 23, 87–94.

Shaw, A.K. (2020). Causes and consequences of individual variation in animal movement. *Movement ecology*, 8, 1–12.

Snell, R.S., Beckman, N.G., Fricke, E., Loiselle, B.A., Carvalho, C.S. & Jones, L.R. *et al.* (2019). Consequences of intraspecific variation in seed dispersal for plant demography, communities, evolution and global change. *AoB Plants*, 11, plz016.