SPECIES COMPOSITION AND SPATIAL ECOLOGY OF AMAZONIAN UNDERSTORY MIXED-SPECIES FLOCKS IN A FRAGMENTED LANDSCAPE

A Dissertation
Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

The School of Renewable Natural Resources

by
Karl S. Mokross
B.S., Universidade Federal de São Carlos, 2001
M.S., Instituto Nacional de Pesquisas da Amazônia, 2004
December 2014
Dedicado a Thomas C. e Marina.
ACKNOWLEDGMENTS

Aos mateiros do PDBFF; poucas pesquisas na Amazônia teriam sido realizadas sem sua ajuda. Vocês merecem uma enorme parcela de crédito neste trabalho. Jairo Lopez e Léo Marajó foram grandes professores, obrigado por me ensinarem a caminhar pela mata.
À todos que me ajudaram em campo, muito obrigado por sua ajuda. Elton Lima Retroz, Pablo Hendrigo, Beatriz Souza, Ana Carlena Vilela, Angelica Nunes e Marconi Campos.
À equipe do PDBFF por toda a ajuda logística: motoristas, limpadores de trilhas e gambiarras com pouco equipamento, mas muito talento. Rosely, Batatinha, Seu Martins, Ari, Cabocão, Juninho, Josimar, Adriane e Zé Luiz.
Ao meu orientador Phil Stouffer; obrigado por me passar um projeto que abriga tantas perguntas interessantes e pela interminável paciência.
À equipe administrativa da School of Renewable Natural Resources: William Kelso, Allen Rutherford, Nedra Ghoram e Karen Chambre.
À minha banca examinadora que me acompanhou durante o desenvolvimento de minha pesquisa e me deu excelentes sugestões e críticas: James Van Remsen, Thomas Sherry, Thomas Dean e Christopher Carlton.
Aos meus orientadores passados por terem investindo em uma criatura desprovida de quaisquer talentos acadêmicos: Tania Sanaiotti, Carlos Roberto Sousa e Silva, e Luiz Eduardo Anelli.
Aos amigos de Manaus, Sergio Marquez, Manoela Borges, Gabriela Zuquim, Jonas Gebara e Fábio Rohe.
Aos professores da LSU e seus insights inspiradores: Kyle Harms, Bruce Williamson, e Michael Kaller.
Ao povo do Bird Lunch por contribuir com ótimas discussões, grande parte do meu crescimento acadêmico aconteceu nas Sextas entre as 12:00 e 13:00.
Aos doutores Scott Saleska e Michael Lefsky por cederem os dados de Lidar.
À National Science Foundation, assim como a American Ornithological Union pelo financiamento desta pesquisa.
Ao passarinho com mais personalidade da Amazônia, o ipecuá. Que seu canto ainda anime muitos bandos pela Amazônia.
À minha família pelo apoio Beverly, Jenny, Thomas, Bernhard, Glaucia, Karen, Barbara e Ervino.
À Marina Corrêa Côrtes, pelos conselhos, amizade e carinho.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS .................................................................................................................. iii

ABSTRACT .................................................................................................................................. vii

CHAPTER 1: INTRODUCTION .................................................................................................... 1
  LITERATURE CITED ................................................................................................................ 5

CHAPTER 2: SPATIAL BEHAVIOR OF AMAZONIAN UNDERSTORY MIXED-SPECIES FLOCKS IN A HETEROGENEOUS LANDSCAPE ................................................................. 7
  INTRODUCTION ..................................................................................................................... 7
  METHODS ............................................................................................................................... 12
    Study Site ............................................................................................................................ 12
    Data collection .................................................................................................................... 14
    Trajectory analysis ............................................................................................................ 14
    Home ranges next to secondary forests ........................................................................ 16
    Home range layout in the landscape ................................................................................. 18
  RESULTS .................................................................................................................................. 18
    Trajectory analyses ........................................................................................................... 20
    Home range analyses ......................................................................................................... 23
    Home range layout in the landscape ................................................................................ 26
  DISCUSSION .......................................................................................................................... 27
    Trajectory analysis ............................................................................................................ 30
    Home range layout ........................................................................................................... 33
    Concluding remarks ......................................................................................................... 33
  LITERATURE CITED .............................................................................................................. 35

CHAPTER 3: MODELING THE SPATIAL BEHAVIOR OF UNDERSTORY MIXED-SPECIES FLOCKS ........................................................................................................................ 41
  LITERATURE CITED .............................................................................................................. 46

CHAPTER 4: DECAY OF INTERSPECIFIC AVIAN FLOCK NETWORKS ALONG A DISTURBANCE GRADIENT IN AMAZONIA ................................................................................. 49
  INTRODUCTION ..................................................................................................................... 49
  MATERIAL AND METHODS .................................................................................................. 53
    Study site and data collection ............................................................................................ 53
    Habitat and vegetation characterization ........................................................................... 54
    Network and statistical analyses ...................................................................................... 55
  RESULTS .................................................................................................................................. 58
  DISCUSSION .......................................................................................................................... 60
    Species-level and flock-level changes in network structure ........................................... 64
    Habitat and the consequences of changes in social structure ....................................... 66
    Conclusions ....................................................................................................................... 67
  LITERATURE CITED .............................................................................................................. 68
ABSTRACT

With the ongoing advance of the agricultural frontier in the Amazon basin, it is inevitable that heterogeneous landscapes will play a key role in conservation. These landscapes are mostly composed of patchworks of small forest fragments, secondary forests and roads. Conservation, however must take species interactions into consideration as they play a pivotal part the maintenance of several biological processes in the tropics. One of the most conspicuous interspecific interactions are seen in mixed-species flocks of birds, which in the Amazon, represent one of the best organized systems of bird aggregations. In this research, I assess how flock spatial behavior and species compositions are affected by changes in habitat structure. I followed 29 mixed-species flocks in different landscapes types such as secondary forests, forest fragments of 10 and 100 ha, and mixes of primary and secondary forest patches. As flocks foraged through their territories, I recorded their species composition every 30 minutes and georeferenced their movements every 30 seconds. Flocks spatial behavior was severely affected by anthropogenic features such as forest edges and secondary forests as flocks respond strongly to vegetation height. Using step-selection models, it was possible to reproduce flock movements and show that they prefer taller vegetation and lower areas of topography such as stream valleys. Due to this behavior, flocks avoided areas where canopy height was below 15 meters, and extensive areas of secondary below this height hold unstable flocks that do not persist for long periods. The ones that persisted showed home ranges that were much larger than what was observed in primary forest. Time spent in secondary forest was dependent on vegetation height, but not area, which seems to be shaped by intraspecific interactions. Flock social structure is also severely affected by habitat structure. Flock species richness did not show a predictable pattern, but participation was negatively affected. In fact, our data indicates that flock social structure
may take longer to recover than spatial behavior. Assessing a 30-year mist-net capture dataset, we were able to determine that indeed, decreased species participation seems to be a more important driver in flock dissolution than local extinction.
CHAPTER 1: INTRODUCTION

Currently, deforestation in the Amazon gives signs of slowing down (Nepstad et al. 2014). Yet the area that has been and will be converted spans sub-continental magnitudes. Deforestation is not spatially homogeneous; what is left behind after the expansion of the agricultural frontier is a patchwork of forest remnants, roads and secondary forests (Laurance et al. 2002), so it is likely that these heterogeneous landscapes will play a major role in tropical conservation in the 21st century (Gardner et al. 2009, Didham 2011). The challenge is considerable for ecologists to assess the conservation value of such landscapes as their constituent elements are variable: forest fragments may vary in size, secondary forests are temporally dynamic environments, and species show variable responses (Ewers and Didham 2007). This last item poses a significant logistical challenge; assessing one species, let alone several, can be both costly and time-consuming (Hansbauer et al. 2010).

One important element of the Amazonian avifauna provides a good model for assessing heterogeneous landscapes: understory mixed-species flocks. Considered one of the most organized avian flocking systems, this social system extends throughout most of Amazonian terra firme forests (Munn 1984, Powell 1985, English 1998, Jullien and Thiollay 1998). Meeting at every dawn on a fixed location, a core of five to ten species each represented by one reproductive pair forage through a common territory and defend it from conspecifics (Munn and Terborgh 1979). Other species are known to join these flocks regularly, and some aggregations may include up to 50 species (Martínez et al. 2013). These core species occupy a home range of about 8 ha; these are temporally and spatially stable (Martínez et al. 2013), and are known to vary in size according to vegetation density (Thiollay and Jullien 1998). Mixed-species flocks
are known to have high movement rates, but no directional patterns had been detected until now, and their movements so far have been described as random (Powell 1989).

Amazonian mixed-species flocks are known to be sensitive to disturbances, and tend to disappear in small fragments and selectively logged forests (Stouffer and Bierregaard 1995, Thiollay 1997). Common flocking species have been recorded in regenerating areas outside fragments after about a decade (Antongiovanni and Metzger 2005); yet until now, little was known about their gregarious behavior. It was unknown if individual species were alone, if there were reductions in flock participation, or if the few surviving elements were aggregating as usual. Another important set of questions regards the spatial behavior of flocks in anthropogenic landscape elements such as forest fragments, forest edges and extensive areas of secondary vegetation. No data were available on whether flocks avoid forest edges, whether they occupy secondary forest, and whether they show measurable changes in composition and spatial behavior when moving through these areas.

Few places could have offered a patchwork of such landscape elements allied with logistic support as the Biological Dynamics of Forest Fragments Project. This large-scale experiment began in the late 1970s in the central Amazon and consists of replicated forest fragments of 1, 10 and 100 ha, and a patchwork of secondary forests of different ages. Mist net sampling effort has been carried out continuously beginning before the isolation of these fragments. I analyzed two aspects flocking behavior: spatial ecology and species composition. I intended to move beyond the mere presence or absence of component species in these open systems, and wished to understand how this social system is affected by large man-made landscapes. To do this, I followed flocks and (1) georeferenced their positions every 30 seconds, and (2) noted all species present in flocks during half-hour blocks.
This dissertation is divided into four chapters. The first two chapters deal with the spatial ecology of mixed-species flocks. Initially, I provide a detailed description of flock spatial behavior in a heterogeneous landscape; secondly, I elaborate movement models that reproduce the spatial behavior of understory mixed-species flocks. The two remaining chapters focus on flock composition and social organization; the former explains changes in flocks occupying different landscape elements, and the latter provides a reconstruction of flock social structure spanning 25 years using a long-term database on mist-net captures.

From the spatial behavior aspect, I found a strong response to altered landscapes: flocks avoid areas of lower stature vegetation, and therefore respond to forest edges by tending to move parallel to them, leading to home ranges that align with man-made features. For flocks inhabiting forest edges, time spent in secondary forest tended to increase as forests exhibited taller canopies; however, the area of introgression seems to be influenced by landscape configuration and territorial interactions. Flocks tend to have much larger home ranges in secondary forest, and extensive areas of less-developed secondary forest held unstable flocks that seemingly disbanded or abandoned them.

Through collaboration with Dr. Jonathan Potts, a recent graduate working with Dr. Mark Lewis at the University of Alberta, we created a step-selection model of space use based on flock movement behavior and territorial interactions. We were able to show that flocks, despite a seemingly random movement behavior, prefer to occupy areas of lower topography and vegetation with higher canopies; their occupation of the landscape is also limited by the interactions with neighboring flocks. With these models, we may recreate flock occupation in larger landscapes, thus enabling us to forecast the numbers of territories in heterogeneous landscapes.
In relation to flock organization, we applied techniques developed in the network theory field, which allowed us to measure participation through established metrics. We found that flock participation reduces significantly in altered areas such as 10 ha-fragments and secondary forests. This reduction was not necessarily due to reductions in species richness; some areas actually held higher numbers of participating species, but they tended to remain significantly shorter periods in these aggregations.

Questions remained on flock changes over time when a landscape is drastically altered through forest fragmentation. To assess these changes, I utilized the long-term dataset on mist net captures from the Biological Dynamics of Forest Fragments Project, which spanned 1979 to 2009. I adapted network theory methods in order to analyze mist net data and assessed the co-captures of 20 different species along with the nuclear species Thamnomanes caesius. I demonstrated that flock dissolution was caused by both the reduction of the nuclear species and by reduction of flock participation by species that persist in fragments. In fact, our data suggest that species ceasing association with flocks seems to be a more important driver of flock dissolution than local extinctions of flocking species.

In summary, both the spatial ecology and social relations of Amazonian mixed-species flocks are strongly affected by changes in habitat configuration and structure. Interestingly, movement patterns and social organization do not respond synchronously; data suggest that flock spatial behavior in well-developed secondary forests is similar to what is observed in primary forests, whereas flock organization only seems to fully recover in areas with much more developed forest structure. In general, these results have increased our capacity to forecast the response of a significant part of the understory fauna. Questions remain concerning benefits of flock participation such as reproductive output and survival of individuals in flocks of varying
cohesive values, as well as mechanisms of flock disappearance in very degraded areas.

Ascertaining the conservation value of heterogeneous areas requires a multi-faceted approach. A detailed understanding of such responses will be indispensable in order to create management strategies and conservation planning.

**LITERATURE CITED**


CHAPTER 2: SPATIAL BEHAVIOR OF AMAZONIAN UNDERSTORY MIXED-SPECIES FLOCKS IN A HETEROGENEOUS LANDSCAPE

INTRODUCTION

It is expected that 40% of the Amazon could be deforested by the middle of the 21st century (Soares-Filho et al. 2006). This process is likely to spread from access points such as roads and waterways (Laurance et al. 2001), creating heterogeneous landscapes containing patches of primary tropical forest surrounded by areas of intensive land use and secondary forests. Abandoned areas where forests regenerate may hold subsets of forest-dependent species; as such, the conservation value of human modified landscapes has gained increasing attention (Daily et al. 2001, Wright and Muller-Landau 2006, Marris 2009, Didham 2011, Mendenhall et al. 2014). Secondary forests are important elements in human modified landscapes, because they are expected to grow in overall area in the tropics, and may support a relatively high percentage of primary forest species (Perz and Skole 2003, Barlow et al. 2007). A complicating factor in understanding the biodiversity value of secondary forests, however, is the idiosyncratic and highly dynamic nature of these systems. The regeneration process may be influenced not only by time, but by other factors such as land-use history, distance from seed-dispersal sources, and soil conditions (Mesquita et al. 2001, Wright 2010). Ultimately, increased structural complexity of the vegetation can be a driver of space use for many forest-dependent species (Verschuyl et al. 2008). The interplay between secondary forests and primary forest also factors in; community composition in fragments is strongly related to the habitat quality of the surrounding matrix, because it may facilitate fragment colonization (Laurance et al. 2011). Matrix is defined here as habitat surrounding forest remnants at any stage of regeneration. Additionally, configuration of these landscapes may influence community composition inside habitat remnants. For example,
edge effects may exacerbate area effects on the biota inside forest fragments (Banks-Leite et al. 2010). Such negative effects, however, may be modulated by matrix quality. Given a regeneration process of the matrix, forest-dependent species may eventually inhabit the areas surrounding fragments (Ferraz et al. 2007, Stouffer et al. 2011).

Understanding the biodiversity value of landscape elements in heterogeneous landscapes is challenging. Stationary sampling-based inventories provide information on species presence in human-modified landscapes (Barlow et al. 2007). A serious limitation, however, stems from making inferences on highly mobile animals, because they can switch between different landscape elements (Hansbauer et al. 2010). Given the amount of regeneration in secondary forest along edges with primary forest, individuals might expand their home range into once inhospitable areas and use them more frequently as secondary growth develops. On the other hand, very sensitive species might be repelled by edges, with edge avoidance behavior pushing territory boundaries back into the forest interior. Even for animals using edges or second growth vegetation, movement may be greatly affected. Therefore, spatial behavior of forest-dependent animals in heterogeneous landscapes contain three aspects: first, interactions with forest edges, which may provide clues as to how the animals perceive the neighboring habitat; second, avoidance or inclusion of matrix habitats into their home ranges; and third, changes in movement patterns within the matrix.

To understand space use and movement patterns, there are different ways to use spatial data. More commonly, and on a larger-scale, patterns of space use can be acquired through a compilation of animal locations that generate a utilization distribution (Seaman and Powell 1996). This allows a depiction of how home range boundaries are set and how the spatial activity within its limits is influenced by the landscape. Recent developments in analytical techniques
consider finer scales. Animal movement trajectories reflect the interaction of an organism with the landscape along a time series, thus capturing their immediate behavior towards landscape elements (Calenge et al. 2009b). For example, movements typically become faster and straighter when an individual is in an inhospitable matrix (Hodgson et al. 2011), and some animals may change their orientation when approaching forest edges (Uriarte et al. 2011).

In central Amazonian forests, the effects of forest fragmentation on understory birds are well-studied from the perspective of birds using forest fragments (Robinson 1999, Sigel et al. 2006, Lees and Peres 2008, Martensen et al. 2012) especially at the Biological Dynamics of Forest Fragments Project (BDFFP) in central Amazonia (Stouffer et al. 2011); despite that, there is still an overarching need to understand how birds use the matrix, as it strongly affects community dynamics in fragments (Stouffer et al. 2006). A recent study at the BDFFP showed that movements across fragment boundaries recovered in a high proportion of understory avian guilds to pre-isolation rates at around 26 years as the surrounding matrix regenerated into secondary forests (Powell et al. 2013). Little is known about spatial behavior of individuals outside these fragments, such as how long they remain in the matrix, how far they enter it, or if their movements differ in any significant way.

Among the most sensitive guilds to habitat disturbance, mixed-species flocks form a significant portion of the insectivorous understory avian community in Amazonian tropical forests (Sodhi et al. 2011). Amazonian understory flocks consist of a year-round core of about 10 insectivorous species with overlapping territories led by the antshrike *Thamnomanes caesius* (or its allospecies in Bolivia and Peru, *Thamnomanes schistogynus*). Each core species in a flock is represented by one reproductive pair that defends its territory from conspecifics (Munn and Terborgh 1979). These core species are regularly joined by a long list of species, resulting in
aggregations of 50 species or more (Jullien and Thiollay 1998). Individuals in mixed-species flocks are active foragers, moving during all daylight hours actively seeking prey (Powell 1985). Core species gather in the same location every day at dawn and move through their territory, eventually returning to the vicinity of the gathering point, where individuals roost within about 50 m from each other (K. Mokross pers. obs.) in similar fashion to Central American flocks (Gradwohl and Greenberg 1980). The area used and defended by core members is about 10 hectares, with a negative relationship between total area and foliar density (Jullien and Thiollay 1998). Data from rainforests in French Guiana, Peru, and Brazil show near continuous coverage of abutting understory mixed-species flock territories (Munn 1985, Johnson et al. 2011), which tend to be highly stable in time (Martinez and Gomez 2013).

Amazonian mixed-species flocking species are heavily forest-dependent, reluctant to cross even narrow open areas (Develey and Stouffer 2001) and disappear in selectively logged forests (Thiollay 1997) and small fragments (Stouffer and Bierregaard 1995). Despite that, some species have been detected in secondary forest (Antongiovanni and Metzger 2005). It is an open question as to whether the species detected in secondary forest are forming stable aggregations with stable territories. Based on following flocks, I assess the spatial behavior of mixed-species flocks in a heterogeneous landscape in central Amazonia, and ask three questions across three scales, trajectory, home range, and landscape:

A) At the trajectory scale, do flock movement patterns differ between heterogeneous landscape components? More specifically, are there edge effects on movement directionality and do movement patterns differ among primary forest, edges, and secondary forest? Studies show the negative effects of edges on Neotropical forest dependent birds (Laurance et al. 2004, Ming et al. 2005, Banks-Leite et al. 2010), and Amazonian mixed-species flocks have been shown to
avoid crossing even narrow forest openings such as roads (Develey and Stouffer 2001).

Differences in the movement patterns of species near forest edges and after entering the matrix have also been recorded for other forest-dependent organisms (Crist et al. 1992, Levey et al. 2005, Hansbauer et al. 2008, Uriarte et al. 2011). I expect that flocks will tend to avoid forest edges, favoring movements away from forest edges, towards the fragment or continuous forest interior. I also expect movement patterns (i.e. turning angles and step lengths) will change inside secondary forest, tending towards faster and more rectilinear trajectories associated with less successful foraging (Lima and Zollner 1996). B) At a home range scale: how are flocks home ranges next to forest edges and how do they vary according to vegetation height? Disruption of original habitat can strongly influence forest-dependent animals home ranges and restrict them to forest remnants when the matrix is cleared of all vegetation (Rolstad 1991); as the surrounding vegetation regenerates, understory birds can expand their movements into previously inhospitable areas (Stouffer et al. 2011, Powell et al. 2013). I hypothesize that flocks avoid less-developed secondary forests and enter better-developed secondary forests using it more frequently and more extensively. The elastic disk hypothesis states that animal home range sizes tend to become smaller with higher resource densities (Huxley 1934). This has been demonstrated in Amazonian mixed-species flocks in Guyana through a qualitative categorization of vegetation types and home range sizes; proportion of denser habitats correlated with smaller home ranges (Jullien and Thiollay 1998). It is possible that the amount of vertical vegetation may influence resource density and thus home range size. On the other hand, the strong stratification of the Amazonian avifauna may prevent full utilization of the vegetation column (Walther et al. 2002). I expect that home range area will be larger in less-developed secondary forest, which tends to present lower vegetation heights. C) At a landscape scale, I ask how home range
dispositions varies in different landscape elements such as forest fragments of different sizes and flocks inhabiting areas next to secondary forests. Animals show different intensities of use within their home ranges, and this is a strong factor that shaping overall disposition of home ranges in the landscape. I expect that intensity of use will be correlated with vegetation height and that home range limits will be delimited by forest borders and areas with less-developed secondary forest.

With these results, I attempt to link behavioral processes to space use patterns in heterogeneous landscapes to address biodiversity value of secondary forests for a significant and vulnerable component of the Amazonian rainforest avifauna.

METHODS

Study Site

The Biological Dynamics of Forest Fragments Project (BDDFP) is located at about 80 km north of the city of Manaus, in the state of Amazonas, Brazil. Fragments were isolated between 1980 and 1990 during the creation of cattle pastures and were characterized in three categories, 1, 10, and 100-ha fragments (Lovejoy et al. 1986, Bierregaard and Gascon 2001). Following initial clear-cutting, some areas were burned to create pastures while others were abandoned. Vegetation along larger streams was left standing at 30-40 m strips, creating a mixture of primary forest patches connected to continuous primary forest and secondary forest. Most pastures were inactive by the 1990s, allowing the matrix around the fragments to regenerate. A buffer of 100 m was cleared around some fragments from the early 1990s to the early 2000s, but these areas have regenerated as well. The resulting landscape is structurally heterogeneous due to different management histories (Mesquita et al. 2001) with fragments of different sizes surrounded by secondary forests varying in structure and age (Fig.1). Due to land history use and
edaphic conditions, secondary forest structure at the BDFFP is not correlated to regeneration time, where some patches are still occupied by more open and shorter-stature vegetation after 30 years since clearing. Here I designate well-developed second growth as having a nearly continuous canopy with a mean height of >18m, in some cases, dominated by trees of the genus Cecropia. Less-developed second growth has a patchy canopy <18m, less dense understory and is typically dominated by trees of the genus Vismia (Borges and Stouffer 1999).

Figure1. A. Map of the BDFFP. White boxes represent areas where flocks were sampled. B. LIDAR coverage of one of these areas (Dimona ranch) and examples of home ranges of flocks in different landscape elements. From left to right: 10-ha fragment (yellow), secondary forest (red), primary-secondary forest mix (white) and 100-ha fragment (lime green), other studied flocks are shown lighter shades than the examples. Flocks are generally absent in areas of scrubby vegetation (light grey), such as the buffers around fragments. The flock with a red asterisk was considered as unstable as it was not detected in subsequent sampling periods. Darker colors represent areas of taller vegetation based on Lidar data.
Data collection

Data were collected between June 2009 and August 2011. Flock territories were located in five main habitat configuration types: primary forest (interior and edge); 100-ha fragments (interior and edge); 10-ha fragments; primary-secondary forest mix, which are strips of primary forest that are connected to continuous forest; and secondary forests (Table 1 and Figure 1). Flock activity is conspicuous, enabling them to be followed on foot. The observer (KM) maintained a distance of 15-20 meters from the core of activities. As flocks moved, the observer’s positions were recorded at 30-second intervals with a hand-held GPS unit (Garmin Vista HCX).

Trajectory analysis

To test the influence of edge effects on movement directionality, I used 2-minute intervals between relocations to reduce noise from GPS errors. The effects of forest edges is strong within 20 meters from forest edges such as water vapor deficit, temperature and foliage density (Kapos 1989, Laurance et al. 2002), therefore I gathered the relocations within a 40-m-wide zone, 20 m on each side of the forest border, henceforward called edge. To analyze trajectories, I compiled the absolute angles of the relocations inside the edge, doubled the angles (Batschelet 1981) and applied a Rayleigh test of uniformity (V0 test) in the CircStats R package (Lund 2012) to compare them to the axial angle of the forest edge. I also quantified the average duration and distance of forays into secondary forest by flocks to assess flock behavior next to the matrix in each habitat configuration type.

Movement pattern analyses were performed with the modpartltraj function in the AdehabitatLT library (Calenge et al. 2009b) in the R software. Using Markov models, this function partitions an animal’s trajectory according to different movement behaviors (Calenge et
al. 2009a). I defined three movement behaviors with distinct parameters based on field observations (Table 1): (1) area-restricted behavior (ARB), i.e. when flocks foraged in a small area, reducing their speeds and moving with tight angles; (2) normal behavior, i.e. when flocks foraged at a regular speed, but turned at constant frequency; and (3) fast behavior, i.e. when flock speed was higher than average, and turning angles were high, resulting in ballistic trajectories covering great distances in short time intervals. For this analysis, I only used trajectories that were larger than 50 relocations.

Table 1. Parameters used in the \textit{partlmod} function to define the three movement modes. Step length distributions are normal, while I used a Von Mises distribution for turning angles.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Step length</th>
<th>Angle Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>Wrapped cauchy</td>
</tr>
<tr>
<td>Fast</td>
<td></td>
<td>(\mu = 0), (\rho = 0.5)</td>
</tr>
<tr>
<td>Normal</td>
<td></td>
<td>(\mu = 0), (\rho = 0.5)</td>
</tr>
<tr>
<td>ABR</td>
<td></td>
<td>(\mu = 180), (\rho = 0.5)</td>
</tr>
</tbody>
</table>

I analyzed the proportions of these three movement modes in the three main landscape elements: primary forest (PF), secondary forest (SF), and edge (ED) (20 meters on each side of the limit between primary forest and secondary forest). To probe the behavior of flocks in more detail, each landscape type was categorized into the following features. Primary forest consisted of continuous forest (CF), Primary-secondary mix (MIX), 100-\(\text{ha}\) fragments (100ha) and 10-\(\text{ha}\) fragments (10ha). Edges were subdivided into soft edges (S), when there was a more gradual transition between PF and SF without trails, or hard edges (H), when PF shifted abruptly to SF and a trail separated the elements. Secondary forest was characterized as well-developed secondary forest (Wd) when mean vegetation height was over 15m below 23m, and less-developed secondary forest (De) when mean vegetation height was below 15m (table 5).
For each flock in each landscape feature, I obtained the proportion of movement modes and tested the differences between the observed movement mode proportions in each landscape feature using multinominal regression through the mlogit package in R (Croissant 2013). The dependent variable was defined as the movement mode chosen at each step and the landscape features as individual-specific variables, each flock grouped as an individual (id.var) and within flock, each step as the choice ID (chid.var).

Home ranges next to secondary forests

To map intensity of use within the home range, recorded positions were used to create a quadratic kernel using Geospatial Modeling Environment (GME) software (Beyer 2012), at 1m resolution, 275 bandwidth (i.e., radius of a circle in which points are counted around each pixel) at default scaling factor (i.e., a value that the point density values are multiplied for scaling). I chose these settings because the kernels would be conservative when considering total home range extent. Isopleths were created from kernels at intervals ranging from 99% (entire home range area) to 10% (areas of highest location densities) (Fig.2). I used a LIDAR (Light Detection and Ranging) canopy height model (CHM) (Electronic Supplementary Material, Lidar specifications), and generated the zonal statistics for the vegetation located inside each isopleth, which provided mean vegetation height (Electronic Supplementary Material, Table 1).
Figure 2. Depiction of a flock home range (FL_IB) adjacent to secondary forest. Home range is represented by isopleths ranging from 10% (center in white) to 99% (edges of home range in yellow). Black dots are locations taken at 30-second intervals. The section of home range occupying secondary forest is shaded in dark gray. Red line shows the edge between primary and secondary forest. Height of individual trees appears in green, with darker green representing taller trees.

To quantify the time flocks spent in second growth, I created a polygon over the area of secondary forest and applied a count points in polygon (countpntsinpolys) function in GME for each flock home range. Additionally, I counted the positions inside each isopleth to quantify percentage of time spent in secondary forest. I also extracted the mean vegetation height in every isopleth of the home ranges using the intersect polygon with raster (isectpolyrst) function in GME. To test whether vegetation height was correlated to time spent in secondary forest, I used a linear model (lm) in R software (R Core team 2014) in the portion of the 99% isopleth located in secondary forest.

I tested if home range area was correlated to mean vegetation height by using the same approach, but using the entire home range instead of only the area of secondary forest. For home range size analysis, I also included flocks entirely located in primary forest or secondary forest.
**Home range layout in the landscape**

To test if vegetation height influences the shapes of home range kernels, I averaged canopy height values in 10 X 10-m squares due to the high small-scale variance and constructed a model of space use. The probability of using a particular square $x$ is modelled to be proportional to $f(x|\alpha) = \exp[\alpha C(x)]$, where $C(x)$ is the canopy height and $\alpha$ is a model parameter. The null model is $\alpha=0$, meaning that any square is equally likely to be used as any other. I tested this against the alternative hypothesis that there is some $\alpha>0$ that significantly improves the fit of the model to the data. I used a maximum likelihood approach, seeking to find the alpha that minimises the sum of $\ln[f(x_n|\alpha)]$ over a set of independent fixes $x_n$. This sum is the log-likelihood function $l(x_1, ..., x_N|\alpha)$ where $x_1, ..., x_N$ is the set of independent fixes being used for the test. For this analysis I used the one of the Lidar sets that encompasses the Dimona data set (Table 1), which was the best-sampled area. Because each flock gathers in the same point at the start of each day, each day’s path of motion within the territory is independent of the previous days’. Therefore, I can assume that a recorded position of a flock on one day is independent from a position recorded on another day so I let each $x_n$ be a randomly selected position from a single flock on a single day. To avoid bias from the random selection, I considered all possible sets of such positions $x_1, ..., x_N$ and took the average of the various $l(x_1, ..., x_N|\alpha)$. I denoted this average by $L(x_1, ..., x_N|\alpha)$.

**RESULTS**

I sampled 26 flocks totaling 700 hours of observation in six different landscapes within the BDFFP (Table 1). There were 14 flocks located near edges adjacent to secondary forest; of these, three were inside 10 ha-fragments, six near 100 ha-fragment edges, three near continuous primary forest edges, and two in primary-secondary forest mix.
Of the 14 flocks located next to edges, only 10 could be analyzed in regards to detailed vegetation structure, as there was no Lidar information for flocks located in the Porto Alegre landscape (Table 2).

Table 2. Understory mixed-species flocks sampled in this study. Flocks are arranged by area within the BDFFP, habitat types are: 100-ha/edge: flocks located in 100-ha fragments near edges; 100-ha/interior: flocks in 100-ha fragments at the center of the fragment with no contact with forest edge; 10-ha: flocks in 10-ha fragments; prim./sec.: flocks in areas with patches of primary and secondary forest; primary edge: flocks in areas of primary forest with interface to secondary forest; primary interior: flocks in primary forest with no contact to forest edge; secondary forest: flocks in which the entire home range occupies secondary forest. Total hours sampled indicate the amount of time for the dry season.

<table>
<thead>
<tr>
<th>Area</th>
<th>Flock</th>
<th>Habitat type</th>
<th>Total Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimona</td>
<td>Cap_II</td>
<td>second growth</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Cap_N</td>
<td>second growth</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Central</td>
<td>100-ha /interior</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Dim_10ha</td>
<td>10-ha</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Ig_cmp_flk</td>
<td>prim. /sec.</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Lake_flk</td>
<td>100-ha /border</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>South_Central</td>
<td>100-ha /interior</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Southwest</td>
<td>100-ha /border</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>W400</td>
<td>100-ha /border</td>
<td>39</td>
</tr>
<tr>
<td>Colosso</td>
<td>Col_10ha</td>
<td>10-ha</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Col_cabfríoI</td>
<td>primary border</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Col_Cap1ha</td>
<td>second growth</td>
<td>46</td>
</tr>
<tr>
<td>Florestal</td>
<td>Fl_a</td>
<td>prim. /sec.</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Fl_I_b</td>
<td>primary border</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Fl_II</td>
<td>primary border</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Fl_III</td>
<td>primary border</td>
<td>21</td>
</tr>
<tr>
<td>Gavião</td>
<td>Gav_10_ha</td>
<td>primary interior</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Gavião_I</td>
<td>primary interior</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Gavião_II</td>
<td>primary interior</td>
<td>20</td>
</tr>
<tr>
<td>Km37</td>
<td>Km37_III</td>
<td>primary interior</td>
<td>20</td>
</tr>
<tr>
<td>Porto Alegre</td>
<td>PA_10ha**</td>
<td>10-ha</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>PA1**</td>
<td>100-ha /border</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>PA2**</td>
<td>100-ha /border</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>PA3**</td>
<td>100-ha /border</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>PA4**</td>
<td>100-ha /interior</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td>700</td>
</tr>
</tbody>
</table>

* Unstable flock - see text. ** Flocks located in *Cecropia* secondary forest and were not located under Lidar coverage, therefore could not be analyzed in regards to vegetation height.
Nine flocks were located inside primary forest; of which six were inside continuous primary forest, and three were inside 100 ha fragments. Lastly, three flocks were located entirely within secondary forest.

**Trajectory analyses**

As a whole, flocks moved at an average rate of 3.7 m/min. (n=17699, sd=5.00). Individuals within the flock advanced in different directions from the main movement vector of the aggregation, ranging in height mostly between 1 and 15 m. *Thamnomanes caesius* usually remained in the center of activity, and periodically moved away and making calls, which seemed to attract the individuals in a desired direction. Occasionally, individuals did not follow *T. caesius*, prompting it to move back and attempt another direction. Flocks seemed to avoid open areas and usually circled around large forest gaps.

On few occasions however, most individuals would quickly cross open areas of up to 25 m in length. Apart from these rare events, flocks generally did not enter vegetation < 5 m high.

Flock trajectories tended to be perpendicular to hard edges adjacent to less-developed secondary forest (Fig. 3). Four out of ten flocks showed this response with significant Rayleigh test values (Table 3): two flocks in 10-ha fragments and two in continuous forest edges, both adjacent to less-developed secondary forest. Flocks adjacent to more developed second growth did not move consistently parallel to forest edges. Flocks in 100-ha fragments generally showed little activity along edges, and when doing so, did not show any directional trend.
Figure 3. Trajectory angles in relation to forest edges for a 10-ha fragment flock (Dimona). North and South edges are highlighted in blue, while east and west facing edges are highlighted in red. The rose diagrams indicate the distribution of directions for each edge set, highlighted in the same color code.

Table 3. V test results of absolute angles of trajectory steps within 20 m from forest edge in relation to forest edge angle. Rbar is the mean resultant length, MuO the axial angle of the border that flock absolute angles at each step are being compared to, Total sampled column indicates the number of steps that were used in the analysis, % in EDGE indicates the amount of time spent within 20 m of primary-secondary forest interface. Bold characters indicate significant results.

<table>
<thead>
<tr>
<th>Flock</th>
<th>Rbar</th>
<th>p</th>
<th>MuO</th>
<th>N</th>
<th>% in EDGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Col10ha</td>
<td>0.093</td>
<td><strong>0.000</strong></td>
<td>328</td>
<td>726</td>
<td>25.7</td>
</tr>
<tr>
<td>Dim10ha</td>
<td>0.064</td>
<td><strong>0.007</strong></td>
<td>36</td>
<td>681</td>
<td>19.4</td>
</tr>
<tr>
<td>FLIA</td>
<td>0.104</td>
<td><strong>0.019</strong></td>
<td>298</td>
<td>202</td>
<td>31.4</td>
</tr>
<tr>
<td>FLIB</td>
<td>0.220</td>
<td><strong>0.033</strong></td>
<td>334</td>
<td>36</td>
<td>5.9</td>
</tr>
<tr>
<td>Igcmpflk</td>
<td>0.025</td>
<td>0.175</td>
<td>310</td>
<td>707</td>
<td>23.5</td>
</tr>
<tr>
<td>Lake</td>
<td>0.069</td>
<td>0.263</td>
<td>216</td>
<td>65</td>
<td>10.6</td>
</tr>
<tr>
<td>FLIII</td>
<td>0.024</td>
<td>0.341</td>
<td>334</td>
<td>146</td>
<td>22.8</td>
</tr>
<tr>
<td>SW</td>
<td>0.027</td>
<td>0.393</td>
<td>282</td>
<td>50</td>
<td>8.1</td>
</tr>
<tr>
<td>FLII</td>
<td>-0.019</td>
<td>0.672</td>
<td>334</td>
<td>294</td>
<td>11.4</td>
</tr>
<tr>
<td>w400</td>
<td>-0.128</td>
<td>0.957</td>
<td>36</td>
<td>93</td>
<td>7.6</td>
</tr>
</tbody>
</table>
Two of the 10-ha fragments surrounded by less-developed secondary forest dominated by *T. caesius* showed flock movements into second growth consisted of forays that lasted 9.2 minutes on average (s.e.=2.07, max=82.0) covering distances up to 20 m from the forest edge and always returning to the fragment (Fig. S1). One flock located in a 10 ha-fragment surrounded by well-developed secondary forest dominated by *Cecropia* did not enter secondary forest after 18 hours of observation. Flocks in 100-ha fragments surrounded by well-developed secondary forest (*Cecropia*) used it more frequently and ventured farther away from primary forest than did flocks in a 100-ha fragment surrounded by less-developed (*Vismia*) second growth. In the 100-ha fragment surrounded by *Vismia*, flocks rarely entered secondary forest (4.2% out of 79 hours) and did not move more than 50 m away from the border with forays lasting 11.47 minutes on average (s.e.=5.13, max=62.0); in the 100-ha fragments surrounded by *Cecropia*, flocks entered secondary forest an average of 16% of the sampled time (total of 28.2 hours), moving away from the border up to 65 meters in forays that lasted up 10.37 minutes on average (s.e.=4.71, max=68.5). In primary-secondary forest mix, flocks tended to move further away from primary forest, up to 208 meters from primary forest, lasting 20.27 minutes on average (s.e.=5.13, max=230.0). Flocks next to soft edges near well-developed secondary forest moved up to 175 m away from primary forest in forays that lasted 21.9 minutes on average (s.e.=12.12, max=32). Collectively, flocks near edges showed reluctance to cross forest edges, however, some species such as *Thamnomanes caesius*, *Xiphorhynchus pardalotus* and *Xenops minutus* were more prone to transition into secondary forest, whereas other species such as *Myrmotherula longipennis*, *M. Menetriesii* and *Thamnomanes ardesiacus* tended to remain behind.

In order to assess the movement modes inside landscape features, the *modparltraj* function was applied to 121 different trajectories for 10 flocks located near forest edges (Fig. 4). The
multinomial model was statistically significant (Likelihood ratio test: chi square 1868.6, p<0.001); magnitudes ranged mostly between -0.5 and 1, with the exception of fast movements in 10ha forests and hard edges. Overall, most movement was in Normal mode (45.6%), followed by ARB (30.8%). The least observed movement mode was fast (23.5%). Under model predictions, fast movements tended to increase in proportion inside 10 ha-fragments and hard edges (Fig. 5). Fast movements increased slightly in young secondary forests compared to old, but their overall proportions were not consistently higher than what is predicted for forest and edge areas.

**Home range analyses**

All flocks located near edges entered secondary forest, but the time spent there was positively and significantly correlated to the mean vegetation height in secondary forest (r² = 0.519, F1,8 = 8.64, p = 0.0187)(Fig. 6A). Despite higher use in well-developed secondary forests, the majority of flocks did not concentrate their activities in those areas. Instead, their home range cores were largely in primary forest (Table 4 and Fig. 2). In fact, all of the gathering locations where flock core members met at sunrise were in primary forest. The only exception was one flock located next to the best-developed area of secondary forest, where its home range core was located in the primary-secondary forest interface (Table 2, FLIII). Despite a strong correlation between time spent in secondary forest and mean vegetation height, the area of the home range occupying secondary forest was not correlated to mean vegetation height (r² = 0.221, F1,8 = 2.274, p = 0.17).
Table 4. Percentage of time mixed-species flocks spent in secondary forest. First column (ISOP) shows kernel isopleths decreasing from complete home ranges (99%) to areas of most intense use (10%). Cells are color-coded: higher percentages of time in secondary forest are darker. First row shows mean vegetation height in secondary forest within 99% isopleth. Second row are flock IDs ordered from lowest to highest percentage of total home range area in second growth. Notice that areas of highest intensity of use are not present in most flocks.

<table>
<thead>
<tr>
<th>Hght.</th>
<th>6.9</th>
<th>12.2</th>
<th>9.2</th>
<th>6.4</th>
<th>21.7</th>
<th>21.8</th>
<th>7.1</th>
<th>16.0</th>
<th>14.5</th>
<th>20.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISOP</td>
<td>Lake</td>
<td>SW</td>
<td>W400</td>
<td>Dim_1</td>
<td>FLII</td>
<td>FL_IB</td>
<td>Col_1</td>
<td>Igemp</td>
<td>FL_IA</td>
<td>FLIII</td>
</tr>
<tr>
<td>99</td>
<td>2</td>
<td>3.7</td>
<td>7.2</td>
<td>8.5</td>
<td>15.9</td>
<td>16.3</td>
<td>20.8</td>
<td>27.5</td>
<td>30</td>
<td>37.3</td>
</tr>
<tr>
<td>95</td>
<td>1.8</td>
<td>3.6</td>
<td>7.2</td>
<td>7.2</td>
<td>15.2</td>
<td>16.1</td>
<td>20.8</td>
<td>26</td>
<td>29.7</td>
<td>37.7</td>
</tr>
<tr>
<td>90</td>
<td>1.9</td>
<td>3.5</td>
<td>6.4</td>
<td>5.8</td>
<td>14.4</td>
<td>15.3</td>
<td>20.7</td>
<td>23.7</td>
<td>28</td>
<td>38.7</td>
</tr>
<tr>
<td>80</td>
<td>1.7</td>
<td>2.5</td>
<td>3.8</td>
<td>4.3</td>
<td>12.6</td>
<td>12.4</td>
<td>19.8</td>
<td>19.5</td>
<td>24.6</td>
<td>40.4</td>
</tr>
<tr>
<td>70</td>
<td>0.8</td>
<td>2.7</td>
<td>2</td>
<td>2.9</td>
<td>11.7</td>
<td>8.1</td>
<td>18.9</td>
<td>16.6</td>
<td>20.8</td>
<td>41.6</td>
</tr>
<tr>
<td>60</td>
<td>0</td>
<td>1.9</td>
<td>0</td>
<td>2.3</td>
<td>10.1</td>
<td>2.9</td>
<td>18.7</td>
<td>8.8</td>
<td>17.8</td>
<td>45.4</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>0.6</td>
<td>0</td>
<td>1.2</td>
<td>8.1</td>
<td>1.1</td>
<td>19.2</td>
<td>4</td>
<td>14.5</td>
<td>52.7</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>8.3</td>
<td>0</td>
<td>15.9</td>
<td>3</td>
<td>10.3</td>
<td>55.2</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8.8</td>
<td>0</td>
<td>13</td>
<td>1.3</td>
<td>7.5</td>
<td>58.4</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8.7</td>
<td>0</td>
<td>10.5</td>
<td>0</td>
<td>6.4</td>
<td>71.6</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6.8</td>
<td>92</td>
</tr>
</tbody>
</table>

Overall, total home range area tended to have a negative correlation to mean vegetation height ($r^2 = 0.31, F_{1,8} = 7.638, p = 0.013, \text{Fig. } 3B$), where the smallest home range, 6.7 ha, was in an area of continuous primary forest with a canopy height average of 23.4 m (flock CaboFrio), and the largest home range, 17.1 ha, was in an area of secondary forest with a canopy height average of 13.1 m (flock Cap1ha).
Figure 4. A. Example of a trajectory that was partitioned using the partmodltraj function. The blue triangle represents the beginning of the trajectory, a flock exits to secondary forest, eventually returning to primary forest while changing its movement patterns, the end of this trajectory is represented by a red square. Vegetation is shown by a Lidar-derived canopy height model, where darker shades represent taller vegetation. Movement models are color coded, Area restricted behavior (ABR) in blue, normal movement modes in yellow, fast movements are in red. This frame is a subset of the trajectories in the following panel. B. Primary-secondary forest mix (Ig. Cmp. Flk). C. Primary-secondary forest mix (FLIA). D. 10 ha-fragment (Dimona).
Figure 5. Prediction from the multinomial model on the proportions (represented as percentages) of movement modes inside each landscape element. **PF**: Primary forest, **ED**: Edge (20 m at each side of forest-secondary forest interface), **SF**: Secondary forest. Under **PF**, the types are **CF**: Continuous forest, **10ha**: 10ha-forest fragment, **100ha**: 100ha-forest fragment, **MIX**: primary-secondary forest mix. **ED** types are **S**: soft edges and **H**: hard edges. **SF** types are **Wd**: Well-developed secondary forest and **De**: Less-developed secondary forest. Blue color represent partitions categorized as *Area Restricted Behavior*, red Colors indicate *Fast* movement behavior and Yellow is *Normal* movement behavior. The bars with black outlines represent the averaged proportions of movement modes in each landscape element.

**Home range layout in the landscape**

No flocks were found in 1-ha fragments. Each 10-ha fragment held one home range; the 100-ha fragment held about 10 flocks (Figure 1). All flocks in 100-ha fragments were located adjacent to edges, except for two flocks in the center of the fragment surrounded by neighboring flocks on all sides. Flocks in primary-secondary forest mix centered their home ranges in areas with primary forest, extending the margins towards secondary forests. There were flocks inhabiting secondary forest with no contact to primary forest, but the aggregation in the least developed area of secondary forest (i.e. with lowest mean vegetation height) was not present throughout the entire sampling period. This flock was found in June 2010 being led by only one
female *Thamnomanes caesius*, the same individual plus a male were seen in October 2010, and neither were found in May 2011 after several visits to their former gathering area.

Flock kernel shapes responded to the disposition of anthropogenic features. The response of home range shapes to vegetation, the value of $\alpha$ that maximizes $L(x_1, \ldots, x_N|\alpha)$, is $\alpha = 0.065$. Using the likelihood ratio test, the p-value associated with rejecting the null hypothesis in favor of the hypothesis that canopy height is a predictor of space use is $p = 0.0000033$. Regarding edge effects on home range shapes, this value of alpha means that if an area of primary forest of height 30m is next to an area of secondary forest of height 10m, then birds near the boundary between the two areas will be $\exp[0.065 \times (30 - 10)] = 3.7$ times more likely to be found in the primary forest.

![Figure 6](image)

**Figure 6.** A. For flocks located adjacent to secondary forest: relation between time in secondary forest and mean vegetation height of second growth. B. Relationship between total home range area and mean canopy height.

**DISCUSSION**

The combination of fine-scale movement in relation to vegetation height with habitat configuration extended our predictive power on how mixed-species flocks may occupy human modified landscapes. On a trajectory level, flocks tended to avoid areas with vegetation $<10m$ height and entered less-developed areas of secondary forest with some reluctance (i.e. canopy
heights <15m) but only as relatively short forays which inevitably turned back to the areas of primary forest. Flocks tended to move parallel to either side forest edges that were located next to less-developed areas of secondary forest; the effects of edges on movement directionality tended to disappear with decreased vegetation height discrepancies. The duration and distance of these forays into secondary forest are likely influenced by habitat configuration. For example, flocks in 10-ha fragments used secondary forest less frequently than other habitat configuration types regardless of development of the surrounding matrix. Flocks in 100-ha fragment edges entered it for longer duration and distance than the former. The pressure from neighbors inside 100-ha fragments could be pushing flocks near edges to explore secondary forests, while flocks in 10 ha-fragments were relatively isolated from neighbors in an area that supplied sufficient habitat for a typical home range size. Flocks located in primary-secondary forest mixes showed much longer forays covering longer distances within the matrix, suggesting more pressure to utilize secondary forest due to the little primary forest habitat available.

In regards to movement modes, it was expected that flocks would move faster and in more rectilinear fashion in secondary forests and more slowly and tortuously in primary forests. Differences were not as stark, despite indications that certain movement modes are more frequently adopted in determined landscape elements. The switching between these movement modes seems to be more complex, and the characteristics of landscape elements seems to exert some influence. Ultimately, fine scale adjustments to resource patches seems to play a much more important role.

At the home range scale, flocks occupying areas near edges exhibited a relationship between the proportion of time spent in secondary forests and forest structure, here expressed by canopy height. This suggests that during a regeneration process, flocks located near edges tended to enter
and use secondary forest more. Home range cores, however, were not located in secondary forest until vegetation reached at least 20 m of mean canopy height. Our results also suggest that that despite higher frequency of use, the area occupied in secondary forest did not necessarily become larger. This reinforces the possibility that area of introgression into secondary forest may be more related to habitat configuration and territorial interactions with neighboring flocks rather than secondary forest vegetation height. Another evidence related to how flocks use areas lower vegetation; flocks in secondary forests tended to occupy nearly double the area recorded in primary forest flocks.

At a landscape scale, flocks were not present in 1-ha fragments, most likely because of the low surrounding vegetation resulted by re-isolation of the fragments. On the other hand, 10 ha-fragments were occupied by one home range interacting with forest edge on all sides. Fragments of 100 ha held 10 to 11 flock home ranges encapsulated by fragment borders. The majority of flocks were near edges while two flocks were located at the core. There were home ranges entirely located in secondary forest, but below a mean vegetation height threshold >15 m, flocks seem to be unstable as key members are not consistently present in these areas. Overall, Amazonian mixed-species flocks are occupying the landscape according to the imposed anthropogenic features; home range boundaries tend to be located near fragment edges or areas with less-developed secondary forests. During development of the matrix, once vegetation surpasses an average of 20m, the spatial patterns resemble what is observed in primary forests, however this process is very likely to take longer in areas with history of intensive land use.
Trajectory analysis

Flocks responded to forest edges next to less-developed secondary forests with movements along the interface for extended periods. A possible mechanism underlying this behavior is due to forest edges receiving higher amounts of sunlight, and within a relatively short period from isolation, present higher vegetation density accounting for a high use frequency (Didham and Lawton 1999, Laurance 2004). Occasionally, flocks moved into secondary forest and foraged within 30 meters parallel to the forest edge. This behavior may be possibly due to arthropod spillover from primary forest into adjacent secondary forest (Lucey and Hill 2011). The amount of interaction with forest edges seems to be influenced by a combination of higher vegetation density and micro-topography. Flocks are more likely to move into lower terrain and use higher portions of the topography less frequently (Potts et al. 2014b). Forest edges located on slopes or lower areas are visited more frequently, while higher areas were not visited at all in some circumstances. The mechanisms for higher frequency of use of these distinct landscape features remain to be investigated. A possible explanation lies in a trade-off in resource amounts between the much dryer conditions in fragment edges (Kapos 1989) and higher vegetation density (Laurance and Gomez 2005), as the capacity of vegetation to buffer the microclimatic conditions have been shown to be considerably reduced with habitat fragmentation (Ewers and Banks-Leite 2013).

The fact that certain species remained behind when Thamnomanes caesius led flocks into underdeveloped secondary forest may explain part of the observed spatial behavior near forest edges. Despite individuals of T. caesius being able to cross narrow clearings with relative ease (Laurance and Gomez 2005), when a substantial portion of flock elements became reluctant in heading to a determined area, it tended return and resume calling the flock towards another
direction (K.M, personal observation). This means that flock collective behavior may be significantly affected. Also, the fact that some species showed reluctance to enter secondary forests may help explain some of the differences in flock composition in secondary forest (Mokross et al. 2014).

The analysis on movement modes did not show strong effects on movement modes (Fast, Normal and ARB) in primary or secondary forests. The only patterns standing out were that flocks tend to move predominantly in Fast mode in 10ha fragment interior and along hard forest edges next to less-developed secondary forest, and in ARB mode along soft edges next to well-developed secondary forest. Vegetation height may explain overall space use at home range scales, but not significantly on trajectories. This might be due to the adaptive nature of foraging in mixed-species flocks elements despite the lower resource density, where there are possibly even finer adjustments in individual foraging strategies that would not be apparent at the scale of a trajectory (Newell et al. 2014). Flocks may have to cover larger areas, but the amount of search time on a resource-rich patch may eventually be the same. Interestingly, increased home range size is not determined by less tortuous flock trajectories in secondary forest. It seems that flocks are performing the same types of trajectories, but the area that is covered is more extensive.

Flocks next to forest edges use secondary forests; however, the intensity of use is conditional upon vegetation structure. Vegetation height averaging >20 m is visited at proportions that are close, but still not similar to what is expected in primary forest, whereas vegetation averaging <15 m may be visited at proportions lower than 10%. While I acknowledge that mean vegetation height may not be the proximal cause determining space use, it may efficiently aggregate indirect information about vegetation structure and seems to be a good predictor across avian species (Hinsley et al. 2002, Hyde et al. 2006, Clawges et al. 2008, Goetz et al. 2010). Even though
much of the foraging activity takes place in the understory, vegetation height seems to provide good indication of habitat quality, as flocks are shown to preferentially move to areas with taller tree cover (Potts et al. 2014a). In fact, other important properties of mixed species flocks, such as gregariousness, also seem to be well captured by this variable (Mokross et al. 2014).

Nevertheless, despite the more frequent use of well-developed secondary forest, most flocks still held their territory cores in primary forest, or in vegetation closer in structure to primary forest when no primary forest was available. The only exception to this was one flock located to a well-developed area of secondary forest (Flock FL III), which spent the majority of its time in secondary forest, suggesting that at later stages of development, secondary forest may become fully incorporated as part of a home range (Develey and Stouffer 2001). It is important to highlight that the area occupied in secondary forest is not proportional to time spent in secondary forest; I believe this might be due to territorial interactions between flocks as well as the inverse relation between home range area and vegetation structure (i.e., flocks occupy less area in more developed vegetation). Extensive areas of more developed secondary forests are occupied by flocks forming stable territories; however, total home range area tends to be considerably larger than the average of 8-10 ha that is usually recorded (Jullien and Clobert 2000, Develey and Stouffer 2001), suggesting a possible reduction in arthropod load per area. The relationship between home range size and vegetation density has been shown in previous work conducted in French Guyana (Jullien and Thiollay 1998), which used the proportion of different vegetation types. The present work uses an objective measure of vegetation structure which allows easier predictions of flock occupation in the landscape while also implying that vegetation height, as well as density, may help determine home range extension.
**Home range layout**

There was a gradually increased introgression of flock territories into regenerating secondary forest; however, during initial stages with vegetation <15m it seems to be only a marginal extension of home ranges. This suggests that the boundaries imposed by anthropogenic features gradually dissolve as the regeneration process takes place. The effects of these modifications in the landscape can persist even after 30 years given history of land use. The speed of forest recovery, and subsequently its structure is largely dependent upon stages of turning areas into pasture. Areas that were not burned or grazed showed a more similar structure to primary forest, and subsequently higher use frequency by mixed species flocks. On the other hand, areas that were burned and used as pasture have a Vismia-dominated vegetation that, despite having some stature, is much sparser.

Extensive areas of less-developed secondary forest (mean vegetation height >15 m) away from primary forest were temporarily occupied by unstable and depauperate flocks. The mechanisms underlying their disappearance remain to be investigated. Our data shows that areas with vegetation <10 meter mean height average are largely devoid of any flock activity. The mechanisms underlying this avoidance may be related to lack of resources, predation or both.

**Concluding remarks**

The relation between space-use by mixed-species flocks and vegetation height makes an intuitive statement on the alignment of their home ranges with anthropogenic features. It is a significant departure of the natural disposition of flock territory layout in a landscape, which is determined by topography, vegetation and territorial interactions (Potts et al. 2014a). This implies that a substantial portion of flocks may be forced into sub-optimal habitat by settling in areas that meet only the minimum requirement of vegetation height. This is evidenced by traits
outside of their spatial ecology. For example, flock composition and cohesiveness shows considerable differences in areas where spatial behavior recovers to what is observed in primary forests; overall flock participation is significantly reduced in 10 ha-fragments and in well-developed secondary forests (Mokross et al. 2014). Information on reproductive output and survival also remain to be assessed and may provide crucial information on the value of heterogeneous landscapes to this important component of the Amazonian understory.

A current limitation to understanding mechanisms driving flock space use is the dearth of data on the distribution and dynamics of arthropod prey in tropical forests and second growth. Considering its importance as a resource for not just birds, but for other vertebrates of conservation concern, I encourage research on landscape and structural effects on arthropod biomass in Amazonian forests.

In summary, the available data on flock spatial behavior in this heterogeneous landscape indicates that flocks may show space use patterns that are closely similar to primary forest at about three decades from disturbance if the area has a less-intensive land use history, corroborating results from Powell et. al 2013. Considering that most abandoned areas suffer intensive land use regimes (Nepstad et al. 2014), recovery to pre-disturbance space use patterns may take much longer than I show here.

Human modified landscapes still remain challenging to ecologists as highly dynamic conditions defy simple answers (Gardner et al. 2009) and ultimately require multifaceted approaches. Understanding individual space use in altered landscapes, however, is an invaluable tool to manage landscapes and maintain a significant portion of the avian insectivore guilds along with the ecological services they provide.

Banks-Leite, C., R. M. Ewers, and J. P. Metzger. 2010. Edge effects as the principal cause of area effects on birds in fragmented secondary forest. Oikos 119:918–926.


Croissant, Y. 2013. mlogit: multinomial logit model.


CHAPTER 3: MODELING THE SPATIAL BEHAVIOR OF UNDERSTORY MIXED-SPECIES FLOCKS

Spatial behavior is an important aspect when considering the impact of human activities on native fauna (Lima and Zollner 1996, Morales et al. 2010, Carter et al. 2012). Most resources needed by organisms are scattered in space, and the majority of ecological interactions include some form of movement (Ghazoul 2005, Buskirk and Millspaugh 2006, Holyoak et al. 2008, Smaldino and Schank 2012). Based on the nature and distribution of their resource, some animals occupy large areas moving at random through the landscape, while others occupy exclusive territories which are defended to varying degrees (Buskirk 1976, Kokko 2005, Giuggioli et al. 2011). Understanding the spatial requirements for tropical forest species is a current challenge for effective conservation and management (Schick et al. 2008); little or no information is available for a large proportion of taxa due to logistic limitations in acquiring this type of information. The recent development of new tracking technologies and analytical tools heralds a favorable period for unraveling the spatial ecology of a large suite of organisms (Cagnacci et al. 2010). Historically, much of the conclusions on space use and home range information have been done in post hoc fashion by interpretation of data after their acquisition (Horne et al. 2008), meaning that this area of knowledge has been plagued by lack of predictive power. Currently, detailed information on movement has been incorporated into models of space use (Moorcroft et al. 2006), but these new techniques can be analytically and computationally challenging (Patterson et al. 2008). It is against this background that this chapter unfolds.

Understory mixed-species flocks have been identified about 150 years ago (Bates 1864), and studied more formally in the last 30 years (Munn and Terborgh 1979). It has been understood that flock core species occupy large areas averaging about 8 ha (Jullien and Thiollay
and this requirement may be a strong factor in the local extinction when forest fragmentation takes place (Stouffer and Bierregaard 1995). Flocks also avoid crossing open areas, even narrow roads, thus offering another mechanism explaining their sensitivity to forest fragmentation (Develey and Stouffer 2001, Laurance et al. 2004). Little has been understood beyond that. For investigators studying flocks, their movements approached random walks (Powell 1985), and the only predictable aspect is that individuals gather in the same location at the center of their territory every day at dawn. One investigator was able to detect hints of a pattern in which flocks seemed to move along small valleys for short distances (English 1998). Such valleys are common features of the micro topography in parts of central Amazonian terra firme forests (Laurance et al. 2010). In nature, few animals actually present truly random movement patterns at all times (Bartumeus et al. 2008). Yet, it is still unknown what type of habitat within the understory these flocks might actually prefer, and what would in fact be a minimum requirement in terms of area for flocks.

Mixed-species flocks are important features of the avifauna which organize a large set of species at a specific place and time (Powell 1989). Flocks saturate the landscape with abutting territories next to each other covering extensive areas of the Amazonian understory (Johnson et al. 2011). This is an open system in which individuals are constantly entering and exiting the aggregations, but the nucleus of activity which includes a set of five to ten core species is constantly on the move (Munn 1985, Jullien and Thiollay 1998). The species are highly territorial, and intraspecific disputes arise when two flocks meet, usually when one flock is trespassing into a neighboring territory. During these territorial interactions there is a flurry of vocalizations and chases between conspecifics which can last up to 30 minutes, but little or no physical aggression seems to take place (Munn and Terborgh 1979). Eventually, the intruding
flock moves back towards the center of its territory. These disputes are relatively common and may take place from one to four interactions in a day’s course (K. Mokross pers. obs.). Flock territories are extremely stable, and their shapes may change very little over the course of decades (Martinez and Gomez 2013). Some hints are also given by flocking systems in Central America, where drastic alterations in habitat such as tree falls change territorial limits (Greenberg and Gradwohl 1986). Yet, little is known as to how these territories form: do these flocks territories actually form at random and are crystalized into the landscape? Another pertinent question given by the realities of deforestation is the amount of habitat structure that can actually be removed without flocks dissolution: What are the actual minimum habitat requirements for flocks? These questions boil down to the need of understanding how flocks actually use their habitat and raise the need to understand what is driving their movements. Ideally, it is necessary to reproduce their spatial behavior in order to forecast what may happen in different scenarios of habitat change.

As previously mentioned, the current modelling techniques needed to reproduce flock movement behavior are quite complex. The information spectrum to be covered in understanding their natural history, collecting, organizing data, and developing advanced models is a daunting task. In order to achieve this goal, I began a collaboration with one of the pioneers of mechanistic home range models for territorial animals (Moorcroft et al. 1999) Dr. Mark Lewis at the University of Alberta in Canada, and Jonathan Potts, a recent graduate from England who began his work as a post-doc in his lab. Together, we translated the natural history and empirical movement data of Amazonian mixed-species flocks into a step-selection function that reproduces flock movements in relation to vegetation height and topography. The initial model starts in relatively simple fashion with only two drivers of movement (topography and canopy height),
allowing the inclusion of other factors, thus helping tease apart their importance. The other factors such as memory, resource patch renewal and territorial interactions can be tested in this framework. The main benefit from this model is that it provides indirect information about how flocks relate to their main resource: small invertebrates scattered through the understory vegetation. Insectivory is a difficult trophic interaction to measure; the actual measuring of the amount of resources is an extremely labor-intensive task which may provide inconclusive results (Sekercioglu et al. 2002). Mapping invertebrate abundance is hardly feasible in the way it has done for frugivorous birds for example. Lastly, these models provide a useful tool that allows the forecast of flock occupation in heterogeneous landscapes, giving higher predictive power on the amount of habitat available.

Figure 1. Initial simulations of flock movements in yellow in a 10-ha fragment based on a step-selection function. Empirical flock movement data are shown in black. Vegetation map shows darker green shades as taller vegetation, while lighter shades indicate vegetation below 15 m of height. Here, we test variation in a dimensionless exploratory parameter $q$, in which individuals are more prone to visit areas not previously visited. A. Exploratory parameter high, causing them to cover wider areas within the landscape. B. Exploratory parameter is low, resulting in reduced home range.
Appendices A and B are the two articles which have been published by this collaboration with Dr. Potts as a lead author and I as second author. He graciously allowed the reproduction of the manuscripts in this dissertation in order to complement this body of work on mixed-species flocks. The first article (Appendix A) is entitled “Step selection techniques uncover the environmental predictors of space use patterns in flocks of Amazonian birds”, and provides an explanation of the step-selection function; while the second article (Appendix B), entitled “A unifying framework for quantifying the nature of animal interactions” incorporates territorial interactions in shaping flock home ranges from the original model. The main findings of the first publication are that despite flocks being constantly on the move and showing apparent random walk behavior, they prefer moving towards areas of taller vegetation and lower topography (Potts et al. 2014a). This might be explained by concentration of resources in these areas, which tend to be more buffered in terms of microclimatic conditions for invertebrates (Ewers and Banks-Leite 2013). Resource depletion did not seem to be an important factor; flocks are prone to move back to a previously foraged area at any point, hinting that once an aggregation passes through an area, insects could be adopting cryptic behavior, thus making their foraging less optimal. The second article shows that when incorporating territorial interactions with habitat selection, they are able to better predict flock utilization distributions than based solely on habitat preferences, and suggest that the constant territorial interactions may result in the spatially stable structure in the landscape over time (Potts et al. 2014b)

These models have enabled a quantification of underlying drivers of movement and brought a better understanding of how mixed-species flocks are behaving. Given a prior knowledge of how a landscape is structured we may be able to predict its conservation value for one of the most sensitive guilds in Neotropical avifauna.
LITERATURE CITED


46


CHAPTER 4: DECAY OF INTERSPECIFIC AVIAN FLOCK NETWORKS ALONG A DISTURBANCE GRADIENT IN AMAZONIA

INTRODUCTION

Biological systems are often organized as networks (Girvan and Newman 2002, Proulx 2005, May 2006) and while these networks are nearly ubiquitous, analytical approaches have only recently been applied to identify common properties and understand system-level dynamics. In its simplest form, a biological network can be represented as a graph comprised of nodes (individuals or species) and edges (biological interactions). Network theory offers the ideal conceptual framework to understand the structural complexity of biological systems because it provides metrics to quantify and interpret interactions at the level of individual or species, and document the properties of the system as a whole (Bascompte et al. 2006). Ultimately, these approaches have advanced our understanding of a variety of complex biological processes and types of interactions such as mutualisms (Guimarães et al. 2011), trophic interactions (Stouffer 2005), fitness consequences of social behavior (Ryder 2008, 2009), disease transmission (Naug 2008), and robustness of communities to extinction (Kaiser-Bunbury 2010).

Networks of trophic interactions have been the focus of a substantial body of research (Dunne 2002, Stouffer 2005). The structure of these networks can be influenced not only by

---

This chapter previously appeared as:


It is reprinted by permission of the Royal Society Publishing. Permission letter in appendix D.
intrinsic traits of participating organisms (e.g., phenotype) but also by extrinsic characteristics of the environment (e.g., habitat heterogeneity). For example, habitat modification affects trophic network structure via changes in species richness and frequency of interactions (Tylianakis 2007). Previous work has shown that most ecological networks are resilient to environmental change, but that the threshold at which a community collapses is dependent upon the degree to which species are ecologically redundant and the responses of keystone species to habitat loss (Kaiser-Bunbury 2010, Guimarães et al. 2011, Fortuna 2012). While trophic networks have been fairly well studied, non-trophic interactions, such as the social mutualisms observed in avian mixed-species flocks have received considerably less attention. Ultimately, by characterizing the roles that species play within ecological networks I can begin to understand the assembly of ecological communities (Wey et al. 2008), the predisposition for species to engage in non-random spatiotemporal aggregations (Farine 2012) and how those ecological and evolutionary processes are influenced by environmental change.

Mixed-species flocks are the most complex multi-species aggregations found in terrestrial vertebrates (Munn 1985). Some forms of mixed-species flocking occurs throughout the world, but few reach the temporal stability and interdependency of understory mixed-species flocks within Amazonian rainforests. These flocks of insectivores have a year-round “flock” territory and consist of a core eight to ten obligate species, each represented by a single pair. The territories of these core species overlap exactly in a flock territory of 8-10 ha (Munn 1985, Jullien and Thiollay 1998, Develey and Stouffer 2001). Cinereous Antshrikes (Thamnomanes caesius) play a nuclear role in these mixed species flocks throughout the Amazon basin (Powell 1985) by rallying individuals and giving alarm calls (Martinez and Zenil 2012). Up to 50 other species, mostly in breeding pairs, are known to join the core flock in varying periodicity (Jullien
and Thiollay 1998). Given that flocks are often species-rich and exhibit both spatial and temporal stability (Martinez and Gomez 2013), they represent an important component of the Amazonian understory avifauna (English 1998). Flocks provide direct fitness benefits to participating birds, including improved predator detection and increased foraging efficiency (Beauchamp 2004), yet these benefits may vary by species and be strongly dependent upon habitat context and flock organization.

The Amazon provides an important setting to examine changes in interspecific avian interactions because the area is subject to substantial forest clearing which produces heterogeneous landscapes of primary forest, secondary forest, forest fragments and interspersed roads (Butler and Laurance 2008). These newly fragmented and regenerating Amazonian forests influence the dynamics (Stouffer and Bierregaard 1995, Barlow 2007, Stouffer and Bierregaard 2007) and diversity of avian communities (Sodhi 2008). Given that mixed species flocks in the Amazon are largely forest dependent, they are highly susceptible to habitat disturbances. For example, most flock species avoiding open areas showing reluctance to cross narrow roads (Develey and Stouffer 2001) and often disappear in selectively logged forests (Thiollay 1997) and small fragments (Stouffer and Bierregaard 1995, Maldonado-Coelho and Marini 2000). Despite the detrimental effect of forest clearing on these species, depauperate flocks can still be detected in second-growth and small fragments (Stotz 1993). To date, research on how mixed-species flocks change along disturbance gradients have largely focused on species richness and encounter rates (e.g., number of detections per unit time(Lee 2005, Sridhar and Sankar 2008)), yet no studies have examined how habitat modifications influence interspecific interactions and the stability of flock structure.
Understanding how both species interactions and the subsequent structure of ecological networks change across landscape gradients are important because flocks can affect community dynamics and the fitness of participating species. To date, accurately characterizing interactions within mixed species flocks has remained challenging because flock attendance is dynamic (i.e., many individuals join and leave a flock in both time and space). The analytical framework of network theory can advance our understanding of flock dynamics by characterizing changes at multiple levels of organization (species, i.e., node level; and flock, i.e., network level). Here, I apply network theory to examine how individual species’ interactions and flock-level structure change across a heterogeneous landscape mosaic in the Amazon. First, at the species level, I compare how the number (degree) and frequency (weighted degree) of inter-specific interactions within mixed species flocks varied among primary forests (PF), 100-ha fragments (100-ha), 10-ha fragments (10-ha), a mix of primary and secondary forest (PSF) and secondary forests (SF). Second, I examine how environmentally induced changes at the species-level scale-up to influence flock-level network properties. In particular, I characterize how the connectedness (the distribution of interactions) and cohesion (clustering of species) of flocks change across a landscape gradient. Third, I examine the relationship between vegetation structure and network cohesion as one possible mechanism for changes in social structure among habitats. Given the changes in flock structure among habitats, I also compared species attendance among habitat types to determine if network differences were caused species disappearance or reduction in flock attendance. This work builds a framework for understanding how environmental heterogeneity affects the resilience of complex ecological interactions by examining the integrity and stability of flock networks across a habitat mosaic.
MATERIAL AND METHODS

Study site and data collection

The study was conducted at the Biological Dynamics of Forest Fragments Project (BDFFP) in central Amazonia, Brazil. Fragments were isolated between 1980 and 1990 during the settlement of cattle ranches (Lovejoy 1986, Bierregaard 2001). Following initial clear-cutting, some areas were burned to create pastures while others were abandoned. Most pastures were inactive by the 1990s, allowing the matrix around the fragments to regenerate. A buffer of 100 m was cleared around some fragments from the early 1990s to the early 2000s, but these areas have regenerated as well. The resulting landscape is structurally heterogeneous due to different management histories (Mesquita 2001), with fragments of different sizes surrounded by secondary forests varying in structure and age.

Understory mixed-species flocks gather in the same location every day at dawn and move through their territory until about 13 minutes prior to sunset (Jullien and Thiollay 1998). Flock activity is conspicuous, allowing birds to be followed on foot from a distance of 10-20m. Importantly, the mixed species flocks described here should be differentiated from those observed at army-ant swarms, which are comprised of solitary species that become spatially aggregated around a resource. I followed 21 flocks for at least 17 hours each (mean= 42.1, max=121.4, min.=17.1), totaling 693 hours between March and November of 2010 and 2011. I recorded flock composition in 30-minute time blocks, generating a total of 12,414 species entries. A species was noted as participating with a flock if it was seen within 15 meters of core flock activity for more than 30 minutes. To assess that I had adequately sampled flocks in each habitat, I constructed “sample-based” species accumulation rarefaction curves using program EstimateS (Colwell 2013)( Fig. S1). Based on flock sampling criteria, I also used program
Estimate $S$ and the frequencies of species in the original sampling data to generate a non-parametric estimator of species richness (Chao II) (Chao et al. 2006, Colwell 2013). Likewise, I estimated encounter rate as the number of times a species was detected corrected by total sampling time.

**Habitat and vegetation characterization**

To measure flock territories, flock positions were recorded at 30-second intervals with a Garmin eTrex Vista HCx unit (~10 meter resolution). A quadratic kernel was generated using Geospatial Modeling Environment software (Beyer 2012), a 99% isopleth was generated at 1m resolution, 275 bandwidth at default scaling factor. Vegetation was measured using a LIDAR (Light Detection and Ranging) canopy height models (CHM) provided by Scott Saleska (University of Arizona) and Michael Lefsky (Colorado State University). I generated the zonal statistics for the vegetation located inside the isopleth (Electronic supplementary materials, Table S1).

Flock territories were categorized in five habitat types: Primary forest (PF), if flocks used more than half of the territory in continuous primary forest; 100-ha fragments (100-ha), if flocks inhabited a 100-ha fragment; 10-ha fragments (10-ha), these fragments are only large enough for one flock territory; Primary-Secondary forest mix (PSF), if flocks used more than half of their territory in degraded secondary forest and patches of semi-isolated primary forest; and Secondary forests (SF), if a territory was exclusively in secondary forest. Mean vegetation height was used as an indication of habitat structure and quality.
Network and statistical analyses

I constructed networks for mixed species flocks in all five habitats types. Based on species co-occurrences in each time block, I used the cumulative frequency of associations to construct weighted networks for each flock. Specifically, network edges were defined using species co-occurrences within sampling time blocks. As such, any species associated with the flock aggregation is by default associating with all species present in that sampling time block. Hereafter, I use the term ‘interspecific association’ to describe these interactions. Using spatial proximity to define network associations in this manner is termed the “gambit of the group” such that all individuals within a spatial and temporal range will have reciprocal ties in the network (Croft 2008, Franks et al. 2010). Although, many species will appear accidentally in flocks, our threshold (30-minute sampling time blocks) removes accidental species, which do not accompany the flock for more than a few meters. I chose not to apply filtering techniques to remove low frequency co-occurrences because I used weighted network metrics and were interested in how common and rare species influenced network structure across habitat types. Moreover, data from replicates of independent flocks within habitat types should produce more precise measures of network co-occurrence.

At the species level, for each flock, I calculated unweighted and weighted degree metrics using UCINET (Borgatti 2002) and the R package tnet (Opsahl 2009). Degree is the number of edges (co-occurrences) one given species (network node) maintains with other species in a flock (i.e., species connectedness). Weighted degree is the sum of the frequency of interspecific associations for each node. Networks were visualized using R package ‘network’ (Butts 2012). I calculated the average network degree following (Albert and Barabasi 2002), the degree distribution skewness using R package moments (Halvorsen 2012) and global weighted
clustering coefficient following (Opsahl and Panzarasa 2009) using tnet package in R. To ensure that interspecific associations and network structure could be differentiated from random I used iterative permutation procedures (for methods and results see ESM). Network metrics for replicate flocks were grouped within habitat type for subsequent analyses (see above).

I used a suite of analyses to examine the effect of habitat on species and flock-level network properties. First, I looked at the response of species richness and encounter rate to habitat type. To examine how species richness (Chao II estimator) varied by habitat type I used a GLMM (Generalized Linear Mixed Models) with flock as a random effect and habitat as a fixed effect. All GLMM models used Poisson error distributions and log-link functions unless otherwise noted (see below). To examine how encounter rate (detections per unit time) varied by habitat I used a zero-inflated negative binomial mixed model to account for over-dispersion in the data. The residuals of all models were normally distributed and I compared the effects of habitat type with null models using Likelihood Ratio tests.

Second, I used GLMMs to examine how habitat influenced species-level network metrics (i.e., degree and weighted degree). Given that network data are not independent (Croft et al. 2011), I used flock replicates within each habitat type and included flock identity as a random effect and habitat as a fixed effect to explain variation in degree and weighted degree. To compare across networks in different habitats, I accounted for the number of possible species interactions and sampling time using a log (n or t) offset (McCullagh 1983) where n represented the number of possible interspecific associations within the network (n-1, number of nodes) and t represents sampling time. These corrections enabled us to compare networks with different number of species (nodes) and sampling effort, which is a common problem in network analyses (Croft 2008). Hereafter, I report the corrected values of normalized degree (Freeman 1979) and
weighted degree (i.e. frequency of associations are corrected by sampling time). Maximum likelihood estimates of \( \beta \) coefficients and P-values for fixed effects from each model are reported.

Third, to examine how flock-level network properties varied by habitat type I compare the distribution of species interspecific associations (degree distributions) using Wilcoxon sign-rank tests (see ESM). In addition, given that there was substantial variation in habitat structure among replicates (Table S1) I used least-squares regression to examine how a continuous measure of habitat, vegetation height, influenced flock attendance and cohesion as measured using a weighted clustering coefficient. I used the base package of program R for regressions and lme4 (Bates 2012) and glmmADMB (Skaug 2013) for GLMMs. Graphs were generated using ggplot2 (Wickham 2012).

Finally, I estimated changes in species participation by using detections within each flock. I used primary forest (PF) as the template for comparison under the assumption that it represents baseline flock species composition and attendance rates. To distinguish between species disappearance and decreased attendance, I compared presence/absence and attendance data from altered habitat (100-Ha, 10-Ha, PFS and SF) to those observed in PF. I used all possible pairwise combinations of primary forest flocks and those in other habitats. For example, the comparison of PF (9 flocks) and SF (3 flocks) would generate 27 values. I report percentages of species that disappeared (if only one species was present in a determined habitat type) or decreased attendance (if both species were present in both habitat types) relative to PF. Because some novel species appear in certain habitats and other increase in attendance, reported values do not sum to 100%. The magnitude of change in attendance is reported as averages across flocks sampled in each habitat type where negative values represent decreases.
RESULTS

I sampled 21 flocks across five habitat types with each habitat type receiving a minimum of 86 hrs (Table 1 and S1). Habitat was a good predictor of both species richness ($\chi^2 = 15.85$, $P = 0.003$) and encounter rate ($\chi^2 = 11.92$, $P = 0.017$) as evidenced by those models being better fit than null models (Table S2). Species richness was significantly higher in 10-ha fragments than in intact habitats ($\beta_{100\text{-ha}} = -0.56$, $P = 0.0003$; $\beta_{PF} = -0.48$, $P = 0.0001$; Fig. 1A). In contrast, more intact forest environments (e.g., primary forest and 100-ha) had higher encounter rates than 10-ha fragments and secondary forest after controlling for sampling effort ($\beta_{100\text{-ha}} = 0.88$, $P = 0.003$; $\beta_{PF} = 1.01$, $P = 0.0002$; Fig. 1B).

Figure 1. Mixed-species flocks showed substantial variation in both A) species richness and B) abundance across a habitat gradient in the Brazilian Amazon. Bars represent mean ± standard error.

Models constructed to test for the effect of habitat type on inter-specific associations showed that environmental heterogeneity influenced species level network metrics. Specifically, models that included habitat were a significantly better fit than null models for both normalized degree ($\chi^2 = 22.28$, $P = 0.0002$) and weighted degree ($\chi^2 = 15.69$, $P = 0.003$; see Table S3).
Flocking species in primary forest and 100-ha fragments associated with a proportionately greater number of other species (normalized degree) than did species in degraded habitats (Table 2; Fig. 2A). Likewise, flocks in less disturbed areas also had a higher frequency of interspecific associations (weighted degree) than did flocks in degraded habitats (Table 2; Fig. 2B). Flock identity (random effect) explained only a small portion of the variance in our models, suggesting that results are consistent across habitat type replicates.

Organization of flock networks differed across different habitat types (Fig. 3). A consistent group of core species, particularly *Thamnomanes caesius*, *Xiphorhynchus pardalotus* and *Myrmotherula axillaris*, was present across all habitat types, but the complexity of the network was far greater in primary forest than in 10-ha fragments and secondary forest. An examination of cumulative degree distributions revealed habitat-specific differences in global network structure (Fig. 4). Networks in degraded forest habitats were composed of many weakly associated birds (low degree) with few well-connected species (high degree), which resulted in flocks with low median normalized degree and degree distributions with strong positive skew (i.e., long-tail to the distribution; Fig. 4C-E; Table S4). In comparison, networks of primary forests and 100-ha fragments had a higher median normalized degree and weaker positive skew (Fig. 4A & B; Table S4). Differences in social structure at the level of the flock appear to be in part, driven by vegetation characteristics. Vegetation height was positively correlated with flock cohesion as measured by global clustering coefficients ($R^2 = 0.50$, $F_{1,19} = 19.07$, $P = 0.0003$; Fig. 5A) and species attendance patterns ($R^2 = 0.37$, $F_{1,19} = 11.08$, $P = 0.003$; Fig. 5B).

Comparing presence/absence data and species attendance between disturbed and primary forest habitats shows that decreased attendance rather than disappearance is the primary driver of differences in network structure (Table 3). Specifically, a relatively small proportion of species
were absent in disturbed habitats (2-12%), whereas nearly half of species detected in both habitats decreased attendance rates (47-56%) relative to primary forests. The magnitude of change was largest in the two most degraded habitat types (e.g., 10-Ha and SG; Table 3).

**DISCUSSION**

Interspecific interactions in communities are an essential component of ecosystem function and have important implications for the ecological and evolutionary dynamics of species (Vazquez et al. 2007). To date, research on trophic interaction networks has shown that habitat changes can affect interspecific networks, yet their structure is often resilient to habitat change because species are ecologically redundant (Kaiser-Bunbury 2010, de Visser et al. 2011, O’Gorman et al. 2012).

![Figure 2](image.png)

Figure 2. Box plots show that species in mixed-species flocks in primary forest habitats (PF and 100-Ha) A) had a greater number of interspecific interactions (normalized degree) and B) a higher frequency of interactions (weighted degree) than in degraded forest habitats (SF, PSF and 10-Ha).

Our results corroborate that the interspecific associations that comprise flock networks are also affected by habitat degradation. While there is little known about the functional roles of species within flocks, our results suggest that flock social structure may be comparatively more sensitive than other ecological networks studied to date. Here, I document changes in species
richness, encounter rates, species connectedness and the frequency of interspecific associations within mixed species flocks along a habitat mosaic. Our results suggest that habitat modification and changes in vegetation structure alter flock attendance and subsequent inter-specific associations resulting in reduced flock cohesion and stability. Given that mixed-species flocks host a diversity of understory insectivorous birds these results highlight the potential negative impact of habitat alteration on the dynamics of species interactions.

Figure 3. Example of networks and habitat configurations for three flocks found in Primary forest, 10-ha fragment and Secondary forest habitat types in the Brazilian Amazon. Differences in network structure reflect the decay of interspecific interactions in mixed-species flocks across a disturbance gradient. Edge thickness and transparency in each network are proportional to numbers of interactions. Interaction values at the lowest 10% are set to transparent. Nodes sizes are proportional to flock attendance. Legends for species with participation above 6% are given.

Figure 4. Cumulative degree distributions highlight differences in mixed-species flock network structure across habitat gradient in the Brazilian Amazon. Pairs of histograms for observed and randomized networks show that degraded habitats are characterized by low medians and high skew while more intact habitats have higher medians and lower skew. Dashed lines represent the median normalized degree.
Figure 5. Mixed-species flock cohesiveness (clustering) and attendance show strong positive correlation with vegetation structure. A) Relationship between flock global clustering coefficient and mean vegetation height. B) Relationship between species attendance and mean vegetation height. Y-axis represents the cumulative amount of time all species in a given habitat participated in a flock. Flocks (points in the graph) are color-coded for habitat type (Blue=Primary forest, Red= 100-ha fragment, Yellow=10-ha fragment, Green=Primary-Secondary forest mix, Orange- Secondary forest). Dashed line represents the β coefficient from the model and the gray shaded area is the 95% confidence interval.
Table 1. Summary of sampling for mixed species flocks in five habitat types at the Biological Dynamics of Forest Fragments Project in central Amazonian Brazil.

<table>
<thead>
<tr>
<th>Habitat Type</th>
<th>Replicates (n)</th>
<th>Max Ht. (m) ±</th>
<th>Avg. Ht. (m) ±</th>
<th>Sampling Time (hrs)a</th>
<th>Attendance ±</th>
<th># Spp.b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Forest</td>
<td>9</td>
<td>45.7 ± 1.5</td>
<td>23.8 ± 0.35</td>
<td>304.0</td>
<td>851.9 ± 51.1</td>
<td>109</td>
</tr>
<tr>
<td>100-Ha Fragment</td>
<td>5</td>
<td>46.3 ± 1.3</td>
<td>21.4 ± 1.2</td>
<td>151.5</td>
<td>918.0 ± 60.5</td>
<td>88</td>
</tr>
<tr>
<td>10-Ha Fragment</td>
<td>2</td>
<td>45.0 ± 5.6</td>
<td>16.4 ± 1.8</td>
<td>134.5</td>
<td>717.3 ± 22.2</td>
<td>103</td>
</tr>
<tr>
<td>Primary-Secondary</td>
<td>2</td>
<td>40.4 ± 3.7</td>
<td>16.8 ± 0.90</td>
<td>86.0</td>
<td>797.0 ± 19.9</td>
<td>79</td>
</tr>
<tr>
<td>Second-Growth</td>
<td>3</td>
<td>39.1 ± 5.5</td>
<td>14.4 ± 1.97</td>
<td>88.0</td>
<td>559.5 ± 121.4</td>
<td>82</td>
</tr>
</tbody>
</table>

a) Total sampling time
b) Cumulative number of species observed during sampling

Table 2. The relative changes in the presence/absence and attendance from primary forest to degraded forest for mixed-species flocks. Species were recorded as absent when detected in primary forest but not degraded habitats and decreasing attendance when detected in both habitats but lower relative proportion. Magnitude of change describes the average differences in encounter rate between primary and degraded habitats.

<table>
<thead>
<tr>
<th>Habitat Type</th>
<th>% Spp. Absent</th>
<th>% Spp. Decreasing Attendance</th>
<th>Mean Magnitude of Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-Ha</td>
<td>3.0</td>
<td>47.7</td>
<td>-3.79 ± 2.50</td>
</tr>
<tr>
<td>Second Growth</td>
<td>12.3</td>
<td>47.7</td>
<td>-5.75 ± 3.44</td>
</tr>
<tr>
<td>Primary-Secondary</td>
<td>3.3</td>
<td>56.7</td>
<td>-2.62 ± 4.37</td>
</tr>
<tr>
<td>100-Ha Fragment</td>
<td>2.4</td>
<td>55.1</td>
<td>-1.94 ± 2.50</td>
</tr>
</tbody>
</table>

Species-level and flock-level changes in network structure

Multispecies interactions form the basis of ecological networks, and changes in species presence or behavior can have profound impacts on network structure and ecosystem function (Beyer et al. 2010). Research on food webs suggests that habitat degradation tends to promote homogenization (loss of species) resulting in reduced network complexity and stability (Albrecht et al. 2007, Tylianakis 2007, Laliberte and Tylianakis 2010). Our results, from a non-trophic flock network, corroborate the idea that extrinsic environmental features, like habitat, can affect interspecific social structure. At the species level, our results indicate that flocking birds in small
fragments and degraded secondary forests associated with fewer species (degree) and did so less frequently (weighted degree) than individuals in intact habitats. Structural differences among networks at the level of the flock are best illustrated by habitat specific degree distributions, which show the cumulative effect of changes in species associations across habitat types. In particular, the majority of associations in 10-ha fragments and secondary forests were driven by a few remaining core species, while associations were more evenly distributed across species in intact forest environments.

Our results also highlight that habitat configuration can influence patterns of species richness. Specifically, species richness was higher in some of the more degraded habitats than in large intact forests tracts. These results differ from previous studies in the old world tropics, which found decreases (Lee 2005) or no changes in species richness (Sridhar and Sankar 2008) in mixed species flocks along disturbance gradients. Increases in species richness in 10-ha fragments were not altogether surprising given the available habitat matrix and community composition. In particular, the proximity of fragment borders and secondary forest likely enabled canopy and edge specialists to interact with understory flocks, thereby increasing diversity. Despite high richness and the addition of novel species in degraded habitats, species were encountered less frequently and had reduced flock attendance in 10-ha fragments and secondary forests. These results are consistent with the idea that flocking species are disproportionately affected by habitat disturbance when compared to other guilds (Thiollay 1997, Sridhar and Sankar 2008). Ultimately, increases in species diversity do not appear to have meaningfully modified flock social dynamics.

Our results suggest that changes in flock attendance rather than species loss across the habitat gradient is likely driving changes in flock social structure. Differences in attendance may
result from either reduced bird density in fragmented and sub-optimal habitats (Thiollay 1997, 1999) and/or changes in propensity to join flocks (Dolby and Grubb 1999, Thiollay 1999, Knowlton and Graham 2011, Martinez and Zenil 2012). Previous work in this system and our data suggest that these flocking species may have lower densities but are not completely absent in degraded habitats (Stouffer and Bierregaard 1995, Stouffer 2006). Comparisons of flock participation show that a large proportion of species appear to alter their behavior by reducing flock attendance. Moreover, such changes in behavior also suggest that the costs and benefits of flocking behavior may vary with environmental context (Lee 2005). While I have no data on the mechanisms for decreases in flock attendance, I believe that arthropod prey abundance declines in more degraded environments (Stouffer and Bierregaard 1995, Şekercioğlu 2002, Stouffer 2007). Many permanent flock species are known to have specialized foraging niches, and a reduction in their food resource could increase space use, which might decrease flock attendance. Regardless of the mechanism, reduced attendance and subsequent changes in flock composition will influence the stability, cohesion and integrity of these complex multi-species interactions (Maldonado-Coelho and Marini 2004).

**Habitat and the consequences of changes in social structure**

Flock cohesiveness within the network, as measured by weighted clustering coefficients, was positively correlated with mean vegetation height (Fig. 5A). Importantly, vegetation height is a good proxy for structural habitat complexity (Goetz et al. 2010, Whitfeld et al. 2012) as primary forests and 100-ha fragments also tended to have greater vertical forest stratification than small fragments and secondary forest. The positive relationship between vegetation height and clustering coefficients are best explained by two possible mechanisms. First, vegetation height could have directly influenced network structure if the number of interspecific
associations within flocks is driven, in part, by structural components of the habitat. The documentation of strong vertical stratification within Amazonian bird communities partially support this idea (Walther 2002, Comín 2010). Second, vegetation height could indirectly affect network structure if predation pressure covaries with habitat type. For example, predator communities have been shown to drive flocking propensity and may vary across degraded tropical forests (Thyollay 1985, Chazdon et al. 2009).

Changes in network properties may influence individual performance of species whose natural history revolves around joining flocks. For example, reduced flock attendance may reduce predation avoidance and foraging optimality (Thiollay 1999, Couzin et al. 2002). Flocks in highly disturbed areas were unstable, not lasting more than a few weeks, where pairs of Thamnomanes caesius were inconsistently present. In contrast, flocks in continuous forest are known for their long-term stability, even as individual participants disappear and are replaced (Jullien and Thiollay 1998, Jullien and Clobert 2000). Assuming that changes in flock network structure influence individual fitness, future work should focus on measures of fitness by gathering species level data on foraging efficiency and space use, as well as community level data on predator communities and resource distribution.

**Conclusions**

Forest clearing is one of the largest threats to biodiversity today (Pimm and Raven 2000). In the Amazon, the impact of forest fragmentation on avian species loss is well documented (Stouffer and Bierregaard 1995, Stouffer 2006, 2009), yet changes in behavior, interspecific interactions and community dynamics are less well known. Identifying mechanisms that disrupt ecological processes in human modified habitats is an essential step in mitigating and conserving diverse tropical communities. Network analyses are a powerful tool for quantifying how trophic
and non-trophic interactions and subsequent ecological networks change across landscape gradients because they enable us to quantify the role that species play in community structure and function. Moreover, this approach is likely to be especially useful in the tropics because of the high diversity and subsequent complexity of interspecific interactions. Future research must move beyond simply tallying species lists and towards identifying mechanisms that alter species interactions and community function (Tylianakis 2007). The results presented here advanced our understanding of how non-trophic interspecific interactions and subsequent community structure change along a disturbance gradient. Ultimately, if behavioral interactions and the structure of non-trophic networks tend to be highly sensitive to environmental change, as shown here, a more nuanced approach may be needed when thinking about the resiliency of ecology networks.

LITERATURE CITED


Bates, D., Maechler, M. and Bolker, B. 2012. lme4: Linear mixed-effects models using S4 classes.


CHAPTER 5: HOW AMAZONIAN MIXED-SPECIES FLOCKS BREAK UP AND RECOVER IN FOREST FRAGMENTS: A 30 YEAR ANALYSIS

INTRODUCTION

Forest fragmentation is a pervasive process presently occurring in large tracts of the Amazon basin (Souza et al. 2013); and whereas local extinction or changes in abundance have been the focus of many studies (Robinson 1999, Sigel et al. 2006, Banks-Leite et al. 2010), the effects of fragmentation on interspecific interactions have recently gained more focus as representative of more complex biological processes (Tylianakis et al. 2010, Magrach et al. 2014, Mokross et al. 2014). Mutualistic interspecific interactions are especially important in the tropics and play important role in the maintenance of a large set of biological systems. Gene flow through seed dispersal and pollination (Bawa 1990, Ghazoul 2005), defense from herbivores through an-plant associations (Bronstein 1998), and protection from predators and increased foraging efficiency by interspecific flocking (Powell 1985, Martinez and Zenil 2012) are examples of such processes dependent on species interactions. The negative effects caused by their interruption are difficult to quantify due to logistic constraints in their observation, or the time scale in which effects may manifest (Galetti et al. 2013). Large-scale forest fragmentation studies, for example, require notable logistic efforts which can hardly be repeated when new questions arise, and may take years in providing adequate answers (Didham et al. 2012). Short-term assessments can only portray the last frames of temporal dynamic processes, and may fail to provide relevant information on how such interactions change over time (Pinter-Wollman et al. 2014). One alternative is retrieving information in long-term databases from the perspective of species associations. Handling of large quantities of interrelational data has become feasible due
to increased access to analytical software and techniques such as network analyses, and may hold an opportunity to peer at relevant processes across longer time scales.

The effects of forest fragmentation have been well described in tropical avian communities, especially in central Amazon, where the Biological Dynamics of Forest Fragments Project has operated since the late 70’s (Laurance et al. 2011). Among the most vulnerable guilds in the understory, gregarious insectivorous species are particularly prone to local extinction (Sigel et al. 2006, Sodhi et al. 2011). Some flock species are detected in the regenerating matrix at about a decade or less (Antongiovanni and Metzger 2005) and are potentially able to colonize fragments relatively early (Stouffer and Bierregaard 1995). Yet, despite a subset of flocking species detected in fragments of 1 and 10 ha during this period, there is no evidence of cohesive flocks.

Insectivorous mixed-species flocks are a common feature in the understory occurring throughout terra firme forests in the Amazon basin. These mutualistic aggregations are led by a nuclear species, the Cinereous antshrike Thamnomanes caesius (or the allospecies in Bolivia and Peru, the Bluish-slate antshrike T. schistogynus), which is largely responsible for flock cohesion and movement patterns (Munn 1986). Attending species are almost entirely represented by reproductive pairs that defend their territory from conspecifics while actively foraging through vegetation in the understory (Jullien and Thiollay 1998). A set of core species, which varies from about seven to ten, have completely overlapping territories (Munn and Terborgh 1979). These territories tend to be highly stable, suggesting flocks are an important component in the structure of the avian community in both space and time (Martinez and Gomez 2013). Other species, which can be well over 50, show varying degrees of attendance in these aggregations.
As shown in previous work at the BDFFP, obligate flock followers are known to disappear in smaller fragments (1 and 10 ha) or drastically reduce in abundance. Flock dropouts, or non-obligate flocking species that are frequently seen attending flocks, show mixed responses, but an overall decrease, albeit less dramatic than obligate flocking species. There is evidence pointing to a crowding effect followed by flock dissociation after fragment isolation (Stouffer and Bierregaard 1995). Subsequently, some of the key flocking species have been detected in secondary forest after a decade and are able to recolonize fragments after a few years. Whereas periodic mist net sampling in these fragments has allowed a detection of many flocking species during the regeneration of the surrounding matrix, little is known about the social structure and process in which these flocks reorganize. Presently, flocks have been shown to change in social structure according to varying degrees of habitat disturbance and those drivers are largely due to changes in attendance behavior for many species (Mokross et al. 2014). However, it remains to be known how the process of flock dissolution and its subsequent regeneration occur throughout the isolation history of fragments of different size.

The effects of forest fragmentation on avian wildlife begin by a massive displacement of individuals seeking refuge from recently cleared areas into the forest remnants left standing. During this period, many territorial boundaries change or dissolve entirely, and a crowding effect takes place. Subsequently, there is a great loss of individuals, with increased severity in smaller fragments (Stouffer and Bierregaard 1995). With regeneration of the matrix, species are able to recolonize fragments but rates of local extinction and colonization are still dependent on fragment area (Ferraz et al. 2003, Stouffer et al. 2011).
I wish to test if mist net captures can provide indirect evidence of flock dissolution and recovery. Specific questions are if there are significant differences in species associations or if this is a direct effect of decreases in abundances of flocking species. I also inquire which species stop associating with flocks after fragment isolation and which species remain.

MATERIALS AND METHOD

Study area

The Biological Dynamics of Forest Fragments Project (BDFFP) is about 80 km north of the city of Manaus, in the Brazilian state of Amazonas (2°30’ S, 60° W). Fragments were isolated between 1980 and 1990 during the settlement of cattle pastures (Lovejoy 1986, Bierregaard 2001). However, due to the cut of government subsidies, most pastures were inactive by the 1990s and the matrix around the fragments regenerated. Periodically, a buffer of 100 m was cleared around some fragments from the late 1990s to the early 2000s, but these areas have regenerated as well leaving an area of secondary forest around most fragments that is younger than the adjacent secondary forest. The vegetation is composed of terra firme forest that reaches 30-37 m with occasional emergent that reach up to 55 m (Lovejoy and Bierregaard 1990). Fragments are located in three areas and were isolated between 1979 and 1992 totaling seven fragments of 1 ha, four fragments of 10 ha, and two fragments of 100 ha.

Data collection

A standardized mist net sampling protocol has been conducted since 1979, before the isolation of fragments. Mist nets (NEBBA type ATX, 36 mm mesh, 12 x 2 m) were opened along trails in the interior of fragments from 0600 to 1400. I used eight nets in 1 ha-fragments, 16 nets in 10 ha fragments, and 32 nets in 100 ha fragments. Sampling effort has not been equal
throughout the fragment’s isolation history. The late 90’s and early 2000’s were less intensively sampled compared to the initial years after isolation.

In order to reconstruct flock associations throughout fragment isolation history from mist net data, I consider only the captures of the nuclear species *Thamnomanes caesius* with any other flocking species within 30 minutes of its capture in the same net and one adjacent net. In case an individual *T. caesius* fell at the end of a net lane, I consider the two adjacent nets to the opposite side in order to keep a constant sampling effort. I only considered 20 species that typically flock with *T. caesius*, of which flocking species that forage in the lowest stratum are best represented.

**Data structure**

The minimal sampling unit is a co-capture with *Thamnomanes caesius*, which is a relatively common capture in primary forests at the BDFFP (Stouffer and Bierregaard 1995). This key component of mixed-species flocks is frequently captured with other species within short distances in a net lane. The total number of these co-captures within a considerable sampling interval, however, is still relatively scarce if one wishes to detect trends in species associations. Because of this limitation, I pooled captures within all fragments of the same size within an isolation interval, therefore, data are divided into three fragment size categories: 1, 10 and 100-ha fragments. Since our main question relate to how flocks are reacting in different isolation periods according to fragment size, I believe that this framework provides an outline of changes in interspecific associations with sufficient consistence.

For each fragment type, I pooled captures into 6 time periods with equivalent sampling effort: pre-isolation (pre i); a concentrated sampling period in the year following isolation (i); a period of two years starting after the first year going to the third year after isolation (i2); a period of five years starting after the third year going to the eighth year after isolation (i3), another
interval of five years going from the eight to thirteen years after isolation (i4); and an interval lasting twelve years starting at the thirteenth year after isolation to about 25 years after isolation (present). The duration of sampling intervals has been adjusted using the pooled pre-isolation sampling effort by fragment size as a reference, 8000 net/hours for 1 ha-fragments, 7000 net/hours for 10 ha fragments, and 4000 hours for 100-ha fragments. The last isolation period is relatively long, encompassing 12 years, as the sampling that amounts the same number of net/hours is more scattered during the late 90’s and early 2000’s. Individuals associated with T. caesius are henceforward also referred to as associated to flocks.

Analyses

First, I assessed abundance and species richness in each isolation period for each fragment size. Second, I compared the overall richness to the richness of species captured with T. caesius within the same sampling period.

I obtained the ratio of individuals caught with T. caesius as well as a ratio of total number of individuals captured in fragments (i.e., individuals that are associated to flocks as well as solitary) with T. caesius in each isolation interval.

There are two possible responses from individuals not associated to flocks, they might be locally extinct or still be present in a fragment, but not associated to T. caesius. I determined the number of species in each fragment size at a given isolation period that has either disappeared (was not captured in a given fragment size) or stopped its association to flocks (was captured in a fragment but was not co-captured with T. caesius).
All network measures were corrected for species abundance. This ensured that interactions between two abundant species would not be overrepresented, given their chance of being captured together would be higher. This corrected value is obtained by:

\[ I = \frac{c}{a} \cdot c \]

Where \( I \) is the interaction value after being corrected by species abundance. The value \( c \) is the number of co-occurrences that were observed and \( a \) is the total possible number of co-occurrences which is obtained by multiplying the total overall abundances (i.e. captured individuals) of each species.

Co-captures were compiled in isolation periods and used to build co-occurrence networks to describe flock structure in each fragment type. Three network level measures of flock connectivity are used: mean weighted degree, weighted graph density and marginal species binary graph density excluding \( T. caesius \). Each metric is explained below.

Mean weighted degree takes into consideration the frequency of links between species (i.e. how frequently two species are caught together) and serves as an indicator of how strongly species are interacting in a network (Opsahl et al. 2010). From a data perspective, it means the amount of times species are captured in the same net set.

Weighted graph density is a measure of how saturated a network is. This is based on the total amount of interactions that are possible given the number of nodes and is expressed as a percentage of actual links compared to the total number of possible links (Scott 1988). I consider the total possible number of links based on individual abundance of each species caught in a fragment in each isolation period, given by multiplication of each species’ abundance in a dyadic association.
Marginal species binary graph density is based on the previous measure, but instead of taking frequencies into account, only considers the presence or absence of interactions between flocking species other than T. caesius. This provides a measure of how cohesive flocks are. Network saturation is determined by the presence or absence (binary) linkages among typical flocking species. T. caesius and its edges (i.e., links between two species) are excluded in this analysis since the acquisition of this data is structured around this species.

RESULTS

Our database parsed from a total of 23202 captures, of these, there were 6191 captures of individuals associated to flocks (1443 in 1ha-fragments, 2479 in 10-ha fragments, and 1929 in 100 ha-fragments), totaling 410 T. caesius and 5781 co-captures from re-isolation to 2009. Changes in species richness and overall abundance

Species abundance in 1- and 10-ha fragments increased significantly during the first year after isolation, presenting a larger increment in 1 ha-fragments, the two subsequent periods in 1-ha fragments showed a drastic reduction in overall abundance and a slight recovery from 12 years post-isolation onwards. Species abundance in 10-ha fragments showed a similar trend, with a stronger upswing in abundance in the last two post-isolation periods. Both 1 and 10 ha-fragments still show lower abundances compared to the pre-isolation period (Fig.1A). Abundance in 100 ha-fragments showed a slight drop after isolation with an increase in abundance compared to pre-isolation abundances 12 years after isolation (Fig. 1A).
Figure 1. A. Number of captures. B. Ratio of species caught with the flock nuclear species *Thamnomanes caesius*. C. Ratio of overall captures with *T. caesius*. Interval periods are: *pre.i*: pre-isolation, *i*: one year after isolation, *i2*: second year after isolation, *i3*: three to eight years after isolation, *i4*: eight to 13 years after isolation, *present*: 13 to 25 years after isolation.

Table 2. Number of species captured in fragments but not associated with *T. caesius* (S) and number of species not captured in fragment at each isolation periods at 1, 10 and 100 ha-fragments. Rows represent isolation periods. *Pre i*: Pre-isolation; *i*: first year after isolation; *i2*: second to third year after isolation; *i3*: third to eighth year after isolation; *i4*: ninth to thirteenth year after isolation; *present*: fourteenth to twenty-fifth year after isolation.

<table>
<thead>
<tr>
<th></th>
<th>1 ha</th>
<th></th>
<th>10 ha</th>
<th></th>
<th>100 ha</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S</td>
<td>D</td>
<td>S</td>
<td>D</td>
<td>S</td>
</tr>
<tr>
<td><em>Pre i</em></td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td><em>i</em></td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td><em>i2</em></td>
<td>14</td>
<td>6</td>
<td>9</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td><em>i3</em></td>
<td>10</td>
<td>9</td>
<td>12</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td><em>i4</em></td>
<td>7</td>
<td>12</td>
<td>8</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td><em>present</em></td>
<td>10</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 2. Richness of flocking species caught in A. 1 ha-fragments, B. 10 ha-fragments, and C. 100 ha-fragments. Closed symbols represent richness of all individuals caught in fragments, open symbols represent species richness of individuals caught with *T. caesius*. Interval periods are: *pre.i*: pre-isolation, *i*: one year after isolation, *i2*: second year after isolation, *i3*: three to eight years after isolation, *i4*: eight to 13 years after isolation, *present*: 13 to 25 years after isolation.
The ratio of *T. caesius* to the pooled number of co-captures followed a similar trend to overall abundances, except that the number of co-captures for each *T. caesius* dropped right after isolation in 10-ha fragments (Fig. 1B). There were no co-captures in 1 ha-fragments from 2 years after isolation (i2) onwards. No *T. caesius* was captured two years after isolation in any 1 ha-fragments; in the subsequent isolation periods *T. caesius* were captured, but without co-captures. The ratio of overall species abundances captured in fragments to the abundance of *T. caesius* did not follow the same trends as the ratio of *T. caesius* and pooled co-captures (Fig 1C). Instead, in 1 ha-fragments the number of potentially flocking species per *T. caesius* individual were stable in the first year after isolation, since there were no *T. caesius* captures in the third to eighth year after isolation, the ratio for 1 ha fragments at this time period is essentially null. Eight years after isolation onwards, the number of captures per individual *T. caesius* rises substantially (Fig. 1C). This ratio increased in 10 ha-fragments in the third to eighth (i2) year after isolation, reaching similar estimates to pre-isolation after the thirteenth year after isolation (Fig. 1C). The ratio between the overall abundance of flocking species and *T. caesius* remained essentially stable in 100 ha-fragments (Fig. 1C).

**Species richness and composition**

There was an overall reduction in species present in fragments after isolation followed by a slight recovery. The most dramatic trend was seen in 1 ha-fragments, where there was a continuous drop in richness from 19 to in the first to third year after isolation (i2) to eight species captured in the thirteenth year after isolation to present (i4 and present). At these same intervals, associated species to *T. caesius* went from 16 to none (Fig. 2A). The species that were more strongly associated to *T. caesius* in relation to their abundances in 1 ha-fragments before isolation were *Automolus infuscatus, Myrmotherula longipennis, Xenops minutus, Philydor*
erythrocercus and Myrmotherula axillaris (Table 1). After isolation there was an increase for association values for all species, with the strongest positive change in Thamnomanes ardesiacus, Epinecrophylla gutturalis, Hylophilus ochraceiceps and Myrmotherula axillaris (Table 1).

After the first year since isolation, most species were captured in 1-ha fragments at some point but were not associated to T. caesius, this was consistent for Glyphorynchus spirurus, Mionectes macconnelli, M. axillaris, E. gutturalis, Rhynchoclyclus olivaceus and Xiphorhynchus pardalotus (Table 2).

Overall species richness in 10 ha-fragments decreased from 19 to 15 species between the eighth to thirteenth year after isolation (i3), with an upswing to 17 the penultimate isolation period (i4). The number of species co-occurring with T. caesius, however, reduced drastically at that same period (i3) going from 16 to three species (Fig. 2B). The species most strongly associated to T. caesius corrected for abundance before isolation were Myrmotherula menetriesii, X. pardalotus, Deconychura longicauda, H. ochraceiceps and E. gutturalis respectively (Table 1). After isolation, there was an increase in association with M. axillaris, T. ardesiacus and M. longipennis. In the following isolation period from one to three years after isolation (i2), there is a decrease in association to T. caesius, yet, all species were present in 10 ha-fragments, and the foliage-gleaner A. infuscatus, and the antwren M. longipennis are still consistently associated to T. caesius. In the two subsequent isolation periods spanning from three to thirteen years after isolation (i3 and i4), flock participation is reduced to less than half of species, reaching its lowest in i3 where associations are reduced to only T. ardesiacus and A. infuscatus. Throughout the isolation history of 10 ha fragments, no species were consistently associated with T. caesius. The species that were mostly present were A. infuscatus, which was present from 8 to 13 years form
isolation but was not captured with *T. caesius* at i4 and *X. pardalotus*, which was absent at i3. Again, from the third year since isolation to present, most species were absent from co-captures with *T. caesius* despite being present in fragments. The species that consistently stopped being co-captured from i2 onwards were *Certhiasomus stictolaemus* and *Myiobius barbatus*. *Hylophilus ochraceiceps* showed the same response from the eighth year (i4) onwards. The species at the last isolation period (present) show decreased association values compared to pre-isolation periods, the most relevant ones being *X. minutus*, *M. axillaris* and *E. gutturalis*.

Overall richness in 100-ha fragments did not change substantially, oscillating between 17 and 20 species, yet, the number of species co-captured with *T. caesius* decreased from 14 to 9 right after isolation (i) and recovered at the eighth year after isolation (i3) (Fig. 2C). The species more strongly associated to *T. caesius* in the pre-isolation period are *T. ardesiacus*, *X. pardalotus*, *E. gutturalis*, *M. longipennis* and *M. menetriesii* respectively. During the first year after isolation, 10 species were present but were not associated to *T. caesius* (Table X), the number steadily increases in subsequent isolation periods. *M. barbatus* and *Campylorhamphus procurvoides* were not captured with *T. caesius* at any isolation interval except for the last one (present) for the former and the first year after isolation for the latter. Overall, the species most prone to remain in fragments but not be captured with *T. caesius* were *R. olivaceus*, *M. barbatus*, *M. menetriesii*, *M. macconnelli*, *C. stictolaemus* and *D. longicauda*. The species that remained present in flocks with *T. caesius* were *T. ardesiacus*, *X. pardalotus*, *A. infuscatus*, *M. longipennis* and *G. spirurus*. *Glyphorynchus spirurus*, however, is very abundant in smaller fragments after isolation, with low participation numbers.
Network structure

Mean weighted degree in 1 ha-fragments had a sharp increase during the first period after isolation, afterwards, dropping to zero through all post-isolation periods. This value drops consistently in 10ha-fragments after isolation until eight years after isolation (i3), where then values increase (Fig. 4A). The recovery is at a slower pace than the loss of species, and at the last isolation period (present), values are at half the pre-capture values. The amount and strength of connections in 100 ha-fragments dropped during the first two isolation periods (i and i2), from which there was a spike surpassing pre-isolation levels after 8 to 13 years after isolation (i3). There is a subsequent drop from 13 years after isolation to present in relation to pre-isolation periods.

Weighted graph density, a measure of overall density and strength of links between species that were present in the fragments decreased steadily from isolation onwards for all fragment sizes. There was no crowding effect for this measure in any of the fragment sizes. Fragments of 10 ha showed a constant decrease until thirteen years from isolation (i3), from there showing a slight rise to values similar to ones detected right after isolation (i). 100-ha fragments showed an overall drop in weighted graph density from pre-isolation (pre-i) to the last isolation interval (present) (Fig. 4B). There is a sharp drop from 29% right after isolation (i) to 10% at the third year since isolation (i2) followed by a strong upswing to 22% at the following interval (i3), eventually decreasing to 10% again at the last interval (present).
Table 1. Capture abundance (white bars) and flock participation (inset in black) by the 20 main species in central Amazonian mixed-species flocks. Bars in red represent species captured in fragments, but not associated to *T. caesius*; red X represent species not captured and possibly absent. Species are divided by foraging strategies, ** are species predominantly foliage gleaners; + are aerial foragers, ++ woodcreepers. Maximum value is set to 50, captures extrapolating the maximum are indicated with numbers inset on right side.

<table>
<thead>
<tr>
<th>Species</th>
<th>1 ha-fragment</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>10 ha-fragment</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>100 ha-fragment</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 ha-fragment</td>
<td>10 ha-fragment</td>
<td>100 ha-fragment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pre.i</td>
<td>i</td>
<td>i2</td>
<td>i3</td>
<td>i4</td>
<td>present</td>
<td>pre.i</td>
<td>i</td>
<td>i2</td>
<td>i3</td>
<td>i4</td>
<td>present</td>
<td>pre.i</td>
<td>i</td>
<td>i2</td>
</tr>
<tr>
<td>Thamnomanes caesius</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Thamnomanes ardesius</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Aournemouth infuscatus</strong></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Philydor erythrocerus</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Myrmotherula longipennis</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Epinecrophylla gutturalis</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Myrmotherula menetriessi</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Myrmotherula axilans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hylophlus ochraceiceps</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>+ Mionectes macconnelli</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Myiobius barbatus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hylophlus muscioperus</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Rynchoecycus olivaceus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Xiphophorus pardalotus</strong></td>
<td>42</td>
<td>43</td>
<td>45</td>
<td>41</td>
<td>40</td>
<td></td>
<td>42</td>
<td>43</td>
<td>45</td>
<td>41</td>
<td>40</td>
<td></td>
<td>42</td>
<td>43</td>
<td>45</td>
</tr>
<tr>
<td>Certianomus stictaera</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Campylohamus procuroideos</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Sittasomus griseicapillus</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Dacrydium megacephalum</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Xenops minutus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2. Richness of flocking species caught in A. 1 ha-fragments, B. 10 ha-fragments, and C. 100 ha-fragments. Closed symbols represent richness of all individuals caught in fragments, open symbols represent species richness of individuals caught with T. caesius. Interval periods are: pre.i: pre-isolation, i: one year after isolation, i2: second year after isolation, i3: three to eight years after isolation, i4: eight to 13 years after isolation, present: 13 to 25 years after isolation.

Graph density, the overall co-occurrence between species that were captured with T. caesius, sharply increased immediately after fragment isolation for all three fragments size categories. This density sharply declines three years after isolation (i2), where 10 and 100 ha fragments reach 20% density coming down from 46% and 76% respectively, 1 ha-fragments reach zero and remain so for the rest of its isolation periods. Eight years after isolation (i3) 10 ha fragments still decrease in network density reaching a value of zero, but rebounds to 29% at 13 years (i4) after isolation and 39% in the last interval (present). Graph density at 100 ha-fragments at the last three isolation intervals oscillate to higher densities followed by soft decline and a subsequent rebound. These values are not as extreme as seen during the first post-isolation periods (Fig. 4C).
DISCUSSION

Changes in species richness and overall abundance, Species richness and composition, Association to T. caesius

Flocking species composition seems to suffer consistent changes where the patterns seen in species in smaller fragments are dampened in larger fragments. With the exception of 1 ha-fragments, potential flocking species richness do not seem to suffer significant reduction after fragment isolation, in fact, typical flocking species detected in fragments but unassociated to *Thamnomanes caesius* accounted for the majority of non-interactions. This result indicates that local extinctions may not be the principal driver for flock reduction in fragments. Some species, on the other hand, tend to persist and remain strongly associated to *T. caesius* throughout most isolation periods, such as *Automolus infuscatu*, *Xiphorhynchus pardalotus* and *Thamnomanes ardesiacus*. Anecdotal evidence from periods after isolation indicate flock dissolution during the initial periods after isolation corroborate with the data obtained in mist nets (Stouffer and Bierregaard 1995). Additionally, the current association of remaining species were concordant ground observation data obtained in 2009 (Mokross et al. 2014), for example, *T. ardesiacus*, is prone to avoid open areas and tends to remain in fragments, associating to flocks.

Flocking activity does not seem presently occurring in 1 ha-fragments; individuals *T. caesius* were absent during the second interval after isolation (i2), but reappear in the capture record afterwards. The number of captures, however, is comparatively low, and no other flocking species are co-captured. There is indication of a subset of flocking species present in 1 ha-fragments, but after nearly three decades, despite a permeable matrix of low-stands secondary forest, flocks are either not forming, or existing aggregations are not cohesive enough. A third possible mechanism is the formation of aggregations which involve different species; however, to our knowledge, none would be as cohesive or involve as many species as *T. caesius* flocks.
This is corroborated with our previous study; flocks that were actively followed in low-standing secondary forests showed lower numbers of participating species and a significantly reduced flock participation (Mokross et al. 2014). Very degraded areas of secondary forest that were sampled elsewhere in the same landscape showed to be rarely visited by flocks and little permanence of *T. caesius*, thus rendering unstable aggregations over long periods (Mokross in prep.)

The effects on gregarious behavior in flocking activity is considerable in 10 ha-fragments, where despite the presence of a large number of potentially participating species and enough area to hold one flock territory, there is a considerable reduction in the number of co-captures. Currently, 10 ha-fragments seem to be recovering, but despite close numbers of overall captures in the present compared to pre-isolation, associations to *T. caesius* are considerably low. As shown in previous work at the BDFFP, species abundances and richness decreases after an initial crowding effect in the first year after fragment isolation in 1 and 10 ha-fragments. This crowding effect is not apparent in 100-ha fragments, which show little reduction in the first year after isolation and an upswing in capture numbers that surpass pre-capture rates between three to eight years after isolation. It is possible that flocks organize new territories in 100 ha-fragments after two to three years, and eventually, provide refuge for displaced species, thus explaining a higher overall capture density in comparison to pre-isolation periods. It is interesting to note that during the crowding effect in smaller fragments, association rates to *T. caesius* are not increasing; this is especially noticeable in 10 ha-fragments.
Figure 3. Networks of flocking species co-captured with *Thamnomanes caesius*. Nodes represent each species present at an isolation interval at each of the three fragment size categories, the scale of node size to abundance is exponential. Colored nodes represent number of individuals caught with *T. caesius*, while larger transparent nodes are proportional to the overall abundance of species caught in a fragment. After isolation, nodes are color coded according to changes in relation to pre-isolation associations with *T. caesius* at a given isolation interval: red indicates less association to *T. caesius* and green indicates higher association compared to pre-isolation periods. Edges are corrected by to overall abundance in a fragment, meaning that very abundant species with low association frequencies to are given a lower association score.
Network structure

Graph density between flocking species show that the number of species simultaneously captured decreases persistently in all fragments. From a flock perspective this mean that the number of individuals co-occurring in a given flock at any given moment are much lower, at times being reduced to only two or three species foraging together. 10 ha-fragments show a slight recovery towards the end of its isolation history, when the surrounding matrix has become more permeable, allowing the recolonization of species once absent. The low graph density at the last isolation period, however, shows that flocks are still not as aggregate as in pristine environments. This result is consistent with our results from aural and visual censuses from the same 10 ha-fragments during 2010 and 2011, showing these flocks to be very diverse through the course of sampling equivalent to approximately five days, yet, the number of co-occurring species is relatively low when compared to primary forest flocks (Mokross et al. 2014).

Figure 4. Network properties for potential flocking species A. Mean weighted degree, which is a measure of how frequently species are associating. B. Weighted graph density, indicating how well species are associating in comparison to their overall abundance in a fragment at a given isolation interval. C. Marginal species binary graph density is a measure of how species other than T. caesius are captured together.

Despite a sharp increase in overall abundance during the crowding effect, other indicators of flock cohesiveness, such as mean weighted degree and weighted graph density actually dropped. Weighted graph density, another indicator where overall linkages in relation to the
number of individuals present in fragments, seem to weaken in subsequent periods after isolation for all fragments despite the spike in abundances due to birds seeking refuge.

One possibility for the oscillation seen in several measures from isolation to present in 100 ha-fragments, is that flock territories were severely re-structured after isolation. As these fragments dimensions are quite large (roughly 1 km on each side), the process could have taken long enough that crowding effect might have not been as severe as seen in smaller fragments, which were isolated in much less time. Currently, 100 ha-fragments are occupied by 10 to 11 flocks (Mokross et al. 2014b in prep), of which, only two are located at the center with no contact to edges. The remaining 8 to 9 flocks are located at the fringes. It is possible that during the initial isolation periods, there could have been a major relocation of territorial borders, a process that may take considerably more time in terms of reaching stability, which data suggests happening between three to eight years (i3). The subsequent upswing observed between three to eight years could be an influx and concentration of individuals from the surrounding landscape. Curiously, one of the best-structured flocks observed in our sampling from 2010-2011 was at the center of a 100-ha fragment. Whereas I do not discard the influence of habitat structure at the center, this setting within a landscape context seems like another plausible mechanism.

Considerations about mist net-derived sampling

Despite the limitations of mist net sampling (Remsen and Good 1996), long-term databases such as the BDFFP avian capture monitoring program are able to provide an outline of the dynamics in species association patterns. It is possible, to some degree, to have an idea of species composition in flocks throughout isolation periods. Nevertheless, finer points such as species identities should be carefully weighted, as capture probabilities are variable among species and association patterns could be masked due to differences in foraging habits. In this
case, it is highly recommended that observations be backed by thorough understanding natural history of the species in question.

Clustering coefficient would be another way of measuring flock cohesiveness, but considering the potential for false negatives provided in mist net capture data, I believe that an overall network measure such as saturation may provide better estimates.

Concluding remarks

Interspecific flocking is an important strategy that reflects on the fitness and survival of a large subset of species in tropical forests. Organized systems, such as the understory aggregations in Amazonian forests may suffer considerable negative effects from de-structuring due to habitat alterations. Through a well-sampled long-term dataset, it is possible to perceive that differences seen in abundances do not fully explain the dissolution of flocks. Also, the use of this inter relational technique on mist net capture databases opens opportunities for further lines of inquiry that pertain to interspecific associations in other well-sampled sites, given that inferences are well-grounded on natural history as well as fully understanding the limitations of mist net sampling.

LITERATURE CITED

Banks-Leite, C., R. M. Ewers, and J. P. Metzger. 2010. Edge effects as the principal cause of area effects on birds in fragmented secondary forest. Oikos 119:918–926.


APPENDIX A: STEP SELECTION TECHNIQUES UNCOVER THE ENVIRONMENTAL PREDICTORS OF SPACE USE PATTERNS IN FLOCKS OF AMAZONIAN BIRDS

INTRODUCTION

Understanding and quantifying the drivers behind animal movement and space use is a fundamental goal for ecology (Nathan et al., 2008). It is of particular importance in situations where landscapes are changing, making prediction vital for informed conservation (Thomas et al., 2004). The Amazon rainforest is a prime example of a rapidly changing ecosystem, mainly due to wide-scale deforestation (Laurance et al., 2011; Fearnside, 2005; Nepstad et al., 2014). The mixed-species insectivore bird communities that live there are key players in the ecosystem, influencing trophic cascades through herbivorous insects and plants (Mäntylä et al., 2011). Therefore building predictive models of their behaviour is of great importance for understanding how to maintain Amazonia’s rich biodiversity (Chapin et al., 2000).

These flocks are found in practically all terra firme forests in the Amazon basin. They are composed of a wide variety of insectivore species that actively forage in the vegetation (Munn, 1984; Powell, 1989; Mokross et al., 2014). They spend practically the whole daytime searching the different strata and substrates in the vegetation, with high consumption rates. This makes them important contributors to the species richness of the Neotropical avifauna (Powell, 1989).

---

1This chapter previously appeared as:


It is reprinted by permission of Jonathan Potts. Permission letter in appendix E.

Ecology and Evolution, which is published under Wiley & Sons, is under the terms of the Creative Commons Attributions License. No permissions are required for academic reuse.
They typically comprise at least 20 different species at any point in time, and may contain as many as 60, with different species making use of various specialist niches found in the forage (Munn & Terborgh, 1979). Many species are are frequent flock attendants but leave occasionally (flock dropouts), either by switching between flocks, or by having smaller territories than the core species (Jullien & Thiollay, 1998). However, each flock has a core composed of 5-10 species that are consistently present and share the same overlapping territory, each breeding pair defending its territory from conspecifics (Munn & Terborgh, 1979). In the flocks studied here, the Cinerous antshrike (Thamnomanes caesius) plays a nuclear role by giving alarm and rally calls that maintain flock cohesiveness (Munn, 1986). Typically, movement decisions appear to be made by the Cinerous antshrike, but occasionally the core species fail to follow and another direction is taken.

Space use for these flocks is very stable with territory shapes changing little in two decades (Martinez & Gomez, 2013). The core species gather in the same location at dawn every day, usually in a central position within its territory, and will begin foraging from there until sunset where they roost in relatively close vicinity to each other (Powell, 1985; Jullien & Thiollay, 1998; Martinez & Gomez, 2013). The purpose of this paper is to begin a process of disentangling the behavioural drivers behind these movement paths, then to use this understanding to build a predictive model of space use patterns in insectivore bird flocks. Linking animal movement to space use in a quantitative, analytic fashion is vital for predicting the effects of environmental changes on animal populations (Morales et al., 2008). The factors driving the animals’ movement ultimately determine the size and structure of the space that they use in order to meet their everyday needs. By uncovering how these movement processes give
rise to spatial patterns, it would be possible to predict the types of terrain that would be used were the environment to be perturbed, by anthropogenic effects or otherwise (Nathan et al., 2008).

In this paper, we make an important step towards this end, by identifying and quantifying some of the key environmental factors that influence Amazonian bird flock movement, then using them to construct a predictive model of space use. Our approach begins by using a step-selection function (SSF) (Fortin et al., 2005) to test three hypotheses regarding the drivers behind the flocks’ movement decisions. Such techniques, recently reviewed by Thurfjell et al. (2014), have proved invaluable for determining the different drivers of movement in various animal populations. These include foraging decisions in Elk (Forester et al., 2009), memory processes in Bison (Merkle et al., 2014), mechanisms for coexistence of large carnivores (Vanaka et al., 2013), and wolf-ungulate predator-prey interactions (Latombe et al., 2013).

We then derive a master equation from the SSF to link these processes to the emergent space use patterns, following the program started by Moorcroft & Barnett (2008) to integrate resource selection and mechanistic territorial models. The hypotheses we test are that (1) flocks are more likely to move into areas with taller canopies than shorter, (2) flocks tend to move away from higher ground and towards lower, (3) flocks leave some time for the resources to renew before re-visiting a tree they have recently visited.

Taller canopies are expected to be preferable for birds since they tend to contain a larger mass of resources (Basset et al., 1992). Furthermore, both leaf abundance and tree height are known to be positively correlated with insect biomass in certain rainforest trees (Ellwood & Foster, 2004; Campos et al., 2006). In the system studied here, some birds have been seen moving all the way up to the sub-canopy and foraging there (KM, PS pers. obs.), suggesting that
the flocks are making use of the entirety of each tree, so benefitting from the greater available biomass in taller trees. On the other hand, lower ground can support more buffered conditions from wind turbulence and sunlight from outside the forest cover (Ewers & Banks-Leite, 2013), and naturally hold higher air and soil moisture levels (Baraloto & Couteron, 2010) which could potentially increase arthropod loads per vegetation volume (Williams-Linera & Herrera, 2003; Chan et al., 2008).

We begin by examining these two covariates in order to develop a basic methodological framework that we can easily extend to build more complicated models. These could include other factors driving the birds’ movement decisions, such as memory (Smouse et al., 2010), territoriality (Potts et al., 2013), or tighter movement patterns in dense foliage (Jullien & Thiollay, 1998). Building a model one parameter at a time is advantageous since we gain a clear understanding of exactly how, and to what extent, each environmental factor influences flock movement. Though starting with a more complex model may lead to more accurate predictions, it would make it harder to disentangle the relative effects of each model parameter on the resulting space use.

That said, the relative effects of canopy height and topography on movement are interesting in themselves. Indeed, prior to the study, we were unclear whether the two effects were too closely related to be distinguishable. For example, it is known that tree mortality is correlated with steeper slopes in this part of the Amazon (Castilho et al., 2006), so there is an expectation of more disturbance, hence lower canopies, in steeper areas. Trees may also be shorter on slopes due to leaching of soil minerals. One outcome of the hypothesis testing will be to see if both parameters are having an individual and separate effect on bird movement, or whether their effects are closely correlated.
MATERIALS AND METHODS

The step selection function model

Our model for bird-flock movement is based on a step selection function (SSF) approach (Fortin et al., 2005). Following the formalism initiated by Rhodes et al. (2005), but extended here to take into account correlations in the movement, we write the probability \( f(x|y, \theta_0) \) of finding an animal at position \( x \), having traveled from \( y \) in the previous step, given that it arrived at \( y \) on a bearing of \( \theta_0 \) as follows

\[
f(x|y, \theta_0) = \Phi(x|y, \theta_0)w(x,E) \int_\Omega \Phi(x'|y, \theta_0)w(x',E)dx'
\]

Here, \( w(x,E) \) is a weighting function that depends upon the animal’s position \( x \) and some environmental covariates \( E \) (Forester et al., 2009), \( \Phi(x|y, \theta_0) \) is the probability of being at \( x \) in the absence of habitat selection, given that the animal was previously at \( y \) and had arrived there on a bearing of \( \theta_0 \), \( \Omega \) is the study area, and bearings are measured in an anti-clockwise direction from the right-hand half of the horizontal axis. Each step takes a fixed amount of time \( \tau \). The function \( \Phi(x|y, \theta_0) \) allows us to take into account the fact that animals may be more likely to take steps of a particular length, and the distribution of such lengths can be derived from empirical data. For computational purposes, we truncated the step length distribution at steps of greater than 100m, since these never occur in our data. We include the angle \( \theta_0 \) into this formulation to allow for the possibility of correlations between successive movement bearings.

For the purpose of testing hypotheses (1) and (2), \( w(x,E) \) is a function of the canopy height \( C(x) \) and the topography (i.e. elevation above sea level) \( T(x) \), both measured in meters (m). We test two candidate formulations for \( w(x,E) \)

\[
w_a(x, \alpha, \beta) = \exp[\alpha C(x) - \beta T(x)]
\]

(2)
\[ w_b(\mathbf{x}, \alpha, \beta) = C(\mathbf{x})^\alpha T(\mathbf{x})^{-\beta} \]  

Notice that Eq. 3 can also be written as \( w_b(\mathbf{x}, \alpha, \beta) = \exp[\alpha \ln(C(\mathbf{x})) - \beta \ln(T(\mathbf{x}))] \), in keeping with the original formulation of the step selection function from Fortin et al. (2005). Since we would expect the birds to be more likely to move towards lower ground than higher, we place a minus sign before the \( \beta \) in each equation, so that \( \beta \) is expected to be positive. We treat canopy height and topography as two separate variables, noting that there is little or no correlation between the two (\( R^2 = 0.007 \)).

To test hypothesis (3), we assume that the resource amount (i.e. insect biomass) at the start of the day \((t=0)\), is proportional to the canopy height. This relationship was observed by Campos et al. (2006), who gave linear relationships between tree height and biomass for various insect species. As the birds move through an area, they deplete the resources, which take a time \( G\tau \) to renew. The resource amount present at a site at time \( g\tau \) after having been visited is assumed to be \( R(\mathbf{x}, t, G) = gC(\mathbf{x})/G \) as long as \( g < G \), otherwise \( R(\mathbf{x}, t, G) = C(\mathbf{x}) \). Here, \( t \) is the time since start-of-day and a unit of resources is implicitly defined to be the maximum amount of usable resources sustainable by a tree per meter of tree height. At time \( t=0 \), we assume \( R(\mathbf{x}, 0, G) = C(\mathbf{x}) \). As with hypotheses (1) and (2), we test two candidate formulations for \( w(\mathbf{x}, E) \)

\[ w_c(\mathbf{x}, t, \alpha, \beta, G) = \exp[\alpha R(\mathbf{x}, G, t) - \beta T(\mathbf{x})] \]  

\[ w_d(\mathbf{x}, t, \alpha, \beta, G) = C(\mathbf{x}, t, G)^\alpha T(\mathbf{x})^{-\beta} \]  

Notice that when \( G=1 \), we have \( w_d(\mathbf{x}, \alpha, \beta) = w_c(\mathbf{x}, t, \alpha, \beta, 1) \) and \( w_c(\mathbf{x}, \alpha, \beta) = w_d(\mathbf{x}, t, \alpha, \beta, 1) \).

Data collection methods

Flock activity is conspicuous, allowing birds to be followed on foot. As flocks moved, geolocations were recorded at 30 second intervals with a hand-held GPS unit (Garmin Vista HCX, equipped with Wide Area Augmentation System coverage ensure reliable precision under
canopies). The observer maintained a distance of 10-20m from the flocks to ensure no alarm or avoidance behavior was induced in the birds. Observer distance is not in perfect lockstep with the flock, yet the average distance to the approximate center of the flock could be maintained to an accuracy of a few meters. Where possible, we used the location of a Cinereous Antshrike as the flock location. This species was usually conspicuous in the center of the flock. Other more active species typically spread out over a radius of 5-10 m, depending on the size and speed of the flock.

Though GPS error can be around 10m, it is mainly caused by the relatively slow movement of the ionosphere (Parkinson & Spilker, 1996) which only changes by a few centimeters during 30 second intervals. Indeed, evidence from using hand-held GPS for tracking butterflies suggests that the median drift (i.e. absolute error) between consecutive 15 second relocations is only 8cm (Severns & Breed, 2014). Therefore it is reasonable to assume that the measured step lengths and turning angles accurately reflect reality.

Compared to other available methods, these data reflect well the movement of flocks on a small spatio-temporal scale. They provide a high resolution of time sequence that is not possible in radio-telemetry studies, and presently no other techniques allow the gathering of detailed spatial data for passerines of this size. Unlike remote telemetry, this method also allows the direct observation of behavior, so the observer can directly verify whether the recorded spatial locations are corresponding well with the canopy height and topographical maps.

For measuring canopy heights, we used a LIDAR (Light Detection and Ranging) canopy height model (CHM). Similarly, topography (Digital Elevation Models DEM) was acquired using small footprint airborne LIDAR. The derived (post-processed) images from the LIDAR data are 1m/pixel resolution, which we transformed into 10m lattices by bilinear interpolation. LIDAR
data were collected by airborne laser scanning using a Hexagon-LEICA ALS50 PHASE II MPiA sensor of 150 kHz, at 800 m altitude, with 24 degrees opening, 118 MHz pulse rate, 58 Hz scan rate, 3.7 points/m² density. Swaths were of 340 meters wide, spaced at 240 meters. Post-processing used a Forest Service methodology to generate DEM and CHM at 1 m²/pixel [see Stark et al. (2012) for more details on LIDAR data collection and analysis].

Sampling was restricted to areas within LIDAR coverage which did not span more than 1.5 km², and of which five were scattered along the study area at an average of 6 km from each other. The choice of flocks was mainly restricted to locations where the entire home range would be inside this LIDAR coverage (i.e. away from edges of the Canopy height models and Digital elevation models). We analyzed 6 different flocks from the Dimona LIDAR dataset, which was the largest (2.8 km by 1.5 km) and best-sampled area, and also the one that presented the highest variability in vegetation types. This area falls within the Biological Dynamics of Forest Fragments Project (BDFFP), about 70 km north of Manaus, Brazil (see http://pdbff.inpa.gov.br/ for maps).

Data were gathered during the dry seasons between June and November during 2009-2011 and each flock was tracked for between 5 and 11 days. Each flock gathers in one particular place each day, then moves around the forest for a total of about 11.5 hours during the day, before each bird goes back to its roost for the night. Flock composition was sampled every half-hour to check that cohesiveness was being maintained. Flocks were taken from a variety of different habitat types to ensure the greatest generality in our findings, and minimize the effects of spatial autocorrelation. These included areas predominated by secondary forests, areas of primary forest away from edges, and areas near forest edges. Flocks were initially found based on their dawn gatherings. As they were first followed, it was unclear where they would go, so it
is unlikely there was a bias to the flocks chosen. If the flocks moved into areas that were difficult for the observer to reach, bearings in relation to the observer were taken in order to make the necessary corrections in the data. In these cases, the observer did not lose the flock from sight.

**Parametrizing the models from the data**

The first step in parametrizing the models is to calculate the step length and turning angle distributions, i.e. the distance between successive positions and the angle an animal turns through from one move to another, respectively [see e.g. Crist et al. (1992)]. Since these depend upon the temporal resolution $\tau$ (i.e. the time between successive position fixes), we use both $\tau=1$ minute and $\tau=5$ minutes, deriving two different sets of step length and turning angle distribution for the different values of $\tau$. The value $\tau=1$ minute is chosen because bird flocks tend to move from one tree to another at an average of approximately every 1 or 2 minutes. Though their movement is a continuous rather than discrete process, the model is formulated so this timescale roughly represents the small-scale decisions that the birds make regarding whether they stay in a tree or choose to move to another. We also examine the case $\tau=5$ minutes to determine whether the decisions about where to move can instead be viewed as taking place on a timescale longer than a single jump between trees. In other words, the birds might only be considering the next tree they move to when deciding where to go ($\tau=1$ minute), or they might be thinking a few trees ahead when they make this decision ($\tau=5$ minutes).

The step length distributions are fitted to both a Weibull distribution (Forester et al., 2009) and an Exponentiated Weibull (EW) distribution (Nassar & Eissa, 2003), using the Akaike Information Criterion (AIC) to determine the best model, whereas we fit the turning angles to a von-Mises distribution (Marsh & Jones, 1988). The Weibull, EW and von Mises distributions have the following forms, respectively.
\[ \rho_1(x|a, b) = \frac{a}{b} \left( \frac{x}{b} \right)^{a-1} \exp \left[ -\left( \frac{x}{b} \right)^a \right], \quad (6) \]
\[ \rho_2(x|a, b, c) = \frac{ac}{b} \left( \frac{x}{b} \right)^{a-1} \exp \left[ -\left( \frac{x}{b} \right)^a \right] \left[ 1 - \exp \left[ -\left( \frac{x}{b} \right)^a \right] \right]^{c-1}, \quad (7) \]
\[ V(\phi|k) = \frac{\exp[k\cos(\phi)]}{2\pi I_0(k)}. \quad (8) \]

Since the rainforest canopy consists of distinct treetops whose widths are each roughly 10m across, we split the terrain \( \Omega \) into a grid \( S \) of 10m by 10m squares. This allows us to associate a value of \( C(s) \) and \( T(s) \) to each square \( s \) in \( S \), respectively the mean canopy height and mean topography of the square. Canopy heights ranged from 50m to essentially none, which we set to be 1m for the purpose of the model (a value of zero metres for the canopy height would give an identically zero probability of moving there, which is biologically implausible). Topography ranged from 40m to 115m. Parametrizing Eq. 1 from the data therefore requires maximizing the following likelihood function

\[ L(X|E) = \prod_{n=2}^{N} \frac{\Phi(x_n|x_{n-1}, \theta_{n-1})w(x_n, E)}{\sum_{s \in S} \Phi(s|x_{n-1}, \theta_{n-1})w(s, E)}. \quad (9) \]

where \( X = \{x_0, x_1, \ldots, x_N\} \) are the consecutive positions of a flock, \( \theta_n \) is the bearing from \( x_{n-1} \) to \( x_n \), \( \Phi \) is the product of the best-fit step length and turning angle distributions, and \( w \) is either \( w_a \), \( w_b \), \( w_c \) or \( w_d \), depending on which model we are fitting.

To test hypothesis (1), we fix \( \beta = 0 \) and find the value of \( \alpha \) that maximises \( L(X|E) \), which we call \( \alpha_m \). We then use the likelihood ratio test to compare the resulting value of \( L(X|E) \) with the value of \( L(X|E) \) when both \( \alpha \) and \( \beta \) set to zero. For hypothesis (2), we fix \( \alpha = \alpha_m \) and find the
value of $\beta$ that maximises $L(\mathbf{X}|E)$, again using the likelihood ratio test to compare this value of $L(\mathbf{X}|E)$ with the one where $\alpha = \alpha_m$ and $\beta = 0$.

This technique of fixing $\alpha = \alpha_m$ when testing hypothesis (2) means that we are only testing for topographical effects on movement that are additional to the effects of canopy height. This is to address the question of whether these effects can be separated (see the last paragraph of the Introduction). We then find the values of $\alpha$ and $\beta$ that maximise $L(\mathbf{X}|E)$ by varying both parameters simultaneously, giving best fit values denoted by $\alpha_{bf}$ and $\beta_{bf}$. We use a Markov bootstrap method with 100 bootstraps to find standard errors for $\alpha$ and $\beta$ (Horowitz, 2003). Hypothesis (3) is tested by fixing $\alpha = \alpha_{bf}$ and $\beta = \beta_{bf}$ and finding the value of $G$ that maximises $L(\mathbf{X}|E)$, then using the likelihood ratio test to compare the resulting value of $L(\mathbf{X}|E)$ with the value of $L(\mathbf{X}|E)$ when $G=1$. For each maximization calculation, we use the Nelder-Mead simplex algorithm (Lagarias et al., 1998), as implemented in the Python `maximize()` function from the SciPy library (Jones et al., 2001).

**Constructing the space use distribution**

We use two methods for constructing the space use distribution from the parametrized SSF (Eq. 1), via simulation analysis and through constructing the master equation and numerically deriving its steady-state solution. For the former approach, we simulate one particular flock’s movement on the grid $\mathbf{S}$ using the jump probabilities given by SSF. Since the flock gathers in one particular place each day, and moves around the terrain for a total of about eleven-and-a-half hours during the day, we start the simulated birds at the gathering point and run the simulation for 138 time steps, each step representing $\tau = 5$ minutes (giving 11 hours 30 minutes in total), taking a note of all the positions at which the flock landed after each step. We repeat this 100 times, representing 100 days, giving 13,800 simulated positions in total. In the
data, we tend to have around 10 days per flock. However, we use 100 here to average out some of the stochasticity. From these simulated positions, we calculate the 50%, 60%, 70%, 80%, and 90% Kernel Density Estimators (KDEs), using a fixed kernel method with smoothing parameter
\[ h = \sigma n^{-1/6} \] where \( \sigma = (1/2) \sqrt{\frac{2}{\sigma_x^2 + \sigma_y^2}} \) and \( \sigma_x, \sigma_y \) are the standard deviations of the simulated data in the x- and y-directions respectively (Worton, 1989). KDE calculations are performed using Python. In addition to simulation analysis, we also construct the master equation for the probability density function \( u(x, \theta, t) \) of the animal being at \( x \) at time \( t \) having traveled there on a bearing of \( \theta \). This allows us to compare our results with the predictions of Barnett & Moorcroft (2008), who mathematically analyzed the step selection function (Eq. 1) in the simpler case where the turning angle distribution is uniform. They proved that the steady state (time-independent) solution \( u^*(x) \) is proportional to \( w(x, E) z(x, E) \), where \( (x, E) = \int_\Omega \Phi(x' | x, \theta) w(x, E) dx' \) is a local averaging of \( w(x, E) \). We examine to what extent this result extends to our more complicated situation of a correlated random walker. We use Eq. 1 to construct the following master equation
\[
\begin{align*}
u(x, \theta, t + \tau) &= \int_{-\pi}^{\pi} d \theta_0 \int_{0}^{r_{\text{max}}} dr \frac{\phi(x | y_\theta(r), \theta_0) w(x, E)}{\int_{\Omega} dx' \phi(x' | y_\theta(r), \theta_0) w(x', E)} u(y_\theta(r), \theta_0, t) \\
\end{align*}
\] (10)
where \( y_\theta(r) \) describes the locus of points \( y \) upon which the animal could approach \( x = (x_1, x_2) \) at bearing \( \theta \), i.e. \( y_\theta(r) = (x_1 + \cos(\theta + \pi) r, x_2 + \sin(\theta + \pi) r) \), with \( r \) denoting the distance between \( y_\theta(r) \) and \( x \) (Potts et al., 2014). Here \( r_{\text{max}} \) is the distance along this line from \( x \) to the boundary of \( \Omega \) and so gives the upper endpoint of integration. To calculate the steady-state distribution, we solve Eq. 10 iteratively until \( |u(x, \theta, t + \tau) - u(x, \theta, t)| < 10^{-8} \) for every value of \( x \) and \( \theta \). The area \( \Omega \) for this calculation is defined to be the 95% KDE of the flock used for the simulations. We used zero-flux boundary conditions, which models the fact that the birds
are confined within their territory. Calculations were coded in C and it took approximately 2 hours to find a single steady state distribution.

Note that in these methods, we are separating the fitting of the turning angle and step length distributions from the fitting of the weighting functions. This makes the maximization procedure far faster and means the algorithms are more likely to converge to the global maximum. However, if the weighting function w gives a particularly strong selection for an environmental covariate and/or the step length distributions are fat-tailed, then this separation may cause inaccuracies in the resulting model. To test that this is not the case, we calculated the mean and standard deviation of the step length and turning angle distributions from the above simulations to verify that the weighting function had not significantly altered them.

RESULTS

Step length and turning angle distributions

For both cases $\tau=1$ minute and $\tau=5$ minutes, the best fit step length distribution is an Exponentiated Weibull (EW) distribution (Fig. 1). For $\tau=1$ minute, $\Delta \text{AIC}=126.9$ between EW and Weibull. For $\tau=5$ minutes, $\Delta \text{AIC}=14.6$. The step length distributions both increase from 0m initially, before decaying (Fig. 1). However, this is not an indicator that birds are more likely to move a medium length distance than a very short distance, but is simply due to there being less area in the annulus of radius between $r$ and $r+\delta r$ when $r$ is smaller. If $\delta r$ is small then the total amount of area into which a flock can move, given that it moves a distance between $r$ and $r+\delta r$, is approximately $\delta r \times 2\pi r$, which is proportional to $r$. To find the relative preferences of the birds to move a particular distance, it is therefore necessary to divide the probability density, $P(r)$, by the distance moved, $r$. If we do this for our data on the 1 minute temporal resolution, we find that
\( P(r)/r \) is approximately \( 0.044 \exp(-r/4.75) \) and for the 5 minute time-scale

\( P(r)/r \approx 0.0080 \exp(-r/11.3) \), both of which decay monotonically as \( r \) increases.

**Hypothesis testing**

The tests indicate that there is a significant effect of both canopy height (hypothesis 1) and topography (hypothesis 2) on the flocks’ movement (Table 1). Furthermore, these aspects of the landscape each affect bird movement separately, rather than being highly intertwined. However, accounting for resource renewal, so that birds are less likely to re-visit trees that they have recently visited, does not improve the model fit (hypothesis 3). The conclusions are the same both for \( \tau=1 \) minute and \( \tau=5 \) minutes, so we cannot conclude anything about the temporal resolution on which decisions are made.

Figure 1. Step length and turning angle distributions. Panel (a) shows the empirical step length distribution (bars) for data where the temporal resolution is \( \tau=1 \) minute, together with the best fit Exponentiated Weibull distribution (solid curve). The latter is given in Eq. 7, with \( a=1.06, b=6.90 \) and \( c=1.82 \). The bars in panel (b) denote the empirical turning angle distribution for the same data, whereas the curve denotes the best fit von Mises distribution, given in Eq. 8 with \( k=0.336 \). Panels (c) and (d) are analogous to (a) and (b) respectively, except they use the data set where \( \tau=5 \) minutes, rather than \( \tau=1 \) minute. Here, \( a=1.26, b=17.2, c=1.55 \) and \( k=0.637 \).
To put these in a biological context, consider two trees, equally accessible over a 5 minute interval and on ground of equal elevation, but one A% taller than the other, e.g. if one is 30m high and the other 20m high then A=50. Then the birds are \((1+A/100)^{0.277} = 1.5^{0.277} \approx 1.1\) times more likely to move to the taller tree than the shorter, i.e. about 10% more likely. The effect is more dramatic when considering the difference between a completely deforested area with essentially no canopy (which we set to 1m for the purposes of the model) and primary forest with, say, 30m canopy. Here, A=3000, so the flocks will be around 160% more likely to move to the primary forest.

Conversely, suppose that both trees are of equal height but one tree is ground B% higher above sea-level than the other. Then the birds are \((1+B/100)^{1.697}\) times more likely to move to the tree on lower ground. For example, an decrease from 50m to 40m elevation leads to a \(1.25^{1.697} \approx 1.460\) increase in probability of moving there, i.e. they are 46% more likely to move to the 40m elevation.

The weighting function \(w_b\) (Eq. 3) provides a better fit to the data than \(w_a\) (Eq. 2) for \(\tau=5\) minutes. The AIC for \(w_b\) is lower than that for \(w_a\) (\(\Delta AIC=3.8\)). Though the AIC for \(w_b\) for \(\tau=1\) minutes is slightly lower than for \(w_a\) (\(\Delta AIC=0.1\)), the change in AIC is not large enough to be considered good evidence that \(w_b\) is better than \(w_a\). In Table 1 we detail the results for the function \(w_b\) and its generalization \(w_d\) (Eq. 5). Results for \(w_a\) and \(w_c\) (Eq. 4) are qualitatively similar.

**Space use distributions**

Figure 2 compares the simulated space use with the empirical data on flock positions. The KDE contour lines for the simulated data are quite tightly packed around the edge of the empirical data points, suggesting that the model is giving a reasonable prediction of space use.
patterns. However, the extent of the simulated home range is clearly larger than the empirical home range. Though separating the fitting of the step length and turning angle distributions from the environmental interactions may mean that the fit is less accurate than if all parameters were fitted together, it turns out that the mean of the simulated data’s step length distribution is $20.05\pm0.95$ m (95% confidence intervals), compared with 20.09 m from the data. The standard deviation of the simulated step lengths is $13.55\pm2.01$ m as compared with 13.23 m from the data. Similarly, the standard deviation of the turning angles from simulation output is $82.1\pm8.7$ degrees as compared with 82.7 degrees from the data, and the mean is $-0.2\pm6.9$ degrees, as compared with $-1.7$ degrees from the data. Therefore including the weighting function does not significantly change the step length or turning angle distributions.

Figure 2. Plots of simulated and real data. Both panels show the empirical data for one flock (dots) together with the 50%, 60%, 70%, 80% and 90% kernel density estimation curves for the simulated data (black curves). See the Methods section for details on how the simulations were performed. The colors underlying panel (a) denote the canopy height, whereas in panel (b) they give the topography, i.e. height of the ground above sea level.

**Comparison with analytic results**

Previous work showed that if there is no correlation in an animal’s movement, the steady-state space-use distribution is proportional to $w(x, E)z(x, E)$ as long as the turning angle distribution is uniform (Barnett & Moorcroft, 2008, Eq. 13), where $z(x, E) =$
By numerically deriving the steady space-use distribution for our model, we show that this result breaks down when we include correlation in the movement process. Figs. 3a and 3b compare the analytic result to the numerical one in the specific example of our Amazonian bird flock model, in the case \( w(x, E) = w_b(x, \alpha, \beta) \) (see Eq. 3). However, if we assume that the turning angle distribution is uniform, then the analytic solution is very similar (Figs. 3b and 3c).

**DISCUSSION**

We have constructed a step selection function (SSF) to test three hypotheses about the drivers behind Amazonian bird flock movement decisions. We found that flocks have a tendency to move towards areas covered by higher canopies, but move away from areas of higher ground. The preference for higher canopies is likely to be due to the greater abundance of resources, through enhanced micro-climatic conditions in the understory and more foraging substrate (Basset et al., 1992). Lower ground is likely to be preferred because it has a moister environment.
that can hold a higher insect biomass (Chan et al., 2008; Bueno et al., 2012). Though these aspects are related, we found no evidence of correlation between topography and canopy height, and each appear to have their own separate effects on flock movement.

The flocks appear to be just as likely to move back to a place that they have recently visited than one that they have not visited for a while. This suggests that when they visit a tree, they do not deplete the resources completely, but leave the tree in the knowledge that there is still food to be found there. Whilst it may seem advantageous to stay at a tree as long as it is profitable to do so, in order to conserve energy (Houston et al., 1993), this frequent movement from tree to tree might be a tactic to avoid predators. Alternatively, insects may temporarily be adopting cryptic behavior on the presence of birds, thus forcing the birds to move on quickly as insects become rapidly harder to find.

We tested different functional forms for the selection weighting, something that is rarely done in literature on step selection functions but could be important (Lele et al., 2013). Although we would be surprised if the functional form were to change the outcome of hypothesis testing, it could very much affect the resulting parameters that are used to build the mechanistic model. For example, an exponential effect of the canopy height vastly increases the relative attraction to very high canopies as compared with a power law effect, since this is effectively the difference between a linear and a logarithmic scaling (see the note after Eq. 3). This has the potential to vastly change the predicted space use patterns. Therefore it is vital to consider functional form when using step selection techniques to build mechanistic models.

Our SSF approach enabled us to run simulations that were used to predict the utilization distribution (UD) of a flock, thereby relating the small-scale movement decisions to the large-scale space use patterns. While the resulting simulated UD captured certain qualitative aspects of
the empirical data (Fig. 2), it overestimated the home range size. In comparison, a straightforward random walk model, based on the empirical mean step length distribution, would give a normal distribution with the 90% contour approximately 395m from the gathering point. This contour would overlap the corresponding (outer) contour from Fig. 2, but would be circular, whereas the simulation contour is far from symmetric. Therefore, though certain features of space use are being predicted by our model, there must be some other aspect of the birds’ movement decisions keeping them far more spatially confined than our current model predicts.

We propose two plausible mechanisms that might explain this confinement. First, these flocks are highly territorial (Develey & Stouffer, 2001), so interactions with neighboring flocks may cause each flock to use less space than they would otherwise. The mechanism of conspecific avoidance has been shown to give rise to spatial confinement in various species of canid (Lewis & Murray, 1993; Moorcroft et al., 2006; Potts et al., 2013). These all deal with avoidance via scent marking, whereas territories in birds are defended via vocalizations and direct interactions (Munn & Terborgh, 1979). However, the generic modeling framework from Potts et al. (2013) could be used to constructed coupled SSFs, whose weighting functions $w$ depend both upon the position of the individual and on interactions with neighbors. These interactions may either be direct or mediated by vocal, visual or olfactory cues.

Second, memory effects, with birds having a preference to move back towards places they have frequently visited, can cause spatial confinement. Theoretical studies by Briscoe et al. (2002) have described such a mechanism in wolf (Canis lupus) populations, and the general results of Tan et al. (2001) show that memory can severely constrain the amount of area used in a given time period. Though it is tricky to determine empirically what constitutes a bird’s
cognitive map of the environment, it is generally considered that memory is an important factor in the spatial confinement and site-fidelity of many animals (Smouse et al., 2010).

By turning our SSF into a master equation for the spatio-temporal probability distribution of the flock’s position, we compared our results to a recent approximate analytic prediction by Barnett & Moorcroft (2008) that applies when the turning angle distribution is uniform. However, their results fail whenever there is correlation in the animal’s movement at any time-scale, a fact noted in Barnett & Moorcroft (2008) but not emphasized in their ecologically motivated paper Moorcroft & Barnett (2008). The more the correlation, the worse the prediction is likely to be, so it is necessary to take care when applying these results to empirical data. Though the correlation in the birds’ movement greatly affected the movement patterns, when we removed any intrinsic correlation from our movement model, the predictions of Barnett & Moorcroft (2008) were visually very good (Fig. 3).

Though our results are not testing conservation decisions per se, the application of these models could provide basis for informed management decisions for a subset of the avian community that is known to be very sensitive to forest disturbances. By providing information on how a combination of two important habitat features influences habitat use and how these flocks anchor their home ranges, this would allow for more realistic estimations of areas that are more important to these species. Also, the drivers related to resource abundance and renewal provide important insights on the nature of the relationship of insectivorous birds and their resource, a topic that has challenged researchers for years (Sherry, 1984; Şekercioğlu et al., 2002). These results also have the potential to be extended to closely related species in other regions of Amazonia. For example, in Southwestern Amazonia flocks are lead by T. schistogynus rather than T. caesius (Munn & Terborgh, 1979) which may behave differently. It is an
interesting future challenge to analyse these differences rigorously. The dynamic and collective nature of bird flock decisions is also likely to have an impact on behavioural decisions. As we refine our model to make it more accurate at predicting space use, it will likely be necessary to take these effects into account.

LITERATURE CITED


INTRODUCTION

Recent years have seen an explosion in the number of studies devoted to collective animal movement modelling, largely enabled by the availability of cheap, fast computational power and vastly improved tracking data (Sumpter 2006, Sumpter 2010). They have succeeded in explaining a wide variety of patterns observed in nature due to the movements and interactions of animals (Deneubourg and Goss 1989, Couzi et al. 2002), such as bird flocking (Nagy et al. 2010), ant raids (Deneubourg et al. 1989), and fish schooling (Hoare et al. 2004). Furthermore, in the last few years, the collective behavior paradigm has been extended to include territorial and home range patterns, which arise from conspecific avoidance mechanisms rather than those of alignment or attraction (Moorcroft and Lewis 2006, Giuggioli et al. 2011, Potts et al. 2013).

Despite these myriad advancements, collective animal models remain disparate and varied, with no quantitative formulation of a unifying framework encompassing the variety of interaction mechanisms. In particular, territorial models have followed two separate modelling paradigms. They first began with Lewis & Murray (Sumpter et al. 2012) and has been developed

---

2This chapter previously appeared as:


It is reprinted by permission of Jonathan Potts and the Royal Society Publishing. Permission letter in appendix E.
through a number of papers over the past 20 years, e.g. Moorcroft et al. (Herbert-Read et al. 2011, Katz et al. 2011), culminating in the book by Moorcroft & Lewis (Moorcroft and Lewis 2006). It involves constructing partial differential equations (PDEs) either from details of the underlying movement and interaction processes or from more phenomenological descriptions, and then using these equations to derive territorial patterns mathematically. The second approach is based more on statistical physics, analyzing the individual movement and interaction processes themselves in discrete space, without taking a mean-field continuum limit (Giuggioli et al. 2011, Gautrais et al. 2012). A recent review explains the biological lessons that can be learned from these models (Strandburg-Peshkin et al. 2013).

Their differences notwithstanding, both approaches employ a similar method of data analysis, by fitting the emergent territorial patterns to positional data. Whilst this is a reasonable way of testing hypotheses about the underlying causes of spatial patterns (Katz et al. 2011), it is not sufficient for concrete quantification of the underlying movement and interaction processes, since many different model processes can give rise to the same emergent spatial patterns. Therefore when constructing a unifying framework for extracting interaction mechanisms, it is important that the quantification happens on the same level that the model is constructed, rather than on the level of the model’s emergent features.

In areas outside territory formation, several recent examples from the collective behavior literature have pioneered a variety of techniques for doing exactly this (Pérez-Escudero 2011). For example, fish repulsion-alignment-attraction mechanisms have been measured using several different techniques (Mann et al. 2014, Ballerini et al. 2008, Petit et al. 2013), as have the geometric nature of their interactions (Mann 2011) and their decision processes (Moorcroft et al. 1999, Moorcroft et al. 2006). Others examples include alignment and leadership decisions in bird
flocks (Lewis and Murray 1993, Nagy et al. 2010, Potts et al. 2012). Though there have been several theoretical studies aimed at more general application (Moorcroft et al. 1999, Potts and Lewis 2014), the techniques and models used to extract interaction mechanisms are varied and are often designed only for use with one specific system.

Parallel to the collective animal literature, many studies have sought to understand and predict space use patterns by examining interactions between animals and their environment. Resource selection analysis is perhaps the widest used tool in this regard, and has a long history (Manly et al. 2002). Recently, this has been integrated with animal movement processes by constructing step selection functions (Fortin et al. 2005, Rhodes et al. 2005, Forester and Rathouz 2009), which allow model building by rigorously deriving parameter values from the data. These, in turn, have been used to build mechanistic models to derive space use patterns from the underlying movement processes and animal-environment interactions (Thurfjell et al. 2014), representing the first step in unifying resource selection with mechanistic models.

Some studies in the step-selection function literature have factored into their analysis either positions of other individuals (Moorcroft and Barnett 2008) or traces left in the environment by animals (Vanak et al. 2013). However, to model simultaneously more than one interacting group of animals, so that it is possible to build a mechanistic model to predict the resulting space use patterns, would require having different interacting movement kernels for each group. For example, in a territorial system there would be one function for each group maintaining a territory. These would then have to be coupled together so that each function depends on the animals modelled by the other.
In this paper, we present a modelling framework that unifies movement with both animal environment and inter-animal interactions. The inter-animal interactions may either be direct, or mediated by a stigmergent process (Latombe et al. 2013, Theraulaz and Bonabeau 1999) such as pheromone deposition or visual cues. Our framework includes as special cases both step selection functions and the two approaches to mechanistic territorial modelling mentioned above. Though we focus specifically on combining territorial interactions into the concept of step selection, our framework also happens to incorporate a variety of other collective motion models, suggesting far broader application (see Table 1). As such, our framework provides a useful way to codify movement and inter-action processes, giving a generic starting point for modelling these processes and a clear way of testing which combinations of them best describe the underlying data. This will both help future researchers in model construction and provide a concrete means by which to compare and contrast different modelling approaches.

We show how to use our model to test hypotheses about the interaction mechanisms underlying territorial behavior, by application to movement data on a community of territorial insectivore bird flocks in the Amazon rainforest. Parameter values for a model of movement and territorial interactions naturally arise from this hypothesis testing. This model can then be analyzed either using the PDE techniques of Moorcroft & Lewis (Moorcroft and Lewis 2006) or by simulation analysis (Giuggioli et al. 2011, Gautrais et al. 2012). This enables the spatial territorial patterns to be derived from the underlying movement and interaction processes, which can be compared with spatial data. We demonstrate how to make this comparison quantitative, thereby giving a technique for determining which the processes are the key drivers of space use in the study population. This ability of our framework to derive ‘macroscopic’ patterns in a
quantitative and non-speculative fashion from a wide variety of ‘microscopic’ movement and interaction processes represents a significant step forward in the general program of building a ‘statistical mechanics for ecological systems’ (Fig. 1) (Giuggioli et al. 2013).

Figure 1. The role of coupled step selection functions in linking movement to emergent spatial patterns. This schematic represents the central place of coupled step selection functions for the program of constructing a statistical mechanics for ecological systems. Different candidate models M1, . . . , MN can be tested against ‘microscopic’ movement and interaction data using the techniques in the Methods section. The best models can then either be simulated or mathematically analysed to derive spatial patterns. These, in turn, can be compared to the ‘macroscopic’ spatial distributions in the data (see Methods) to test whether the mechanisms being modelled are sufficient for accurate predictions of spatial patterns.

MATERIAL AND METHODS

Modeling framework

Our model is based around the notion of a step selection function (Fortin et al. 2005). However, simultaneous modelling of various interacting animals, or groups of animals, requires having a different step selection function for each animal or group. Therefore, instead of having
one function that models all agents, as with previous approaches, we construct a different
function for each agent and link them together with a coupling term. We use the term ‘agent’
here to refer to either a single animal, or a group of animals that are modelled as moving together
as a single entity, for example a pack or a flock. The result is what we call a system of coupled
step selection functions (CSSFs), where each function has the following form
\[
f_{i}^{t+\tau}(x|y, \theta_0) \propto \phi_i(x|y, \theta_0)\mathcal{W}_i(x, y, E)\mathcal{C}_i(x, y, \mathcal{P}_i^t),
\]
represented pictorially in Fig. 2. The function \(f_{i}^{t+\tau}(x|y, \theta_0)\) is the probability of agent \(i\)
moving to position \(x\) at time \(t+\tau\), given that the agent was at position \(y\) at time \(t\) and had arrived
there on a bearing \(\theta_0\). The term \(\phi_i(x|y, \theta_0)\) represents the movement process of agent \(i\),
disregarding the effect of the environment or other agents. For example, this could contain the
step length and turning angle distribution for a correlated random walk (Boyce et al. 2002).

The function \(\mathcal{W}_i(x, y, E)\) is a weighting function containing information about the
desirability of moving across the environment \(E\) from position \(y\) to \(x\). For example, if there is a
partial barrier to movement between \(y\) and \(x\) then \(\mathcal{W}_i(x, y, E)\) may be lower than if the barrier
were not there. On the other hand, if \(x\) were in a very desirable habitat for the agent compared to
\(y\) then \(\mathcal{W}_i(x, y, E)\) would be higher than if the habitats were equal in quality. See Fortin et al.
(2007) for a good example of the variety of animal-environment interactions that can be
modelled this way.

The collective aspects of motion, i.e. the agent-agent interactions, are represented by \(\mathcal{C}_i(x, y, \mathcal{P}_i^t)\). The term \(\mathcal{P}_i^t\) represents both the population positions and any traces of their past
positions left either in the environment or in the memory of agent \(i\). For example, if the agents
were schooling fish then perhaps the pertinent interactions would be direct (Eftimie et al. 2007).
However if the agents were ants then $\mathcal{P}_i^t$ might represent the pheromones left by other ants, to which ant $i$ responds by tending to move up the pheromone gradient (Denoubourg et al. 1989). As a third example, if the agents were territorial bird flocks then $\mathcal{P}_i^t$ might include the memory that the birds in flock $i$ have of past territorial conflicts or vocalizations. This enables us to convert ostensibly non-Markovian processes, such as memory and correlations, into a one-step Markov process, possibly requiring high dimensions to encapsulate $\mathcal{P}_i^t$ appropriately. Since $f_{i,t|t}(x|y, \theta_0)$ is a probability, it must integrate or sum to 1, depending on whether continuous space or discrete space is being used, respectively. Therefore we use the $\propto$ sign in equation 1, noting that this becomes an equality if the right hand side is divided by the integral (continuous space) or sum (discrete space) over the possible target positions $x$.

We demonstrate the generality of our formalism by showing that it reduces to ordinary step selection functions (Fortin et al. 2005), resource selection functions (Jones et al. 2001), and a variety of previously published examples of collective motion models. The latter include models of trail-following ants (Deneubourg et al. 1989), collective patterns in animal populations through alignment and attraction (Couzin et al. 2002, Bovet and Benhamou 1988), and territorial canids (Moorcroft and Lewis 2006, Gautrais et al. 2012, Potts and Giuggioli 2013).

It is possible to generalize equation (1) further by writing the right-hand side as an arbitrary function of $x, y, t, \theta_0, \mathcal{E}$ and $\mathcal{P}_i^t$. This would enable the construction of dependencies between the three aspects of movement, environmental interactions, and collective interactions. For example, if the animal’s speed varies over time due to seasonal changes, or if the turning angle distribution is effected by habitat type, and so forth. However, the models from both
previous collective animal behavior studies and the step/resource-selection literature tend not to incorporate such dependencies, since they appear as examples of equation (1). Therefore, for simplicity, we treat the functions \( \varphi_i \), \( W_i \) and \( C_i \) as independent.
Application to bird data

As a demonstration of how to apply our model, we use movement data on a community of territorial insectivore bird flocks in the Amazon rainforest. These flocks are multi-species, with around 5-10 mating pairs consistently present sharing a territory (Munn and Terborgh 1979). Each pair will defend its territory from conspecifics, using a mixture of vocalizations and direct territorial conflicts (Jullien and Thiollay 1998). The birds from each flock meet together at a ‘gathering point’ at dawn every day, usually in a central position within their territory, from where they forage within the territory for around 11-12 hours, moving together as a flock. We use flock movement data from eleven different territories to test hypotheses about the territorial interaction mechanisms used by the birds. We focus, for simplicity, on the vocal aspect of interactions. Vocalizations make neighboring flocks aware of areas they have recently visited, causing the neighbors to alter their movement processes in or near these areas.

We test three hypotheses: whether (1) flocks are likely to avoid areas that neighbors have visited in the past, due to the vocalizations made there, (2) flocks tend to move back towards their gathering site having visited such an area, (3) the time since the area was visited by a neighbor affects the response of the flock, so that old vocalizations are ignored. This demonstrates the ability of our modelling framework to select between competing theories about the nature of interaction mechanisms.

We analyzed movement of 11 different flocks in the Amazon rainforest over 3 years during the dry season between June and November. The study site is about 70 km north of Manaus, Brazil (see http://pdbff.inpa.gov.br/ for maps). They were each tracked for between 4 and 18 days. The flock positions were recorded every minute during the time that they were active. Flock activity is conspicuous, so that birds can be followed on foot. As flocks moved,
geolocations were recorded with a hand-held GPS unit (Garmin Vista HCX). The observer maintained a distance of 10-20m from the flocks to ensure no alarm or avoidance behavior was induced in the birds.

To examine which territorial interaction processes best fit these data, we constructed a coupled step selection function (Eq. 1) where the terms \( \phi_i(x|y, \theta_0) \) and \( \mathcal{W}_i(x, y, E) \) were obtained from a previous study on the same population (Potts et al. 2014). In that paper, we found that setting \( \phi_i \) to be a product of the exponentiated Weibull distribution (Nassar and Eissa 2003) for the step lengths and a von Mises distribution (Marsh and Jones 1988) for the turning angles fitted the data well. This led to the following distribution

\[
\phi_i(x|y, \theta_0) = \frac{ae}{b} \left( \frac{|x-y|}{b} \right)^{a-1} \exp \left[ -\left( \frac{|x-y|}{b} \right)^a \right] \times \\
\left\{ 1 - \exp \left[ -\left( \frac{|x-y|}{b} \right)^a \right] \right\}^{c-1} \frac{\exp[k \cos(\theta - \theta_0)]}{2\pi I_0(k)},
\]

where each agent \( i \) is an individual flock, \( \theta \) is the bearing from \( y \) to \( x \), \( a = 1.06 \pm 0.03 \), \( b = 6.90 \pm 0.34 \), \( c = 1.82 \pm 0.11 \), \( k = 0.336 \pm 0.015 \) (error bars are 1 standard deviation) and \( I_0(k) \) is a modified Bessel function of the first kind. The best fit model from Potts et al. (2014) for the \( \mathcal{W}_i \) term is \( \mathcal{W}_i(x, y, E) = C(x)^\alpha T(x)^{-\beta} \), where \( C(x) \) and \( T(x) \) are, respectively, the forest canopy height and topography in meters, at position \( x \). The time-interval \( \tau \) is 1 minute and the best fit values for the parameters are \( \alpha = 0.0952 \pm 0.037 \) and \( \beta = 1.658 \pm 0.345 \) (error bars are 1 standard deviation). These were derived by performing the model fit whilst neglecting interaction mechanisms (see (41) for details). For the interaction term \( \mathcal{C}_i(x, y, P_i^f) \), we used various modifications of the territoriality model from Potts et al. (2013), used to uncover behavior al mechanisms in a red fox (Vulpes vulpes) population. Specifically, \( P_i^f(x) = T \) if any flock \( j \neq i \) is at position \( x \) at time \( t \), and \( P_i^f(x) = \min\{P_i^{t-\tau}(x) - \tau, 0\} \) otherwise. Though \( T \) represented the
longevity of scent cues in Potts et al. (2013), here it represents the amount of time a bird will remember a conspecific bird call from a particular location, and so respond to this memory when in that location. The Cinereous Antshrike from each flock tends to make a call about every 2-5 minutes, which can be detected by other birds at a distance of about 50 meters (Karl Mokross, pers. obs.). In the model, we implicitly assume, for simplicity, that birds make calls each time they move and that they are always heard by neighboring flocks.

To test hypothesis (1), we examined whether using the following coupling function

\[
C_i(x, y, \mathcal{P}_i) = \left[ \frac{T_* - \mathcal{P}_i(x)}{T_*} \right]^{\gamma}
\]  

(3)

gives a better fit to the data than the case of no interactions, \(C_i(x, y, \mathcal{P}_i) = 1\). For hypothesis (2), we used the following coupling function

\[
C_i(x, y, \mathcal{P}_i) = V(\kappa I[\mathcal{P}_i(y) > 0], \theta - \theta_g)
\]  

(4)

with \(T_* = \infty\), where \(V(\lambda, \psi)\) is a von Mises distribution (Marsh and Jones 1988), \(I[X]\) is an indicator function taking value 1 if \(X\) is true and 0 otherwise, \(\theta\) is the bearing from \(y\) to \(x\) and \(\theta_g\) is the bearing from \(y\) to the gathering point. For hypothesis (3), we used the coupling function from Eq (4), but with \(T_*\) a finite free parameter, to test whether allowing \(T_*\) to be finite significantly improves the fit.

We fitted the various models to the data using a maximum likelihood technique, whereby we found the free parameters that maximize the product over \(i\) and \(n\) of \(f_{t_i}^{t_{i,n-1}}(x_{i,n+1}|x_{i,n}, \theta_{i,n})\), where \(x_{i,0}, \ldots, x_{i,N_i}\) are the positions of flock \(i\) at times \(t_{i,0}, \ldots, t_{i,N_i}\). To find this maximum, we used the Nelder-Mead simplex algorithm as implemented in the Python maximize() function from the SciPy library (Jones et al. 2001). For hypothesis (1), the free parameters are \(T_*\) and \(\gamma\).
For hypothesis (2), the free parameter is $\kappa$, and for (3) they are $T^*$ and $\kappa$. The p-values for hypothesis testing were obtained using the likelihood ratio test.

One of the strengths of the coupled step selection function approach is that the result of hypothesis testing and/or model selection naturally gives rise to a mechanistic movement model, given by the particular version of equation (1) that corresponds to the best fit model and parameter values. This enables one to determine the space use (i.e. home range) patterns that emerge from the model. We test whether the patterns that emerge from the best model that includes resource selection, topographical selection and territorial interactions are a significantly better fit to the data than the same model without the territorial interactions.

To do this, we constructed a simulation model for the bird flocks, whose movements each step are determined by drawing from the time-dependent probability distribution from Eq. 1 with the best-fit parameter values found by the hypothesis testing technique above. Since each flock gathers in one particular place each day, and moves around the terrain for a total of about eleven-and-a-half hours during the day, we started the simulated birds at the gathering point and ran the simulation for 690 time steps, each step representing $\tau = 1$ minutes (giving 11 hours 30 minutes in total), taking a note of all the positions at which the flock landed after each step. We repeated this 100 times, representing 100 days, giving 69,000 simulated positions for each flock, from which we calculated home ranges using the Kernel Density Estimation (KDE) method. We also ran identical simulations except where the model has $\mathcal{C}(x, y, P_\text{i}^t) = 1$, so that no territorial interactions were included.
Figure 2. Where next? A typical coupled step selection function, giving the probability of an animal’s next move, dependent on territorial marks and resource quality. This is determined both by the strength of territorial marks of conspecifics, given in panel (a), and the quality of the resources (b). The strength of territory marks in this example does not change in the Y-direction, so that animal 1 has territory on the left and animal 2 on the right. The probability of animal 1’s (resp. animal 2’s) next position after some time interval τ, given that it’s current position is in the middle of the landscape (black dot), is shown in panel (c) (resp, panel d). As each animals moves, it marks the terrain causing the territorial profile to change over time, which in turn influences the other animal’s movements. This causes a coupling between their respective step selection functions.

To test which model performed better at predicting space use, we compared the Kullback-Leibler (K-L) distance (Burnham and Anderson 2005) between each model’s KDE distribution and the KDE distribution for the data. The K-L distance differs by a constant from 1/2 times the average Akaike Information Criterion (AIC) of a single sample from the data’s KDE distribution (see Burnham & Anderson (2005) for details). Therefore the difference in AIC (ΔAIC) for two different models of the same data distribution can be thought of as twice the difference in K-L distance, by considering a single KDE distribution as a single data sample. We
have 11 flocks, so 11 KDE distributions. The ΔAIC is twice the sum of the differences in K-L distance across these flocks. We use this value to assess whether the resulting model is better at predicting space use, as opposed to just movement choices, than the model with no territorial interactions. To test whether the models are a good fit to the data, we used a Pearson’s chi-squared test, treating each 10m by 10m square as a single data bin. For this, we used the positional data rather than the smoothed data.

RESULTS

Framing existing models as coupled step selection functions

Step selection and resource selection. Step selection functions are simply single examples of equation (1) with the collective term \( C_i(x, y, P^t_i) \) equal to 1 (Fortin et al. 2005, Forester et al. 2009, Latombe et al. 2013, Vanak et al. 2013). In other words we just consider one animal at a time, and how it interacts with its environment, without attempting to use the results to construct a mechanistic model of interacting animals. Resource selection functions are similar, but the environment-independent movement term \( \phi_i(x|y, \theta_0) \) is replaced with an availability function, which can take whichever form the user feels is appropriate for study (Boyce et al. 2002, Rhodes et al. 2005).

Individual based territory models. The selection of studies by Giuggioli et al. (2009), Giuggioli et al. (2012), Potts et al. (2013) and Potts et al. (2014) modelled territorial interactions using moving agents on a square lattice. The initial model from Giuggioli et al. (2011) has agents performing nearest neighbor random walks and depositing scent as they move. The scent remains for a finite time \( T \), the so-called active scent time, after which it is no longer considered as ‘active’ by conspecifics. Each animal’s movement is restricted by the fact that it cannot move into an area that contains active scent of a neighbor.
This can be framed as a coupled step selection function where \( \varphi_i(x|y, \theta_0) = 1/4 \) if \( x \) is the lattice site either immediately above, below, to the right, or to the left of \( y \), and \( \varphi_i(x|y, \theta_0) = 0 \) otherwise. Additionally, since this model does not include any environmental interactions, we set \( \mathcal{W}_i(x, y, \mathcal{E}) = 1 \). The term \( \mathcal{P}^t_i(x) \) represents the presence of scent at position \( x \) and time \( t \), so

\[
\mathcal{P}^t_i(x) = \begin{cases} T & \text{any animal } j \neq i \text{ is at position } x \text{ at time } t, \\ \min\{\mathcal{P}^{t-\tau}_i(x) - \tau, 0\} & \text{otherwise}. \end{cases}
\] (5)

Then the collective interaction term is

\[
\mathcal{C}_i(x, y, \mathcal{P}^t_i) = \begin{cases} 1 & \text{if } \mathcal{P}^t_i(x) = 0, \\ 0 & \text{otherwise}. \end{cases}
\]

The coupled step selection function formalism (equation 1) gives a natural way of incorporating environmental interactions into such territoriality models, an aspect of this approach hitherto lacking, as noted in Giuggioli et al. (2013).

**Advection-diffusion territory models.** The type of territorial models described in Moorcroft & Lewis (2006) provide several other examples of coupled step selection functions. We describe an individual-level model in a 1D interval \([0, 1]\) that has as its continuum limit the original advection diffusion model of Lewis & Murray (1993). To do this, we first

\[
\phi_i(x|y) = \frac{\exp(-|x - y|/a)}{2a},
\]

where \( a \) is the average step length, and \( \mathcal{W}_i(x, y, \mathcal{E}) = 1 \). This means that the intrinsic movement of each agent (pack of wolves) is a random walk with no correlation, and we are ignoring the effects of the environment on movement. There are two agents in the model, so \( i \in \)
The collective action is mediated by scent deposition so that $P_i^t(x)$ represents the scent mark density of pack $1 - i$. Marking by individual $i$ occurs at a rate $l + mP_{1-i}^t(x)$, where $m$ is typically a monotonic increasing function, representing the tendency of wolves to mark more heavily when conspecific marks are present. $P_i^t(x)$ is governed by the following equation

$$P_i^t(x) = (1 - \mu \tau) P_i^{t-\tau}(x) + \delta(x_{i-1} - x)[l + mP_{1-i}^t(x)] \tau$$

(8)

where $x_i$ is the position of agent $i$ at time $t - \tau$ and $\mu$ is the scent decay rate. Packs have a tendency to move back towards their home range center on encountering foreign scent.

Assuming that the home range center of pack 0 is to the left of the study area and pack 1 to the right, the collective interaction term is given by

$$C_0(x, y, P_0^t) = I(x > y) \tau [D/a - C P_0^{t-\tau}(x)] + I(x \leq y) \tau [D/a + C P_0^{t-\tau}(x)]$$

(9)

$$C_1(x, y, P_1^t) = I(x > y) \tau [D/a + C P_1^{t-\tau}(x)] + I(x \leq y) \tau [D/a - C P_1^{t-\tau}(x)]$$

(10)

where $D$ and $v$ are parameters, which can be determined by model fitting, and $I(X)$ is an indicator function that is equal to 1 if $X$ is true and 0 otherwise.

Now we move from an individual description to positional probability density functions.

Let $u(x, t)$ (resp. $v(x, t)$) be the probability distribution of pack 0 (resp. pack 1). For notational convenience, we rename the scent levels of packs 0 and 1 to $p(x, t)$ and $q(x, t)$ respectively. Then standard theory, e.g. Moorcroft & Lewis 2006 (chapter 2), means that the limit as $\tau \to 0, a \to 0$ of $u(x, t)$ is governed by the following advection-diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2}{\partial x^2}[d_u(x, t)u(x, t)] - \frac{\partial}{\partial x}[c_u(x, t)u(x, t)],$$

(11)
where the advection and diffusion functions \([c_u(x, t)\) and \(d_u(x, t)\) respectively] are the following limits

\[
\begin{align*}
c_u(x, t) &= \lim_{\tau \to 0} \frac{1}{\tau} \int_{-\infty}^{\infty} (y-x)\phi_0(x|y)C_0(x, y, q)dy, \\
d_u(x, t) &= \lim_{\tau \to 0} \frac{1}{\tau} \int_{-\infty}^{\infty} (y-x)^2\phi_0(x|y)C_0(x, y, p)dy.
\end{align*}
\] (12)

This theory is built by constructing the master equation for \(u\). Implicit in the construction is the so-called ‘mean-field’ approximation, which assumes that the covariance between the scent mark density and the position of the pack is (approximately) zero. A direct calculation shows that \(c_u(x, t) = Cq(x, t)\) and \(d_u(x, t) = D\). The equation for \(v(x, t)\) is analogous, but with \(\phi_0, C_0, c_u, d_u,\) and \(q\) replaced by \(\phi_1, C_1, c_v, d_v,\) and \(p\) respectively. Therefore \(c_v(x, t) = -Cp(x, t)\) and \(d_v(x, t) = D\).

The advection diffusion equations for this system of coupled step selection functions are then

\[
\begin{align*}
\frac{\partial u}{\partial t} &= D\frac{\partial^2 u}{\partial x^2} - C\frac{\partial}{\partial x}[qu], \\
\frac{\partial v}{\partial t} &= D\frac{\partial^2 v}{\partial x^2} + C\frac{\partial}{\partial x}[pv].
\end{align*}
\] (13)

Furthermore, the continuous-time limits of the scent marking equations (8) are as follows (Moorcroft 1997, chapter 3)

\[
\begin{align*}
\frac{\partial p}{\partial t} &= u(l + mq) - \mu p, \\
\frac{\partial q}{\partial t} &= v(l + mp) - \mu q.
\end{align*}
\] (14)

Equations (13) and (14) form the system studied in Lewis & Murray (1993). This process can be generalized to derive advection diffusion equations describing territorial pattern formation in two dimensions (Couzin et al. 2002).

**Alignment-and-attraction models.** Equation (1) also reduces to a variety of collective motion models other than territorial ones, including trail-following ants (Deneubourg et al. 1989) and
collective patterns in animal populations through alignment and attraction (Couzin et al. 2002, Eftimie et al. 2007). Here we address one of these modelling frameworks (Couzin et al. 2002) with the others left to the Supplementary Information. To write the model from (Couzin et al. 2002) as a CSSF, we first notice that each animal, $i$, has a fixed speed, $s_i$. Therefore we set $\phi(x|y, \theta_0) = \delta_D(|x - y| - s_i \tau)$, where $\delta_D$ is the Dirac delta function. $W_i(x, y, \mathcal{E}) = 1$ since there are no environmental interactions in the model from Couzin et al. (2002). All the other animals in the population can influence animal $i$’s subsequent movement, so

$$\mathcal{P}_i = (y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n, \theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_n),$$

where $y_j$ is the position of animal $j$ at time $t$, having arrived there on a bearing of $\theta_j$. The model incorporates attraction, alignment and repulsion. Repulsion occurs if there are other animals within distance of $r_r$ from animal $i$, to ensure that animals do not collide. If there is no repulsion then animal $i$ will align with any others that are greater than a distance of $r_r$, but less than a distance of $r_o$, from $i$. They will also be attracted to animals $j$ such that $r_o \leq |y_j - y_i| \leq r_a$ (see Couzin et al. (2002) for details). To aid in writing the interaction term, we let $\theta_r(P^t_i)$ be the repulsion angle, which is the bearing given by the vector

$$v_r = -\sum_{j \neq i} \frac{y_j - y_i}{|y_j - y_i|} I(|y_j - y_i| < r_r). \quad (15)$$

We also define an alignment and attraction angle, $\theta_a(P^t_i)$, which is the bearing given by the direction of

$$v_a = \sum_{j \neq i} \frac{y_j - y_i}{|y_j - y_i|} I(r_o \leq |y_j - y_i| \leq r_a) + \sum_{j \neq i} \left( \frac{\cos(\theta_j)}{\sin(\theta_j)} \right) I(r_r \leq |y_j - y_i| < r_o). \quad (16)$$
The interaction term from Couzin et al. (2002), section ‘Behavioural rules: description’, is then

\[
C_i(x, y, \mathcal{P}_i) = \begin{cases} 
\text{SG}(\theta - \theta_r) & \text{if there is a } j \neq i \text{ such that } |y_j - y_i| < r_r, \\
\text{SG}(\theta - \theta_a) & \text{if there is a } j \neq i \text{ such that } |y_j - y_i| \leq r_a \\
& \text{but no } k \neq i \text{ such that } |y_k - y_i| < r_r, \\
\text{SG}(\theta - \theta_0) & \text{otherwise,}
\end{cases}
\] (17)

where \(\text{SG}(\psi)\) is a spherical Gaussian.

**The example of Amazonian bird flocks**

When we apply our technique to data on Amazonian birds, there is no significant improvement in fit \((p = 0.60)\) if we model birds as having a tendency not to go into areas from where they have heard conspecific bird calls in the past (hypothesis 1 from the Methods section). However, when flocks are modelled as being allowed to move into neighboring territories, but then having a tendency to retreat in the direction of the gathering point (hypothesis 2), we observe a significant improvement in fit \((p = 0.022)\). If we assume that the territorial cues have a finite lifetime (hypothesis 3), the fit becomes worse, suggesting that birds are able to remember these cues for a very long time after they have been made.

To demonstrate the space use patterns that arise from these results, we constructed simulations using the gathering point attraction model, used to test hypothesis 2, with the best fit parameters of \(T^* = \infty\) and \(\kappa = 0.0597\) (Fig. 3). For 9 of the 11 flocks, the resulting Kernel Density Estimator (KDE) distributions are closer to those of the data than the KDE distributions without territorial interactions (see Table 2). Furthermore, the resulting difference in Akaike Information Criteria (\(\Delta AIC\)) between the two models is \(\Delta AIC = 4.07\), giving reasonable evidence to suggest that the model including territorial interactions is better at predicting space use patterns than that
without. This is demonstrated pictorially in Fig. 3b, which shows that the model including territorial interactions is more highly peaked at the center and includes a lower density of outliers.

Table 2. Fitting models both with and without territorial interactions to data on bird flock movement. For each flock, the Kullback-Leibler (K-L) distance between the data’s Kernel Density Estimator (KDE) distribution and the model’s KDE distribution is given. For all but two of the flocks, the model that includes territorial interactions performs best, shown by a positive difference in column 4.

<table>
<thead>
<tr>
<th>Flock</th>
<th>K-L with interactions</th>
<th>K-L no interactions</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central</td>
<td>0.868</td>
<td>1.236</td>
<td>0.367</td>
</tr>
<tr>
<td>North</td>
<td>1.018</td>
<td>1.442</td>
<td>0.424</td>
</tr>
<tr>
<td>South Central</td>
<td>0.673</td>
<td>0.826</td>
<td>0.152</td>
</tr>
<tr>
<td>South West</td>
<td>1.020</td>
<td>1.317</td>
<td>0.297</td>
</tr>
<tr>
<td>Lake</td>
<td>0.902</td>
<td>1.063</td>
<td>0.161</td>
</tr>
<tr>
<td>W400</td>
<td>0.737</td>
<td>0.989</td>
<td>0.252</td>
</tr>
<tr>
<td>Cap II</td>
<td>3.527</td>
<td>3.377</td>
<td>-0.150</td>
</tr>
<tr>
<td>Cap South</td>
<td>1.192</td>
<td>1.465</td>
<td>0.305</td>
</tr>
<tr>
<td>Ig-Cmp</td>
<td>0.779</td>
<td>1.013</td>
<td>0.234</td>
</tr>
<tr>
<td>Cap North</td>
<td>1.125</td>
<td>1.048</td>
<td>-0.077</td>
</tr>
<tr>
<td>North-East</td>
<td>0.967</td>
<td>1.038</td>
<td>0.071</td>
</tr>
</tbody>
</table>

Of the two flocks that are not well-modelled by incorporating territorial interactions, for Cap North we have no data on adjacent flocks (Fig. 3a) so the inability of the model to detect territorial interactions is unsurprising. Cap II, on the other hand, is located in the most degraded area of all flocks in the study. Subsequent observations of the study area suggest that it did not persist over time, as key species either abandoned the area or died. Therefore the territory could well be in the process of moving or degrading during the study period, mechanisms that are likely to be key drivers in shaping the space use, but which are absent from our current model.

For all of the flocks except Cap II, there was insufficient evidence to suggest that the data did not come from the model distribution that included territorial interactions ($p < 0.0001$ for
Cap II, $p > 0.999$ for the others). The same test with the model that excluded territorial interactions suggested that there was only sufficient evidence to reject the hypothesis that the data came from the model for Cap II and Central ($p < 0.0001$ for Cap II and Central, $p > 0.999$ for the others). Therefore we have significantly improved the absolute fit of the Central data by including territorial interactions. Central is the only flock for which we have data on all surrounding flocks so it is precisely the flock for which one would most expect to see improvement of absolute fit.

Figure 3. Space use predictions of bird locks using a coupled step selection function. A. Dots represent recorded positions of bird flocks, whereas the contours detail the space use distributions that arise from a territorial and environmental-interaction model that best fits the movement data (see Methods for details). The colors of the contours for each flock correspond to those of both the positional data points and the text giving the flock names. B. shows the predicted position distributions for the Central flock with territorial interactions minus those without such interactions. Note that no fitting was performed between the model spatial distribution and the bird positions. Instead, the distributions simply emerge from the model’s underlying movement and interaction processes.

**DISCUSSION**

We have constructed a general model for the effects on movement of both animal-habitat and between-animal interactions. We have demonstrated how the model encompasses, as special cases, a variety of disparate collective motion models as well as resource and step selection functions. By fitting a version of our model to data on bird flock locations, we have shown how
it can be used to determine and quantify the nature of territorial interactions, as well as modelling simultaneously the effects of both conspecífics and the environment on movement processes. Since we framed the system as a one-step Markovian model of both the animals and their environment, our framework allows for relatively simple calibration of models, which makes the process computationally fast. This contrasts with methods that fit the movement path as a whole, such as state-space models, which can be difficult to fit (Patterson et al. 2008).

Though we have focused on territorial modelling, so not given an exhaustive demonstration of how our framework might be reducible to all collective behavior models in the literature, we display a variety of different examples, encompassing both direct and mediated interactions, both conspecífic attraction and avoidance processes. These demonstrate the possible wide applicability of our approach, and potential to frame many more models as coupled step selection functions. Encompassing competing models of collective behavior under this unifying frame-work will make future comparisons easier, aided by the methods given here for fitting coupled step selection functions to data. Furthermore, it will enables transference of techniques and results between the hitherto disparate fields of collective motion, resource selection and mechanistic territorial modelling. To give one example, research into ungulate behavior often looks at the effects of the environment on movement but ignores herding interactions (e.g. Fortin et al. 2005), or looks at herding behavior but ignores the resource aspect (e.g. Jullien and Thiollay 1998). Our framework links these two ideas so will help future researchers build and validate models that account for both.

By applying our model to movement patterns of bird flocks, we were able to test hypotheses about the mechanisms behind the interaction processes. Previous studies of mechanisms underlying territorial patterns in populations of scent-marking animals postulated
that they will avoid areas that have recently been claimed by others as their territory (Potts et al. 2013). Here we have shown that the territorial interaction mechanism in bird flocks is quite different. There is no evidence to suggest that they tend to avoid places that have previously been claimed as other flocks’ territories. However, after visiting the outskirts of neighboring territories, they will change their movement processes to include a tendency to retreat back inside their territory. These visitations explain the observed slightly overlapping utilization distributions in the birds’ spatial patterns (see Fig. 3).

Our framework can also be used to build predictive, mechanistic models showing how utilization distributions arise from the underlying movement and interaction processes. To demonstrate this, we used stochastic simulations of the best fit system of coupled step selection functions for the bird data. Recently, step selection functions have been used to construct deterministic master equation (Potts et al. 2014) and partial differential equation models (Moorcroft and Barnett 2008), from which the resulting spatial distributions can be analyzed using well-studied mathematical tools, e.g. Moorcroft & Lewis (2006). Whilst the coupling term in our framework makes such analysis significantly more complicated than for ordinary step selection functions, deterministic mathematical formulations would ultimately enable concrete conclusions to be reached without the need for extensive, time-consuming computer simulations. We therefore hope, in future work, to begin a program of analyzing coupled step selection models mathematically.

Though mechanistic models have previously been proposed to explain space use patterns by examining both movement, territorial interactions and environmental features (Moorcroft et al. 2006), those models fit the emergent space use distribution to relocation data, whereas our model is directly fitted to the movement trajectory itself, enabling the space use distribution to
arise with no additional fitting. The advantage of this is twofold. First, there is no need to throw away data in order to make sure each data point is an independent sample of the spatial distribution from the others (see Moorcroft & Lewis (2006) for details of, and rationale behind, this procedure). Therefore we can use the complete movement trajectory, containing much more information.

Second, fitting the model to the underlying movement choices ensures that the parameter values used to construct the model arise from the movement and interaction processes rather than the emergent patterns. This means that we can assess to what extent these processes predict space use, and where they fail. For example, in the data studied here, the space use of two flocks (Cap II and Cap North) were not predicted by the territorial interaction model as well as by the no-interaction model, unlike the other nine flocks (Table 2). Therefore we can postulate hypotheses about what other processes may be required to predict space use in these instances. On the other hand, fitting directly to the space use distribution implicitly assumes that the mechanistic model describes well all aspects of movement that give rise to the spatial patterns. Consequently, this procedure may cause inaccurate inferences to be made about the parameter values of the underlying processes. In other words, our approach is more cautious, therefore less likely to lead to incorrect results and more likely to reveal the extent to which certain processes fail to predict accurately the spatial patterns.

As an alternative to mathematical models of space use, simulations of individual based models have also been used to attempt to understand animal movement decisions and emergent spatial patterns (Hartig et al. 2011). Typically, they take a pattern-oriented approach (Grimm et al. 1996, Grimm et al. 2005), beginning by including as many aspects of the animal’s movement and interaction processes as are believed to cause the observed patterns. If the empirical patterns,
also called summary statistics, are observed in the model output then the model is simplified to try to understand exactly which of the processes are causing the patterns to emerge. The aim of this approach is to find models that replicate as many of the summary statistics observed in the data as possible, with as few model parameters.

Our approach, on the other hand, is process-based in nature (Evans et al. 2013), seeking to build an individual based mechanistic model by testing hypotheses about the underlying processes one at a time. The key difference is that we test the model parameters against the data for validity on the same level of description at which the model is constructed. The pattern-oriented approach tests the model parameters at a different level of description: that of the summary statistics. However, this is not sufficient for making inferences about the parameter values put into the model. Though analysis of a mechanistic model, individual based or analytic, shows that process A implies pattern B, showing that pattern B replicates the data does not imply that the underlying mechanism is actually process A. Therefore it is not possible, purely using a pattern-oriented approach, to make solidly-grounded inferences about the nature of the mechanisms that have gone into construction of the model. In our approach, we circumvent this issue by testing and parameterizing the model’s mechanisms on the level of description at which they are constructed, then observing the patterns as an emergent feature of the model, which can in turn be compared with the patterns from data.

Recent developments in the collective behavior literature provide many good examples of process-based modelling and model parameterization (Nagy et al. 2010, Herbert-Read et al. 2011, Kat et al. 2011, Pérez-Escudero and de Polavieja 2011, Gautrais et al. 2012, Petit et al. 2013, Strandburg-Peshkin et al. 2013, Mann et al. 2014). However, very few examine the
emergent features of these data-parameterized models and test whether they accurately replicate the population level patterns seen in the data, as we do here.

That said, there are exceptions, e.g. Bode et al. (2010, 2011), Pettit et al. (2013), and these models could, in principle, be used in conjunction with theoretical mechanistic models of pattern formation, such as Goldstone & Janssen (2005), Eftimie et al. (2007), to provide a full story. If they were to be framed under a single overarching methodological framework, such as the coupled step selection functions proposed here, then this would aid sort of the unification of process-based model construction and theoretical process-to-pattern analysis that has recently been sought after (Sumpter et al. 2012).

Table 2. Fitting models both with and without territorial interactions to data on bird flock movement. For each flock, the Kullbeck-Leibler (K-L) distance between the data’s Kernel Density Estimator (KDE) distribution and the model’s KDE distribution is given. For all but two of the flocks, the model that includes territorial interactions performs best, shown by a positive difference in column 4.

<table>
<thead>
<tr>
<th>Flock</th>
<th>K-L with interactions</th>
<th>K-L no interactions</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central</td>
<td>0.868</td>
<td>1.236</td>
<td>0.367</td>
</tr>
<tr>
<td>North</td>
<td>1.018</td>
<td>1.442</td>
<td>0.424</td>
</tr>
<tr>
<td>South Central</td>
<td>0.673</td>
<td>0.826</td>
<td>0.152</td>
</tr>
<tr>
<td>South West</td>
<td>1.020</td>
<td>1.317</td>
<td>0.297</td>
</tr>
<tr>
<td>Lake</td>
<td>0.902</td>
<td>1.063</td>
<td>0.161</td>
</tr>
<tr>
<td>W400</td>
<td>0.737</td>
<td>0.989</td>
<td>0.252</td>
</tr>
<tr>
<td>Cap II</td>
<td>3.527</td>
<td>3.377</td>
<td>-0.150</td>
</tr>
<tr>
<td>Cap South</td>
<td>1.192</td>
<td>1.465</td>
<td>0.305</td>
</tr>
<tr>
<td>Ig-Cmp</td>
<td>0.779</td>
<td>1.013</td>
<td>0.234</td>
</tr>
<tr>
<td>Cap North</td>
<td>1.125</td>
<td>1.048</td>
<td>-0.077</td>
</tr>
<tr>
<td>North-East</td>
<td>0.967</td>
<td>1.038</td>
<td>0.071</td>
</tr>
</tbody>
</table>

Though our model was significantly better at predicting space use than the model free of territorial interactions, it is clear from Fig. 3a that our model does not capture all aspects of the birds’ spatial patterns. However, the strength of our approach is that we can readily add further
behavioral features one at a time, testing the efficacy of each using the techniques detailed here. For example, the birds are known to have direct territorial conflicts, which affect where they move in subsequent days and weeks. Also, the movement is driven by intra-flock interactions, with one particular species, the Cinereous Antshrike (Thamnomanes caesius), playing the main role in maintaining cohesiveness. By using our techniques to test the effect of such behavioral phenomena on movement and space use, we can move towards building truly accurate, predictive models linking movement processes, conspecific interactions and collective behavior, to the emergent space use distributions.

LITERATURE CITED


Levin, S.A. 2012. Towards the marriage of theory and data. Interface Focus, 2: 141-143


## APPENDIX C: SUPPLEMENTARY MATERIAL FOR CHAPTER 4

Table S1. Summary of flocks sampled in the study. Shown are flock ID; habitat type; canopy height measures (maximum, standard deviation and mean); the number of hours each flock was sampled; flock attendance based on sum of normalized presence of all species in a flock; and weighted clustering coefficients.

<table>
<thead>
<tr>
<th>Flock ID/Network</th>
<th>Canopy</th>
<th>Max. ht. (m)</th>
<th>Mean ht. (m)</th>
<th>Hours Sampled</th>
<th>Spp. a</th>
<th>Interaction b</th>
<th>Attendance c</th>
<th>Clustering Coefficient d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central</td>
<td>100-Ha</td>
<td>45.0</td>
<td>24.6</td>
<td>38.5</td>
<td>46</td>
<td>514</td>
<td>929.9</td>
<td>0.79</td>
</tr>
<tr>
<td>Lake Flk</td>
<td>100-Ha</td>
<td>42.6</td>
<td>17.8</td>
<td>18.0</td>
<td>40</td>
<td>450</td>
<td>822.2</td>
<td>0.81</td>
</tr>
<tr>
<td>South Central</td>
<td>100-Ha</td>
<td>49.8</td>
<td>23.7</td>
<td>21.0</td>
<td>47</td>
<td>686</td>
<td>1145.2</td>
<td>0.83</td>
</tr>
<tr>
<td>Southwest</td>
<td>100-Ha</td>
<td>45.5</td>
<td>20.9</td>
<td>25.0</td>
<td>53</td>
<td>671</td>
<td>878.4</td>
<td>0.75</td>
</tr>
<tr>
<td>W400</td>
<td>100-Ha</td>
<td>48.4</td>
<td>20.1</td>
<td>49.0</td>
<td>69</td>
<td>952</td>
<td>814.3</td>
<td>0.69</td>
</tr>
<tr>
<td>Col 10ha</td>
<td>10-Ha</td>
<td>50.6</td>
<td>14.6</td>
<td>51.0</td>
<td>78</td>
<td>860</td>
<td>695.1</td>
<td>0.60</td>
</tr>
<tr>
<td>Dim 10ha</td>
<td>10-Ha</td>
<td>39.4</td>
<td>18.2</td>
<td>83.5</td>
<td>83</td>
<td>994</td>
<td>739.5</td>
<td>0.69</td>
</tr>
<tr>
<td>Cab Frio</td>
<td>PF</td>
<td>46.1</td>
<td>22.3</td>
<td>26.5</td>
<td>50</td>
<td>600</td>
<td>950.9</td>
<td>0.76</td>
</tr>
<tr>
<td>FL IB</td>
<td>PF</td>
<td>48.2</td>
<td>23.0</td>
<td>23.0</td>
<td>51</td>
<td>493</td>
<td>793.5</td>
<td>0.65</td>
</tr>
<tr>
<td>FL II</td>
<td>PF</td>
<td>40.6</td>
<td>23.7</td>
<td>42.5</td>
<td>67</td>
<td>923</td>
<td>841.9</td>
<td>0.74</td>
</tr>
<tr>
<td>FL III</td>
<td>PF</td>
<td>43.5</td>
<td>23.3</td>
<td>26.0</td>
<td>60</td>
<td>844</td>
<td>942.3</td>
<td>0.73</td>
</tr>
<tr>
<td>Gav 10</td>
<td>PF</td>
<td>42.5</td>
<td>23.3</td>
<td>20.5</td>
<td>55</td>
<td>728</td>
<td>1134.1</td>
<td>0.79</td>
</tr>
<tr>
<td>Gav I</td>
<td>PF</td>
<td>49.4</td>
<td>24.7</td>
<td>27.0</td>
<td>51</td>
<td>679</td>
<td>872.2</td>
<td>0.81</td>
</tr>
<tr>
<td>Gav II</td>
<td>PF</td>
<td>43.5</td>
<td>23.9</td>
<td>18.0</td>
<td>45</td>
<td>412</td>
<td>591.7</td>
<td>0.68</td>
</tr>
<tr>
<td>Km37 III</td>
<td>PF</td>
<td>55.2</td>
<td>25.9</td>
<td>20.0</td>
<td>45</td>
<td>451</td>
<td>719.5</td>
<td>0.71</td>
</tr>
<tr>
<td>Km37 V</td>
<td>PF</td>
<td>42.2</td>
<td>24.4</td>
<td>17.0</td>
<td>48</td>
<td>491</td>
<td>820.6</td>
<td>0.70</td>
</tr>
<tr>
<td>FL IA</td>
<td>PSF</td>
<td>43.0</td>
<td>15.9</td>
<td>24.0</td>
<td>52</td>
<td>503</td>
<td>777.1</td>
<td>0.74</td>
</tr>
<tr>
<td>Ig Camp</td>
<td>PSF</td>
<td>37.8</td>
<td>17.7</td>
<td>62.0</td>
<td>66</td>
<td>773</td>
<td>816.9</td>
<td>0.65</td>
</tr>
<tr>
<td>Cap 1ha</td>
<td>SG</td>
<td>43.5</td>
<td>10.8</td>
<td>38.0</td>
<td>54</td>
<td>320</td>
<td>436.8</td>
<td>0.50</td>
</tr>
<tr>
<td>Cap II</td>
<td>SG</td>
<td>41.0</td>
<td>14.7</td>
<td>28.0</td>
<td>49</td>
<td>272</td>
<td>439.3</td>
<td>0.51</td>
</tr>
<tr>
<td>Cap N</td>
<td>SG</td>
<td>32.9</td>
<td>17.6</td>
<td>22.0</td>
<td>43</td>
<td>414</td>
<td>802.3</td>
<td>0.73</td>
</tr>
</tbody>
</table>

a) Represents the number of nodes in the network  
b) Represents the number of edges in the network  
c) Sum of all species encounter rates/hr  
d) Weighted value calculated as a global network statistic
Table S2. Results of the model for habitat effects on Amazonian mixed-species flock species richness and encounter rate, and comparison with null models.

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>Log Likelihood</th>
<th>L-R test</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Species Richness</em> ~ Habitat</td>
<td>49.2</td>
<td>-18.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>57.1</td>
<td>-26.5</td>
<td>15.85</td>
<td>0.003</td>
</tr>
<tr>
<td><em>Encounter Rate</em> ~ Habitat</td>
<td>N/A</td>
<td>-6116.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>N/A</td>
<td>-6122.0</td>
<td>11.92</td>
<td>0.017</td>
</tr>
</tbody>
</table>
Table S3. Results of the models for habitat effects on Amazonian mixed-species flock network properties, and comparison with null models.

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>logLik</th>
<th>L-R test</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized degree ~ Habitat</td>
<td>9209.9</td>
<td>-4599.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>9224.2</td>
<td>-4610.1</td>
<td>22.3</td>
<td>0.0001</td>
</tr>
<tr>
<td>Weighted degree ~ Habitat</td>
<td>715.8</td>
<td>-351.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>723.5</td>
<td>-359.8</td>
<td>15.65</td>
<td>0.0035</td>
</tr>
</tbody>
</table>

a) Both models included flock id as a random effect and model offsets for number of nodes and/or sampling time
Table S4. Differences in the degree distributions of observed and random networks of interspecific interactions sampled in five different habitats are evidenced by median degree, skew and average path length.

<table>
<thead>
<tr>
<th>Network</th>
<th>Median Degree</th>
<th>Quartiles (25-75%)</th>
<th>STD</th>
<th>Skew</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-Ha</td>
<td>0.207</td>
<td>0.12-0.40</td>
<td>0.22</td>
<td>1.110</td>
</tr>
<tr>
<td>10-Ha Random</td>
<td>0.146</td>
<td>0.12-0.17</td>
<td>0.04</td>
<td>0.146</td>
</tr>
<tr>
<td>100-Ha</td>
<td>0.485</td>
<td>0.31-0.72</td>
<td>0.26</td>
<td>0.225</td>
</tr>
<tr>
<td>100-Ha Random</td>
<td>0.250</td>
<td>0.20-0.30</td>
<td>0.07</td>
<td>0.124</td>
</tr>
<tr>
<td>PF</td>
<td>0.423</td>
<td>0.24-0.65</td>
<td>0.25</td>
<td>0.295</td>
</tr>
<tr>
<td>PF Random</td>
<td>0.224</td>
<td>0.18-0.27</td>
<td>0.06</td>
<td>0.167</td>
</tr>
<tr>
<td>PSF</td>
<td>0.274</td>
<td>0.18-0.53</td>
<td>0.25</td>
<td>0.896</td>
</tr>
<tr>
<td>PSF Random</td>
<td>0.184</td>
<td>0.15-0.22</td>
<td>0.05</td>
<td>-0.035</td>
</tr>
<tr>
<td>SG</td>
<td>0.214</td>
<td>0.11-0.42</td>
<td>0.23</td>
<td>1.218</td>
</tr>
<tr>
<td>SG Random</td>
<td>0.125</td>
<td>0.09-0.19</td>
<td>0.07</td>
<td>0.915</td>
</tr>
</tbody>
</table>
Figure S1. Sample (time blocks) based rarefaction curves for each flock sampled in primary forest (blue), 100-Ha fragments (red), second-growth (orange), 10-Ha fragments (yellow) and primary-secondary forests (green).
APPENDIX E: PERMISSION FROM THE ROYAL SOCIETY PUBLISHING TO PUBLISH APPENDIX B

Permission Request

This reuse request is free of charge and you are not required to obtain a license.

Jonathan Potts

Hi Karl,

Things are good, thanks. Congratulations for both your new son and your dissertation - that's a big couple of weeks!

I am very happy with your suggested arrangement regarding dissertation chapters. Let me know how it all goes putting this together - I'd like to read the resulting piece.

Jonathan
VITA

Karl Mokross was born in São Carlos, state of São Paulo, Brazil in 1977. He graduated in Biological Sciences at Universidade Federal de São Carlos (UFSCar) in 2001 and went on to complete a Master’s degree in Tropical Ecology at the Instituto Nacional de Pesquisas da Amazônia (INPA) in 2002. From 2004 to 2008 he worked at Adolpho Ducke Reserve in Manaus filming frog species; as a Caspian tern colony monitor in Oregon; and as wildlife technician at the Kaheawa wind farm in Maui, Hawaii. In 2009 he moved to Baton Rouge to begin a doctoral program in Philip Stouffer’s lab. He currently resides in Rio Claro (SP). He will receive his Doctoral degree in Wildlife from Louisiana State University in December 2014.