



Analysing small-scale aggregation in animal visits in space and time: the ST-BBD method

Tyler R. Bonnell^{a,*}, Pierre Dutilleul^b, Colin A. Chapman^{c,d}, Rafael Reyna-Hurtado^e, Raul Uriel Hernández-Sarabia^f, Raja Sengupta^a

^a Department of Geography, McGill University, Montreal, QC, Canada

^b Department of Plant Science, McGill University, Montreal, QC, Canada

^c Department of Anthropology & McGill School of Environment, McGill University, Montreal, QC, Canada

^d Wildlife Conservation Society, Bronx, NY, U.S.A.

^e EL Colegio de la Frontera Sur ECOSUR, Unidad Campeche, Campeche, Mexico

^f Universidad Veracruzana, Departamento de Neuroetología, Xalapa, Veracruz, Mexico

ARTICLE INFO

Article history:

Received 13 July 2012

Initial acceptance 19 September 2012

Final acceptance 28 November 2012

Available online 9 January 2013

MS. number: A12-00542R

Keywords:

agent-based model

animal movement

beta-binomial distribution

folivore

frugivore

habitat use

monkey

small-scale aggregation

spatiotemporal point pattern analysis

Movement behaviour plays an important role in many ecological interactions. As animals move through the environment, they generate movement patterns, which are a combined result of landscape characteristics and species-specific behaviour. Measuring these ranging patterns is being facilitated by technological advances in collection methods, such as GPS collars, that are capturing movement on finer spatial and temporal scales. We propose the use of a novel spatiotemporal analytical framework (ST-BBD), based on the beta-binomial distribution (BBD) model, to measure small-scale aggregation in animal movement data sets, including two simulated and three collected primate data sets. We use this approach to distinguish different habitat uses of three primate species (red colobus, *Procolobus rufomitratus*, black howler, *Alouatta pigra*, and spider monkey, *Ateles geoffroyi*) and quantify their specific use of the landscape in space and in time, using a parameter of the BBD that measures the variation in sites visited on a landscape. We found that estimates of aggregation in habitat use were higher in the frugivorous spider monkey, compared to the more folivorous howler monkey, and that in the red colobus, aggregation in site visits was dependent on group size and food availability. Applications of this framework to animal movement data could be useful in understanding ecological systems where habitat use is an important factor, such as the relationships between hosts and parasites, or parent plants and seed dispersers.

© 2012 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

Movement is a vital process for many animals, affecting a broad spectrum of ecological processes and patterns observed in nature (Nathan et al. 2008). The study of animal movement behaviour has led to much insight into such fields as foraging ecology (Orians & Pearson 1979; Humphries et al. 2010), wildlife management (Chetkiewicz et al. 2006), plant seed dispersal (Vellend et al. 2003) and disease dynamics (Bartel et al. 2011).

With technological advances providing the increasing availability of high-resolution animal movement data, both in time and in space, there is a need to advance methodologies for analysing movement data. Movement is a spatiotemporal process in which an animal moves over a heterogeneous landscape through time. Thus, states and

conditions of both the animal in question and the landscape are continuously changing. Different methods have been developed to measure movement characteristics, each capturing some aspect of movement. These different measurements can be separated depending on whether the focus is on space, time, or space and time.

One main branch of spatial approaches focuses on point pattern analysis, where movement data are treated as (x, y) locations measured in a two-dimensional space (e.g. home range; Millsaugh et al. 2004). These methods search for patterns, or structure, in the distribution of points, which can provide insight into the behavioural ecology of the study species. Another, more complete approach is followed when a temporal component is added to points, tying them together by the sequence in which they were measured. This approach is characterized by the measurement of 'between-step characteristics' (e.g. step length or angle) at similar time periods. In analysing between-step measurements, the results are often compared to theoretical distributions as a means of interpretation, such as the Lévy walk, random

* Correspondence: T. R. Bonnell, Department of Geography, McGill University, 805 Sherbrooke Street West, Montreal, QC H3A 0B9, Canada.

E-mail address: tyler.bonnell@mail.mcgill.ca (T. R. Bonnell).

walk or correlated-random walk (Viswanathan et al. 1996; Edwards et al. 2007; Humphries et al. 2010). Methods focusing more specifically on the temporal aspects of the data have made use of time series analysis, in which between-step measures (e.g. distance or turning angle) are examined through time to test for autocorrelation in movement behaviour (Dray et al. 2010). Kernel methods have also been applied, classically to (x, y) locations in two-dimensional space, to derive the probability of occurrence on a landscape, and has recently been extended to estimate probability of occurrence in space and time (Keating & Cherry 2009).

Using animal movement data, we were interested in developing a measure to characterize an animal's use of habitats; specifically, to measure the spatiotemporal variation in use of sites within a habitat. In examining movement data, or following an animal in the field, it is often apparent that some sites are used more intensely and more often than others. It is possible that there exists a cyclical pattern of site use, one based on seasons or depletion–replenishment of resources, so that the animal in question may or may not follow along similar routes on the landscape. The relevant method should allow the estimation of variation in these patterns of habitat use in space and time, and provide a statistically sound way to test for differences, for example, among comparable species. Existing approaches to examine spatiotemporal variation (e.g. Mantel test and correlograms; see below, *Variability and Autocorrelation in Animal Visits of a Landscape*) do not capture and compare variation between specific areas and periods, and as a whole are more global than the small-scale focus desired here. Therefore, we evaluated a new approach that uses the beta-binomial distribution (BBD) in a framework where a spatiotemporal grid is created to measure habitat use. Basically, the space and time in which the animal is moving are broken into space–time cubes in which either presence or absence of passage (i.e. Bernoulli trial) is recorded. Space–time cubes have been introduced as sampling units in the field of time geography, and have been used to map spatiotemporal data for subsequent analysis (Miller 2005). Subdividing the space–time cubes into cells within sampling units following the BBD opens the door to the quantification of spatiotemporal aggregation at smaller scales.

The BBD was applied by Hughes & Madden (1993) in an epidemiological context for investigating the aggregation of diseased plants on agricultural landscapes. These authors showed that when there is uneven chance of finding an infected plant through the landscape (i.e. it is more likely to find an infected plant near another infected plant), the BBD can be useful in quantifying the spatial aggregation of infection occurrences. In a spatiotemporal context, the BBD has been suggested for use with animal movement data (Dutilleul 2011, Chapter 4). In this approach, Dutilleul (2011) proposes an extension on the framework of Madden & Hughes (1995) that consists in measuring the displacement of a spatial unit (i.e. point) through time and quantifying the aggregation of visited areas in space and time.

The objective of our paper is to evaluate the spatiotemporal version of the BBD framework (ST-BBD hereafter), as a tool to estimate the intensity to which an animal uses its home range. We first apply the ST-BBD framework to primate movement data simulated under two behaviour models. We then apply it to movement data collected for three primate species in the field, and quantify the intensity of habitat use within each species and compare it between species. The three species chosen vary in the degrees of frugivory and folivory: spider monkeys, *Ateles geoffroyi*, rely heavily on fruit resources (frugivorous), whereas howler monkeys, *Alouatta pigra*, and red colobus, *Procolobus rufomitratus*, typically rely on leaves (folivorous). Foraging for fruits in a tropical forest requires finding the few dispersed trees that produce fruit intermittently. By comparison, when foraging for good-quality leaves, trees are thought to be less dispersed, offering food more often than do fruit trees. Given the varying distribution of resources, both spatially and

temporally, we used the ST-BBD framework to quantify spatiotemporal aggregation accordingly and determine differences in habitat use by these two folivorous and one frugivorous primate species.

Data Sets

Simulated data

We first simulated movement data with an agent-based model (ABM) of primate group foraging, which approximates primate group movement behaviour for use in a spatially explicit epidemiological study (Bonnell et al. 2010). In general, ABMs define characteristics and behaviours of individual agents (e.g. primates, fish, farmers, companies) within a simulation environment, and allow them to interact to create system-level outcomes (for use of ABM in behavioural studies, see: Hemelrijk 2002; Bryson et al. 2007). The ABM here was constructed as a general model, made to fit a wide range of primate group foraging behaviours. In this study, we focus on the effect that a foraging trait called 'weight of remembered sites' has on overall movement behaviour. This trait is a key model component that affects each primate's foraging decision making. In the model, the primate agents use a food site index to assess which site, seen or remembered, is the best site to move towards. This food site index is based on the expected amount of food and the distance to such sites, while the parameter 'weight of remembered sites' gives extra weight to sites that are familiar (i.e. remembered) to the primate agent (equation 1).

The food site index value at site (x', y') from site (x, y) is given by

$$I(x', y') = \frac{D((x', y'), (x, y))}{F(x', y') * w} \quad (1)$$

where $D(\cdot)$ represents the Euclidean distance between two points, $F(\cdot)$ is the primate's evaluation of the amount of food at a given site, and w is the weight parameter applied to remembered sites.

When preference for remembered sites is increased, simulated primates tend to use selected sites intensively, visiting familiar sites often, and travelling along similar routes connecting these sites. We will thus refer to this behaviour model as the 'routing model', because it creates travel routes. On the other hand, with low preference for remembered sites, groups show nonspecific ranging behaviours and no heavy use of a specific area. Accordingly, we will refer to that behaviour model as the 'nonrouting model'.

For each type of model, one simulated group of primates (group size = 72) was made to forage on a 1.5×1.5 km landscape for 6 months. Movement data were recorded for 5 months, after the first month was discounted to sufficiently initialize the memory of individuals in the simulated group. The position of the simulated group was taken every 30 min during the active hours of the day (0700–2000 hours), and recorded as a point (x, y) with a time component (t) . The 30 min interval was selected because it is very feasible logistically speaking and is therefore often used in field studies (Chapman et al. 2002), and it is a crude time period to avoid autocorrelation given the distances animals can cover in half an hour (Reyna-Hurtado et al. 2009). With the chosen total duration (5 months) and sampling time interval (30 min), the final set of simulated data was composed of two point patterns, each with 3900 point locations in space–time (Fig. 1).

Collected data

Our second data source consists of GPS points or tracks taken from field measurements of primate group movement. We make use of movement data collected from a highly folivorous primate (red colobus), a folivore–frugivore primate (black howler), and a highly frugivorous primate (spider monkey). Red colobus data were collected from August 2006 to June 2010 in Kibale National

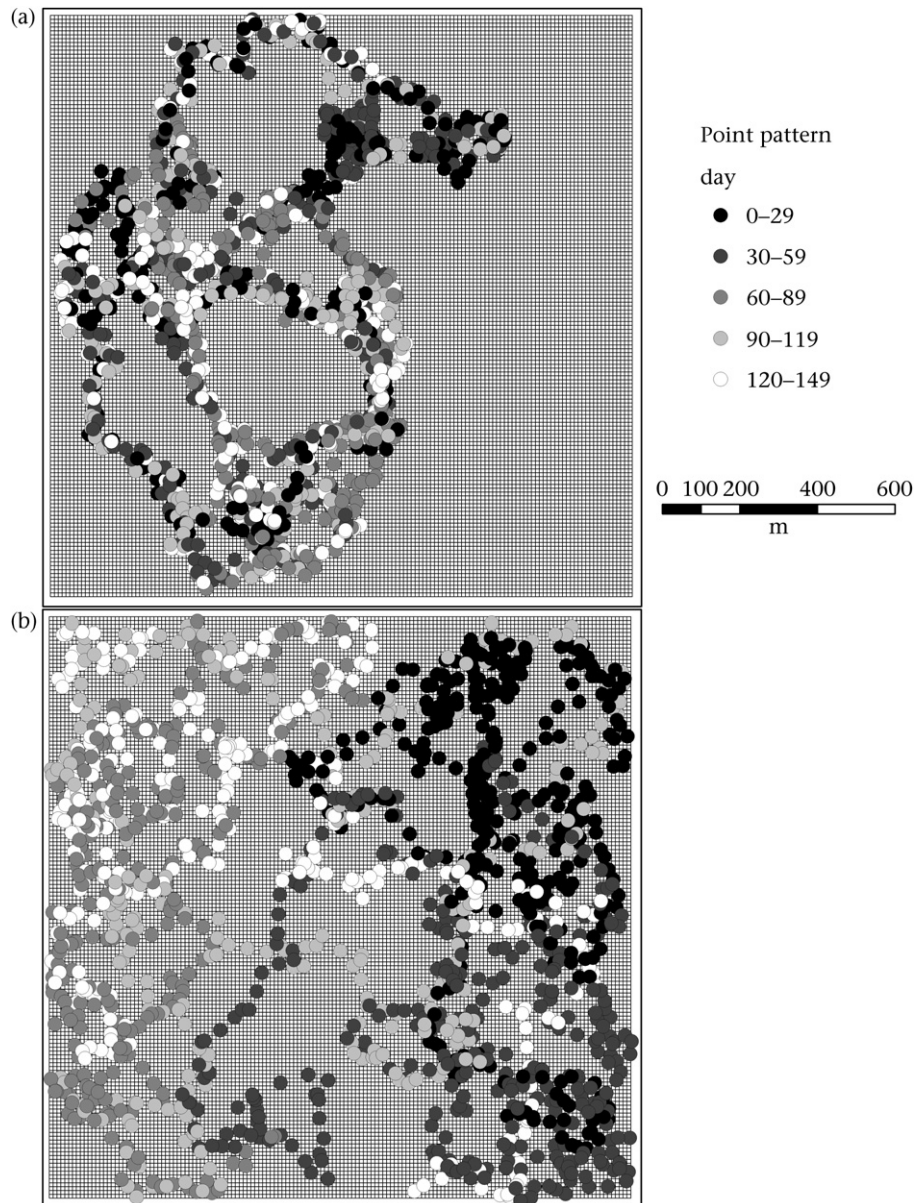


Figure 1. Point patterns generated by simulating primate foraging behaviour: (a) primate agents with routing behaviour and (b) primate agents without routing behaviour. Each simulation, run for 5 months and collecting location every 30 min, created a point pattern of 3900 point locations in space–time.

Park (Uganda), where a group (size 59–104) was followed, and routes were recorded on paper maps showing detailed outlines of a trail system that approximately followed 50×50 m squares. There were 363 follows, with an average of 7 h per follow, for 2564 h of observation in total. A group of black howler monkeys (size 6–8) were followed between January and August 2010 and between May and July 2011 in Calakmul Biosphere Reserve (Mexico), resulting in 214 h of observation with GPS points taken every 15 min. A group of spider monkeys (size 70–75), which display fusion–fission behaviour and travel almost always in smaller subgroups typically of 2–10 individuals, was followed between January and August 2010 and between May and July 2011 in Calakmul, that is, during the same periods as the group of black howler monkeys, resulting in 199 h of follows with GPS points taken every 5 min.

To standardize data collected with different methodologies, each recorded follow was transformed into route data, representing the path the group took. Groups were assumed to travel in

a straight line between sampling periods (30 min for simulated primates; 15 min for howler monkey; 5 min for spider monkey; 'n/a' for red colobus), and resampled at 10 min intervals.

METHODS

The Beta-binomial Distribution

The BBD is discrete and quantitative in nature, and has been applied to describe spatial, temporal and spatiotemporal point patterns with a binary mark (Madden & Hughes 1995; Shah et al. 2002; Dutilleul 2011). This distribution applies to counts of 'successes' (x) among a number of 'trials' (n) per sampling unit. It thus has three parameters: n , the number of trials per sampling unit; p , the expected probability of success in the absence of aggregation, and θ , which measures the variation in the probability of success between sampling units resulting from aggregation and following the beta function (Skellam 1948). Accordingly, θ is the

aggregation index parameter (Griffiths 1973; Madden & Hughes 1995). The probability function of the BBD is given by

$$P(X = x) = \frac{n!}{x!(n-x)!} \frac{\prod_{i=1}^x (p + (i-1)\theta) \prod_{i=1}^{n-x} (1-p + (i-1)\theta)}{\prod_{i=1}^n (1 + (i-1)\theta)}$$

for $x = 0, 1, 2, \dots, n$

(2)

where ! denotes the factorial function; \prod is the multiplication operator; $P(X=x)$ is the probability of x successes occurring in n trials. The mean and variance parameters of the BBD are np and $np(1-p)(1+n\theta)/(1+\theta)$, respectively (Madden & Hughes 1995). In the context of the application of the BBD to analyse movement data, sampling units can be seen as N quadrats of equal area, into which the landscape is first divided and where movement is being observed. Each of the quadrats is then subdivided into n cells of equal area. A cell that has been visited at least once by an animal is considered as 'visited', and x represents the number of such cells ($x = 0, 1, 2, \dots, n$) (Fig. 2). Thus, the landscape is divided into quadrats where the probability of success (i.e. chance of being visited) may be higher or lower depending on the quadrat. The magnitude of this variation in visit frequency is measured by θ , which can also be used as an aggregation index in the analysis of habitat use. This measure of aggregation does not capture the spatial relationships between quadrat counts, but captures the variability in the counts themselves (i.e. how variable are visits on

the landscape, and is the variation greater than expected if there were no aggregation?).

The addition of a temporal component to this approach was first suggested by Dutilleul (2011, Chapter 4), and consists in further subdividing the data set into equal time units (e.g. days, months, years), making each quadrat a space–time cube. It follows that the number of quadrats (N) remains the same, but the total number of cells is equal to the number of cells within a quadrat multiplied by the number of temporal divisions. Accordingly, the point pattern can be analysed temporally, by counting the number of visits per time period (i.e. visits in month 1 are counted on the first layer of cells, visits in month 2 in the second layer, etc.; Fig. 2).

Variability and Autocorrelation in Animal Visits of a Landscape

The variance parameter of a beta-binomial distribution is calculated differently than that for a binomial distribution because the calculation of the former does not assume that the 'trials' are independent (i.e. an animal's visits of cells within a quadrat can be correlated). If Y_i denotes the Bernoulli distribution (visit/no visit) for cell $i = 1, \dots, n$ in a given quadrat, X denotes the number of cells visited within the quadrat and X follows a BBD, then

$$\begin{aligned} \text{Var}(X) &= \text{Var}\left(\sum_{i=1}^n Y_i\right) + 2 \sum_{i=1}^n \sum_{j=i+1}^n \text{Cov}(Y_i, Y_j) \\ &= np(1-p) \frac{(1+n\theta)}{(1+\theta)} \end{aligned}$$

(3)

After isolating the covariances and standardizing them into correlations, it becomes clear that the autocorrelation within

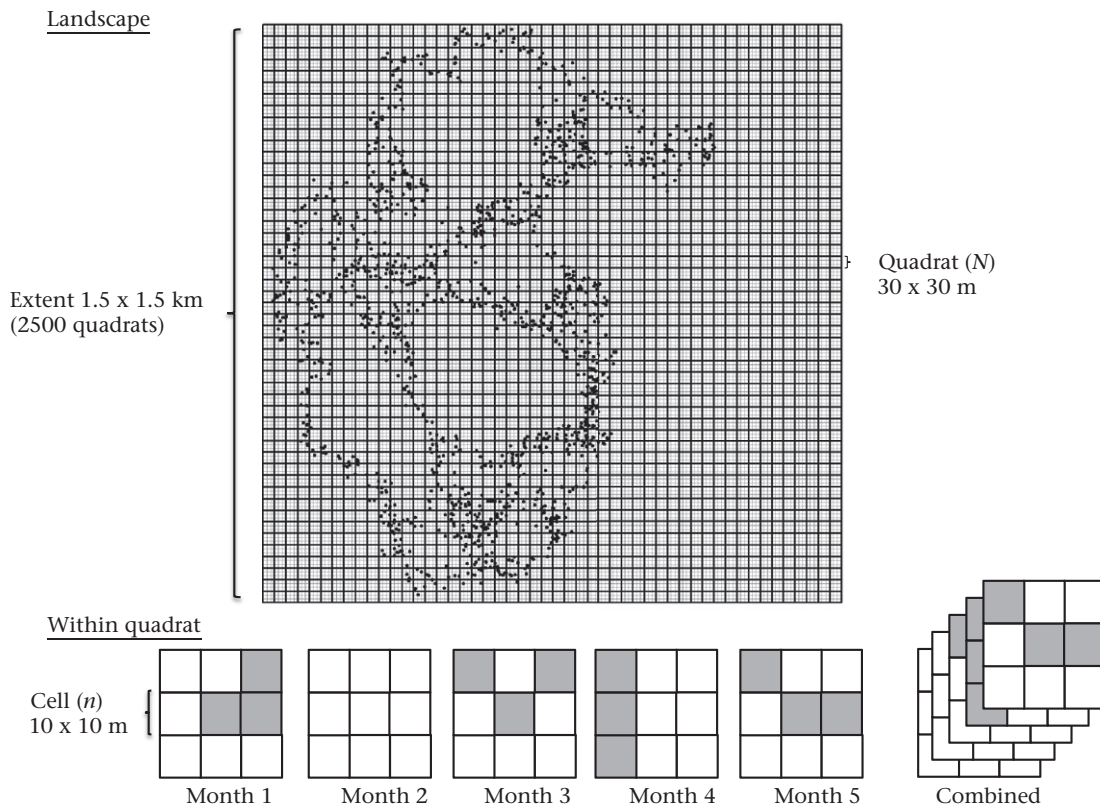


Figure 2. Diagram of the spatiotemporal analytical framework (ST-BBD), based on the beta-binomial distribution (BBD). In this example, the landscape is first subdivided into equal-sized quadrats ($N = 2500, 30 \times 30$ m), quadrats are then subdivided into cells ($n = 9, 10 \times 10$ m), and time is divided into equal units (1 month). Within each quadrat, the cells visited during each time period are combined to provide the final visit count per quadrat.

a quadrat, through the correlations between visits and no visits of cells, is a function of n and θ :

$$2 \sum_{i=1}^n \sum_{j=i+1}^n \text{Corr}(Y_i, Y_j) = \frac{(1 + \theta)}{(1 + n\theta)} - 1 \quad (4)$$

Accordingly, θ can be said to be a measure of the autocorrelation of animal movement within quadrats, and as the variability in quadrat counts grows, so too will the autocorrelation between cell visits within quadrats. If n remains constant in the proposed spatiotemporal framework, the parameter θ can thus be used to answer the following question: 'If a primate has visited a quadrat before, is it more likely to visit it again?'

Another approach, based on the so-called 'Mantel correlograms', has been proposed to characterize autocorrelation in animal movement data sets (Cushman et al. 2005); the Mantel correlogram was originally derived by Oden & Sokal (1986) and Sokal (1986) from Mantel's (1967) test statistic. Whereas the Mantel test statistic essentially is a correlation coefficient between spatial distances and corresponding temporal distances (e.g. times taken during travels), the hypotheses tested in the analysis of Mantel correlograms concern differences in mean time travelled among a certain number of spatial distance classes (Dutilleul et al. 2000, page 149). This approach has been used to assess how reliant the current movement is on past movements (Cushman et al. 2005). The ST-BBD method proposed here and the Mantel correlograms measure autocorrelation in two different aspects of the same phenomenon: habitat use (ST-BBD) versus movement patterns (Mantel correlograms). Given the application or question of interest, it would be important to know which type of autocorrelation it would be better to measure. To highlight the differences in the two approaches, consider an example in which the results would be drastically different. If animals travelled in paths that were similar to previous paths, but were shifted in space, this would result in a significant result in the Mantel test (relationship between distances in space and time), whereas in the ST-BBD approach, since paths did not pass over the same area, it would measure no small-scale aggregation (habitat use).

Variograms and spatial correlograms based on Moran's I and Geary's c statistics (instead of Mantel's test statistic) can be used to analyse and quantify autocorrelation among quadrats, using the number of cells per quadrat visited by animals as raw data or quadrat counts (Madden & Hughes 1995; Dutilleul 2011, Chapter 5). This is likely to be an important factor to consider when addressing questions related to the definition of quadrat size and how the probability of quadrat visit varies with landscape properties (see below).

Implementation of the Spatiotemporal BBD Framework

To create the ST-BBD framework, a program was written in the Java programming language with GeoTools (geotools.org), in which the movement data (e.g. route data) and grid characteristics (i.e. number of quadrats, quadrat and cell sizes, temporal division) are the inputs. The program uses the scale inputs provided by the user to generate a spatiotemporal sampling framework, and outputs the number of cells visited in each quadrat. The analysis of the quadrat counts was then performed with the SAS v.9.2 software (SAS Institute, Chicago, IL, U.S.A.), using a code graciously provided by L. V. Madden, as well as in the R programming language. The SAS code calls the procedure NLMIXED to fit the BBD to the count data of visits per quadrat, where the estimation of p and θ is performed by maximum likelihood. The Bayesian Information Criterion (BIC), a measure of the goodness of fit of a model that incorporates a relevant penalty for the number of parameters (Schwarz 1978),

was used to determine whether the BBD was more appropriate than the binomial distribution. When this was found to be the case, it meant there was sufficient aggregation in the movement data to warrant the use of the BBD.

Analyses

We first evaluated the ST-BBD framework by applying it to the simulated data, using a timescale of 30 days and a spatial grid of 50×50 quadrats (30×30 m each), with 3×3 cells (10×10 m each). The simulated data are controlled in terms of sampling regimes, and include very different movement behaviours, offering ideal data sets to explore the use of the new framework.

We then applied the ST-BBD framework to estimate the levels of aggregation in monthly movement data from the three primate species individually to analyse within-species variation. To control for differences in sampling effort between months, monthly movement patterns were first produced from five follows within each month. Movement patterns were then sampled using a minimum bounding grid, with 30×30 m quadrats subdivided into nine 10×10 m cells, to obtain visit counts per quadrat for the red colobus and the howler monkey (ranges of spatial autocorrelation in counts: red colobus, 51 m; howler monkey, 47 m); the spider monkey was sampled with 60×60 m quadrats subdivided into 36 10×10 m cells (range of spatial autocorrelation in counts: 80 m). For the red colobus, we estimated monthly aggregation within 3-month periods. For the howler and spider monkeys, monthly aggregation was estimated within 2-month periods due to smaller samples sizes.

To compare levels of habitat use between species, sampling effort was controlled by resampling species data to equal observation times. Given the limited data available for the spider and howler monkeys (February–April, June and July), a time span covering 5 months was chosen. A subsample was taken from each primate species, resulting in an average of 176 h of follows for each primate (howler monkey: 174 h, spider monkey: 177 h, red colobus: 178 h). A minimum bounding grid of 25×25 quadrats (60×60 m), with 6×6 cells (10×10 m), was used to obtain spatial and temporal counts for each species (Fig. 2).

Comparisons between θ estimates, obtained by maximum likelihood, are based on 95% confidence intervals, assuming estimates are normally distributed (i.e. an asymptotic property of maximum likelihood estimators).

RESULTS

Simulated Data

Movement data generated with the routing model produced quadrat counts in which a few quadrats had a high number of visits and many quadrats had zero visits (Fig. 3). Comparatively, movement data generated with the nonrouting model produced more quadrats with low numbers of visits and less quadrats with high numbers of visits, but fewer quadrats with no visits (Fig. 3).

When fitting the BBD to quadrat counts, we found that the BBD consistently fitted better than the binomial distribution, indicating aggregation in the data (i.e. the probability of visit was uneven throughout the landscape; Table 1). This difference was less striking, though, in the data simulated with the nonrouting model (Table 1). At the scale of 1 month, when considering movement patterns for each month separately (purely spatial analysis), there was no consistent trend in the aggregation index estimates from the two simulated movement data sets (Fig. 4). When considering all 5 months together (spatiotemporal analysis), the difference between the two estimated measures of aggregation was

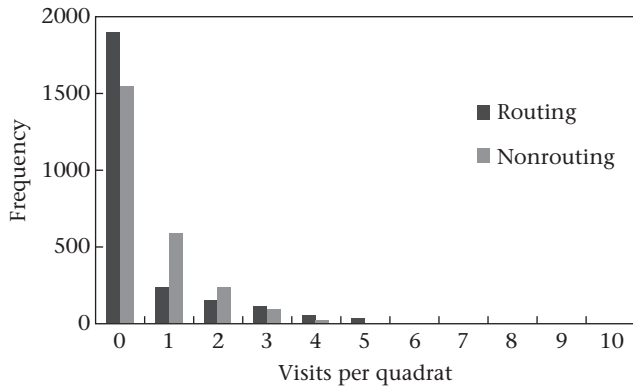


Figure 3. Histograms of the number of visits within quadrats for the simulated data, obtained with the routing versus nonrouting model. A count of visits for each quadrat was done each month ($n = 9$ cells), then all 5 months were combined for the final quadrat count presented in this figure ($n = 45$ cells).

statistically significant (Fig. 4), the estimate dropping in the case of the nonrouting model (88% lower) compared to the routing model (10% lower); this result clearly reflects the amount of overlap between months in the latter case.

Collected Data

To obtain visit counts per quadrat, the movement patterns of the three primate species were sampled for an appropriate quadrat size taking into account differences in spatial autocorrelation and field extent (Fig. 5). Monthly estimates of the visit aggregation index for the red colobus were made for all the months with more than five follows per month (35 months in total; Fig. 6a). Estimates were then computed over 3-month periods, capturing habitat reuse by the red colobus between months (Fig. 6b). This revealed large fluctuations in aggregation over time, with a peak of index estimates in August–February 2008 (Fig. 6b). The estimated aggregation index for all months combined, capturing habitat reuse over the 35 months, was 0.076, that is, 63% lower than the mean (0.204) of estimates over 3-month periods. Sampling effort was not significantly correlated with movement aggregation index estimates (Spearman rank correlation: $r_s = 0.32$, sample size = 35, $P = 0.07$).

Spider monkey movement data showed significantly higher levels of aggregation during the dry season (March–April) than

Table 1
Estimates of the index of aggregation parameter (θ) for the data simulated using two behaviour models: considering individual months and combined months (All)

Month	Binomial (BIC)	Beta-binomial (BIC)	θ	Lower 95% CI	Upper 95% CI	Mean np
Nonrouting behaviour model						
1st	2224.3	2043.9	0.085	0.058	0.111	0.137
2nd	2081.4	2014.6	0.042	0.026	0.059	0.130
3rd	1858.8	1750.4	0.063	0.040	0.086	0.108
4th	1798.7	1764.5	0.031	0.016	0.046	0.107
5th	2001.7	1911.8	0.052	0.032	0.071	0.121
All	5524.2	5321.7	0.012	0.010	0.015	0.603
Routing behaviour model						
1st	1891.9	1786.2	0.061	0.038	0.083	0.111
2nd	2066.6	1934	0.067	0.044	0.090	0.125
3rd	1912.6	1734.4	0.092	0.062	0.123	0.110
4th	1819.2	1750.8	0.046	0.027	0.065	0.107
5th	1780.5	1729.7	0.039	0.022	0.056	0.104
All	6080.2	4737.9	0.055	0.046	0.064	0.554

The Bayesian information criterion (BIC) is presented for the fitted beta-binomial and binomial distributions; lower scores imply a better fit.

during the wet season (July–Aug) (95% CI for the difference between 2-month means: March–April and July–Aug = 0.058, 0.358). The aggregation index estimated by using all months was 0.12, that is, 54% lower than the mean (0.27) of 2-month estimates (Fig. 6c). Sampling effort was not correlated with movement aggregation index estimates (Spearman rank correlation: $r_s = 0.50$, sample size = 5, $P = 0.45$).

Howler monkey movement data showed no significant difference in aggregation between months, February versus March and March versus April (Fig. 6d). The month of April had a total of 36 h, compared to 37 h in March and 40 h in February. The overall estimate of aggregation was 0.23, that is, 22% lower than the mean (0.29) of 2-month estimates.

In between-species comparisons, movement aggregation of the spider monkey (mean index value: 0.192) differed from that of the howler monkey (mean: 0.107) (Table 2). The value for red colobus fell between these values, with a mean aggregation index of 0.157 (Table 2). Again, raising the temporal scale to 5 months (as with the simulated data; see Fig. 4), the aggregation index estimates dropped compared to the average of individual monthly estimates for all three species (red colobus 45%, spider monkey 38%, and howler monkey 33%).

DISCUSSION

Within the ST-BBD framework, movement behaviour that frequently reuses specific areas results in higher variation in the probability of visits on the landscape. Movement behaviour that does not reuse the same areas results in lower values and variability of the probability of visits on the landscape. Using the BBD parameter θ , alias ‘aggregation index’, as a measure of this variability, we were able to quantify differences in habitat use and assess their significance.

The differences in movement behaviour between the routing and nonrouting models were successfully distinguished as a result of the spatiotemporal nature of the ST-BBD. By breaking up space into space–time cubes, movement patterns that overlapped in the same quadrat at different times were captured. In contrast, the purely spatial approach, in which no temporal component was considered, did not reveal important differences between models. This was shown with simulated data, by comparing monthly estimates of θ with ‘combined’ estimates obtained over 5 months (Fig. 4).

The spatiotemporal approach, which incorporates an overlap of months, was useful in distinguishing movement behaviours within each primate species (Fig. 5). In the case of the red colobus, it allowed the detection of significant fluctuations through time, a peak in aggregation centred on August 2007–February 2008, and a slight decrease overall in the aggregation of movement patterns on a monthly basis, suggesting a decrease in revisits to similar sites between months over a long period (Fig. 6b). Spider monkeys showed a significantly higher level of aggregation during the dry season. This could be a consequence of the high use of a single tree species (*Ficus* spp.) during the dry season, whereas fruit resources in the wet season are generally highly available due to the mast fructification of another highly preferred species (*Brosimum ali-castrum*; R. U. Hernández-Sarabia, unpublished data). Similar significant variation was not found within the howler monkey movement patterns (Fig. 6d).

Comparisons between the three species showed that spider monkeys had the highest movement aggregation (Table 2). Overall, spider monkeys travelled farther (~34.8 km during the 177 h subsample) than howler monkeys (~5.3 km) and red colobus (~4.5 km). Such interspecific differences are likely partially induced by the more dispersed food sources for frugivorous primates. The spider monkey movement patterns suggest the

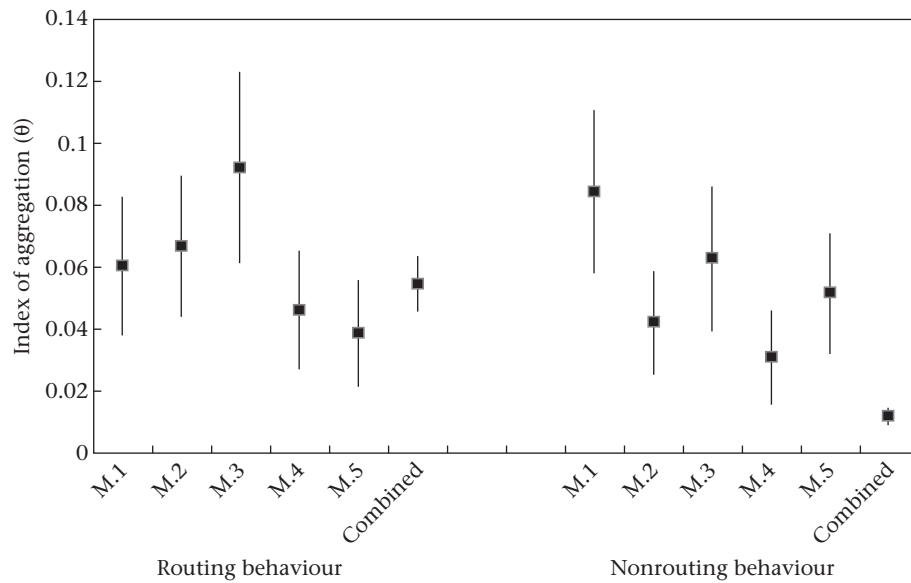


Figure 4. Aggregation index estimates for each month ($n = 9$, $N = 2500$) and for all 5 months combined ($n = 45$, $N = 2500$), for both types of simulated movement behaviour. Bars represent 95% confidence intervals.

heavy use of a central area (sleeping sites), with forays to the periphery, suggesting a central foraging strategy (Chapman 1988). The fact that spider monkeys started at, and often returned to, a few main sleeping areas each night resulted in a few quadrats with higher numbers of visits. Routing behaviour was evident in the movement from and to these central areas (Fig. 5). Also, fruits are more patchily distributed than leaves, and visits of fruit patches could influence the level of aggregation in spider monkey movement data, compared to the other two species, which are more folivorous. Howler monkeys moved more slowly through the landscape and revisited sites less often. Aggregation index estimates for the red colobus showed higher variability within a month and through the standard error of the combined estimate, with no significant difference with the spider and howler monkeys, though.

The amount of data collected for the red colobus allowed exploration of habitat use over a longer temporal extent than that for the howler and spider monkeys, in search of possible explanations for the observed variability in movement aggregation. The observed red colobus group has gone through substantial size changes, from an initial group size of about 59 in July 2006 to about 104 in September 2011. Along with this increase in group size over time, we observed a global decreasing trend in the aggregation index estimates (Fig. 6b). Using a simple linear regression, predicting θ by time since data collection started (31 months) revealed a significant decrease (intercept = 0.248, slope = -0.002 , $F_{1,29} = 5.48$, $R^2 = 0.16$, $P = 0.03$). The increasing group size could have influenced this decrease in aggregation by depleting food sources to a greater spatial extent (i.e. increasing revisit times), or by causing the group to travel farther because of increased intra-group feeding competition (i.e. selecting new areas). Large oscillations in the estimates of θ were also seen at smaller timescales (Fig. 6b), suggesting other influencing factors. The inclusion in the regression of an estimate of food availability, derived from phenology data of tree species in the study area (Chapman et al. 2005), provided gains in predictive power ($F_{2,28} = 7.3$, $R^2_{\text{adj}} = 0.30$, $P = 0.003$). This measure of food availability was also significantly correlated with estimates of θ once the global decreasing trend was removed (Pearson's sample simple linear correlation coefficient: $r = 0.41$, $t_{29} = 2.44$, $P = 0.02$). This suggests that both group size and food availability could be affecting the re-

use of habitat by the group between months, and fits with current theory regarding the group foraging behaviour of the red colobus (Snaith & Chapman 2008).

The temporal and spatial scales chosen in applications of the ST-BBD method will affect the aggregation measured. In our study, the analyses of movement data at a spatial scale of 60×60 m quadrats and on a timescale of 1 month revealed higher levels of spatio-temporal aggregation for the spider monkey. However, the categorization of landscape as visited or not visited during a month did not permit us to determine the type of use of the visited area. An animal simply passing by an area and an animal spending much of the day within a given area would both be considered visiting the site at the time of observation. To capture residency time within patches for primates, temporal scales shorter than 1 month should be considered, thereby measuring a different type of aggregation and habitat use.

Looking further at effects of spatial and temporal sampling and subsampling units on estimates of θ , we varied quadrat and cell size choices using simulated data (routing and nonrouting behaviours) and estimated θ each time. We found that increasing cell size and the total number of quadrats separately produced increased θ estimates; conversely, increasing quadrat size spatially and temporal division separately resulted in decreased θ estimates. These trends make sense when thinking about what θ measures, variability of the probability of visit among quadrats and autocorrelation among cells within a quadrat, and how this variability relates to the within-quadrat spatiotemporal resolution defined by the number of cells within a quadrat for a given quadrat size. If cells are made successively smaller, resulting in more and more cells per quadrat, more detail will be captured in space and time within a quadrat (increased spatiotemporal resolution). Conversely, if spatiotemporal resolution is lowered, there will be less detail in space and time within a quadrat. At the extremes, quadrats would simply be either visited or not visited (lowest spatiotemporal resolution), or there would be an overdetailed gradient in quadrat visits (very high spatiotemporal resolution). Direct comparisons between estimates of θ should therefore be limited to those obtained with the same or similar scale choices (e.g. Table 2).

For between-species comparisons of the level of aggregation in visits, the sampling effort would also need to be controlled as much

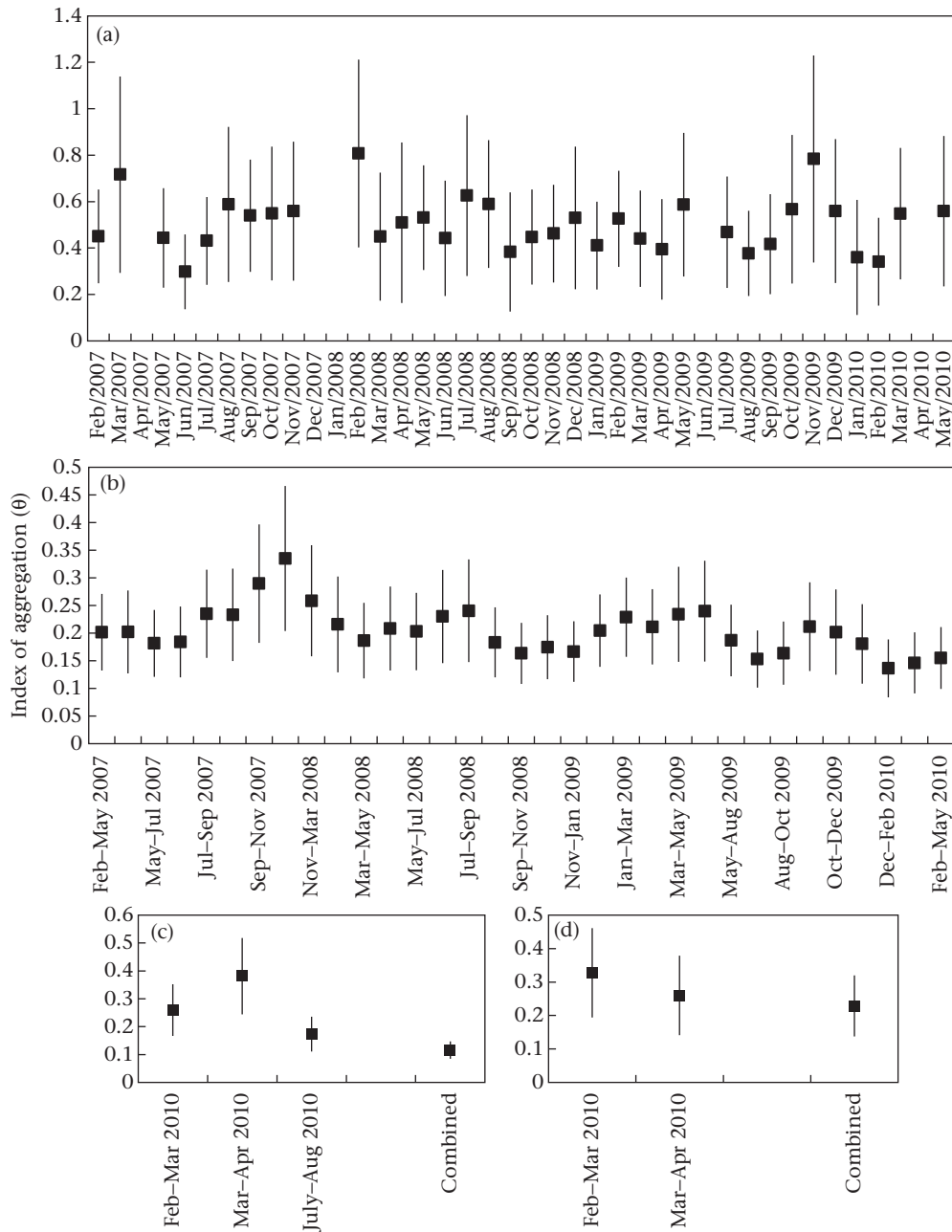


Figure 5. Index of aggregation for monthly movement patterns of the red colobus, spider and howler monkeys. Aggregation index estimates for the red colobus are presented: (a) for each month (purely spatial approach) and (b) for 3-month periods (spatiotemporal approach). Aggregation of movements for (c) the spider monkey and (d) the howler monkey are estimated over 2-month periods (spatiotemporal approach). The 'combined' value is an estimate over all months. Bars represent 95% confidence intervals for the true, but unknown, aggregation index values.

as possible, for example by resampling to equal observation times. Within the red colobus species, a minimum of about 60 h of follow time per month was sufficient for obtaining significant aggregation index estimates; this number of hours might be used as a basis for minimum sampling regime in future studies (Fig. 7).

Sampling rates of movement could also affect the estimated mean visit count of the BBD, as a group would likely visit more novel areas over time. Madden & Hughes (1995) have shown that, although there is no mathematical relationship between the estimated mean of the BBD and the θ estimate, they are often related. From the plant disease literature, we would expect that as the mean number of animal visits (or disease incidents in the plant disease literature) increases, estimates of θ would be increasing first and

then decreasing, following an upside-down U-shaped relationship. In our sampling conducted in space and time (i.e. using space–time cubes), the number of cells per quadrat (n) increased with increasing sampling in time, resulting in shifting values of the mean (np). The resulting trend in θ estimates obtained from 1, 3 or 6 months of collected data was a decreasing parabolic one, as more sampling time was added. Madden & Hughes (1995) suggested the use of the binary power law to examine the relationship between the expected mean and the observed variance.

The results of our evaluation of the ST-BBD method clearly show that it can help develop insight into a species' use of habitat by estimating spatial and temporal aggregation in movement data. Furthermore, the corresponding approach could be useful in

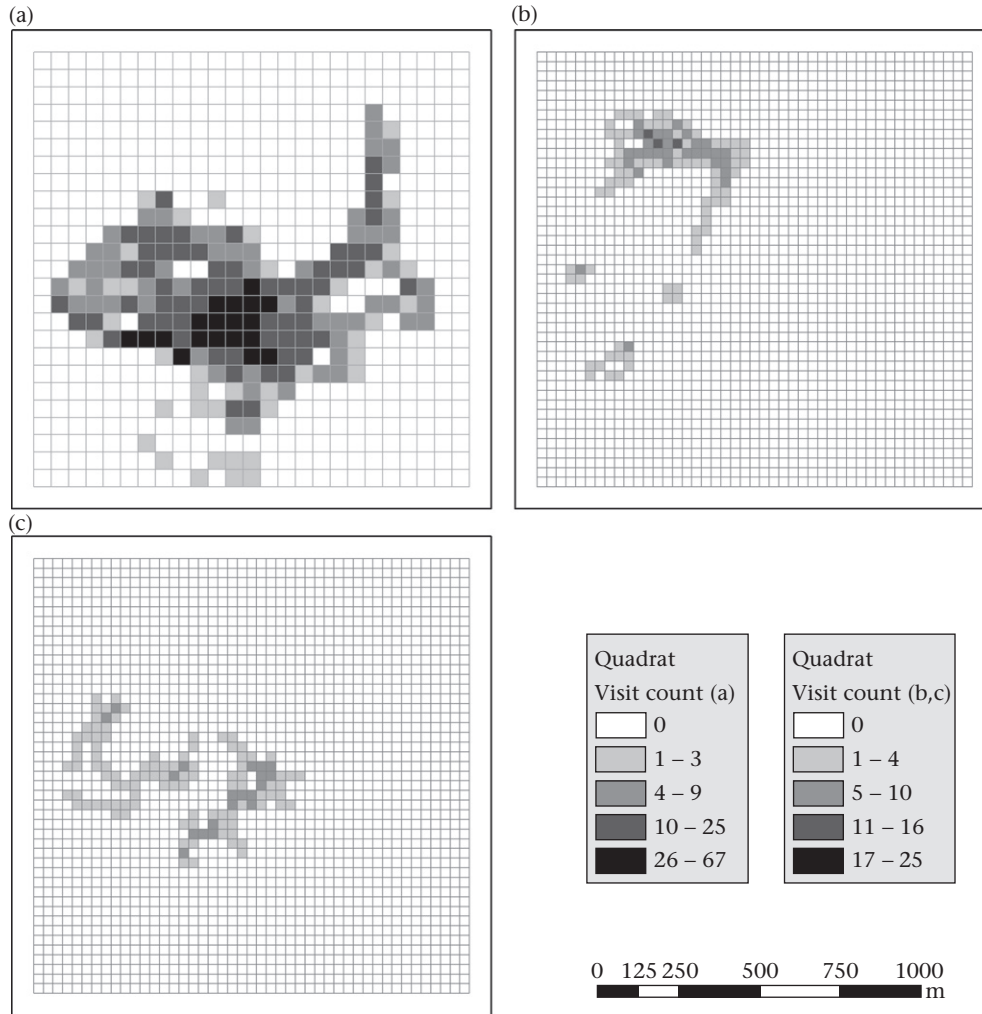


Figure 6. Movement patterns of three primate species: (a) spider monkey, (b) red colobus and (c) howler monkey. The spider monkey sample grid is composed of 25 × 25 quadrats ($N = 625$, 60 × 60 m). The red colobus and howler monkey sample grids are composed of 45 × 45 quadrats ($N = 2500$, 30 × 30 m). Quadrats are shaded based on the number of visited 10 × 10 m cells (a: $n = 1296$; b, c: $n = 45$); quadrats were resampled every month during a 5-month period.

estimating the effects of varying resource distributions (e.g. seasonal) or landscape structures (e.g. corridors, fragmentation) on habitat use. In the near future, especially relevant applications could examine the landscape effects on host–parasite interactions when movement behaviour and habitat use are important factors

Table 2
Estimates of the index of aggregation (θ) for the three primate species

Timescale	Primate	θ estimate	Lower CI	Upper CI	Mean or difference of means
5 months	Spider monkey	0.192	0.137	0.246	0.018
	Howler monkey	0.107	0.049	0.166	0.003
	Red colobus	0.157	0.063	0.252	0.003
Differences	Spider–Howler	0.085	0.005	0.164	0.015
	Spider–Red colobus	0.035	−0.075	0.143	0.015
	Red colobus–Howler	0.050	−0.061	0.161	<0.000

Theta estimates based on 5-month movement patterns are given, together with their 95% confidence intervals. In all cases, the beta-binomial distribution provided a better fit than the binomial distribution based on Bayesian information criterion (BIC).

(e.g. with gastrointestinal parasites with a free-living stage in the external environment, or with tick-borne diseases such as the Lyme disease), and the impact of seed disperser movement on the recruitment of plant seedlings.

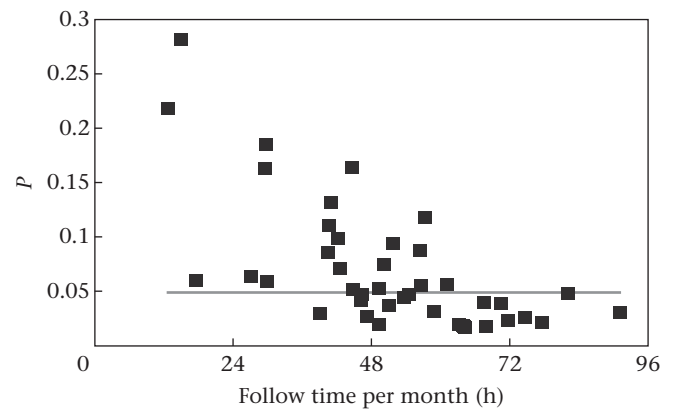


Figure 7. Effect of sampling effort (number of hours of follows per month) on the probability of significance of monthly aggregation index estimates. The effect was tested on the red colobus data (total of 46 months). The grey line represents the significance level of 0.05.

Acknowledgments

Funding for the long-term Ugandan research was provided by Canada Research Chairs Program, Natural Science and Engineering Research Council of Canada, National Geographic, the National Science Foundation and by National Institutes of Health (NIH) grant TW00927 as part of the joint NIH–National Science Foundation Ecology of Infectious Disease Program and the U.K. Economic and Social Research Council to C.C. Funding for field work in Mexico and Uganda was provided by Primate Conservation, Inc., and by a Rufford Small Grant to R.R.H. and U.H.S. Support for the analysis and writing was provided by Natural Science and Engineering Research Council of Canada (Raja Sengupta and C.C.) and Fonds de Recherche Nature et Technologies (Tyler Bonnell). Permission to conduct the research in Uganda was given by the National Council for Science and Technology and the Uganda Wildlife Authority. We would like to extend special thanks to Juan Carlos Serio-Silva for providing fundamental help with the Mexican part of this project, and to Mel Lefebvre and Lauren Chapman for helpful comments on this manuscript.

References

- Bartel, R. A., Oberhauser, K. S., De Roode, J. C. & Altizer, S. M. 2011. Monarch butterfly migration and parasite transmission in eastern North America. *Ecology*, **92**, 342–351.
- Bonnell, T. R., Sengupta, R. R., Chapman, C. A. & Goldberg, T. L. 2010. An agent-based model of red colobus resources and disease dynamics implicates key resource sites as hot spots of disease transmission. *Ecological Modelling*, **221**, 2491–2500.
- Bryson, J. J., Ando, Y. & Lehmann, H. 2007. Agent-based modelling as scientific method: a case study analysing primate social behaviour. *Philosophical Transactions of the Royal Society B*, **362**, 1685–1698.
- Chapman, C. A. 1988. Patterns of foraging and range use by three species of neotropical primates. *Primates*, **29**, 177–194.
- Chapman, C. A., Chapman, L. J. & Gillespie, T. R. 2002. Scale issues in the study of primate foraging: red colobus of Kibale National Park. *American Journal of Physical Anthropology*, **117**, 349–363.
- Chapman, C. A., Chapman, L. J., Struhsaker, T. T., Zanne, A. E., Clark, C. J. & Poulsen, J. R. 2005. A long-term evaluation of fruiting phenology: importance of climate change. *Journal of Tropical Ecology*, **21**, 31–45.
- Chetkiewicz, C. L. B., Clair, C. C. S. & Boyce, M. S. 2006. Corridors for conservation: integrating pattern and process. *Annual Review of Ecology, Evolution and Systematics*, **37**, 317–342.
- Cushman, S. A., Chase, M. & Griffin, C. 2005. Elephants in space and time. *Oikos*, **109**, 331–341.
- Dray, S., Royer-Carenzi, M. & Calenge, C. 2010. The exploratory analysis of autocorrelation in animal-movement studies. *Ecological Research*, **25**, 673–681.
- Duttilleul, P. R. L. 2011. *Spatio-temporal Heterogeneity: Concepts and Analyses*. New York: Cambridge University Press.
- Duttilleul, P., Stockwell, J. D., Frigon, D. & Legendre, P. 2000. The Mantel test versus Pearson's correlation analysis: assessment of the differences for biological and environmental studies. *Journal of Agricultural, Biological and Environmental Statistics*, **5**, 131–150.
- Edwards, A. M., Phillips, R. A., Watkins, N. W., Freeman, M. P., Murphy, E. J., Afanasyev, V., Buldyrev, S. V., da Luz, M. G. E., Raposo, E. P., Stanley, H. E., et al. 2007. Revisiting Levy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature*, **449**, 1044–1045.
- Griffiths, D. A. 1973. Maximum likelihood estimation for the beta-binomial distribution and an application to the household distribution of the total number of cases of a disease. *Biometrics*, **29**, 637–648.
- Hemelrijk, C. 2002. Despotic societies, sexual attraction and the emergence of male 'tolerance': an agent-based model. *Behaviour*, **139**, 729–747.
- Hughes, G. & Madden, L. V. 1993. Using the beta-binomial distribution to describe aggregated patterns of disease incidence. *Phytopathology*, **83**, 759–763.
- Humphries, N. E., Queiroz, N., Dyer, J. R., Pade, N. G., Musyl, M. K., Schaefer, K. M., Fuller, D. W., Brunnschweiler, J. M., Doyle, T. K., Houghton, J. D., et al. 2010. Environmental context explains Levy and Brownian movement patterns of marine predators. *Nature*, **465**, 1066–1069.
- Keating, K. A. & Cherry, S. 2009. Modeling utilization distributions in space and time. *Ecology*, **90**, 1971–1980.
- Madden, L. V. & Hughes, G. 1995. Plant disease incidence: distributions, heterogeneity, and temporal analysis. *Annual Review of Phytopathology*, **33**, 529–564.
- Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.
- Miller, H. J. 2005. A measurement theory for time geography. *Geographical Analysis*, **37**, 17–45.
- Millsbaugh, J. J., Gitzen, R. A., Kernohan, B. J., Larson, M. A. & Clay, C. L. 2004. Comparability of three analytical techniques to assess joint space use. *Wildlife Society Bulletin*, **32**, 148–157.
- Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D. & Smouse, P. E. 2008. A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences, U.S.A.*, **105**, 19052–19059.
- Oden, N. L. & Sokal, R. R. 1986. Directional autocorrelation: an extension of spatial correlograms to two dimensions. *Systematic Zoology*, **35**, 608–617.
- Orians, G. H. & Pearson, N. E. 1979. On the theory of central place foraging. In: *Analysis of Ecological Systems* (Ed. by D. J. Horn, R. D. Mitchell & G. R. Stairs), pp. 155–177. Columbus: Ohio University Press.
- Reyna-Hurtado, R., Rojas-Flores, E. & Tanner, G. W. 2009. Home range and habitat preferences of white-lipped peccaries (*Tayassu pecari*) in Calakmul, Campeche, Mexico. *Journal of Mammalogy*, **90**, 1199–1209.
- Schwarz, G. E. 1978. Estimating the dimension of a model. *Annals of Statistics*, **6**, 461–464.
- Shah, D. A., Clear, R. M., Madden, L. V. & Bergstrom, G. C. 2002. Summarizing the regional incidence of seed-borne fungi with the beta-binomial distribution. *Canadian Journal of Plant Pathology*, **24**, 168–175.
- Skellam, J. G. 1948. A probability distribution derived from the binomial distribution by regarding the probability of success as variable between sets of trials. *Journal of the Royal Statistical Society, Series B*, **10**, 257–261.
- Snaith, T. V. & Chapman, C. A. 2008. Red colobus monkeys display alternative behavioural responses to the costs of scramble competition. *Behavioral Ecology*, **19**, 1289–1296.
- Sokal, R. R. 1986. Spatial data analysis and historical processes. *Proceedings of the International Symposium on Data Analysis and Informatics*, **4**, 29–43.
- Vellend, M., Myers, J. A., Gardescu, S. & Marks, P. L. 2003. Dispersal of trillium seeds by deer: implications for long-distance migration of forest herbs. *Ecology*, **84**, 1067–1072.
- Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Murphy, E. J., Prince, P. A. & Stanley, H. E. 1996. Levy flight search patterns of wandering albatrosses. *Nature*, **381**, 413–415.