How to characterise shared space use networks

Klara M. Wanelik[1*] & Damien R. Farine[2,3,4*]

[1] Institute of Integrative Biology, University of Liverpool, Liverpool, United Kingdom

[2] Department of Collective Behavior, Max Planck Institute of Animal Behavior, Konstanz, Germany.

[3] Department of Biology, University of Konstanz, Germany.

[4] Center for the Advanced Study of Collective Behaviour, University of Konstanz, Germany.

[*] Corresponding authors

Email: klara.wanelik@liverpool.ac.uk; dfarine@ab.mpg.de

ABSTRACT

Studying the social behaviour of small or cryptic species often relies on constructing space-sharing networks from sparse point-based observations of individuals. Such an approach assumes that individuals that have greater shared space use will also interact more. However, there is very little guidance on how much data are required to construct meaningful space-sharing networks, or on how to interpret the relationships generated from such networks. In this study, we quantify the robustness of space-sharing networks to different sampling regimes, providing much needed guidance for informing the choice of sampling regime when designing studies to accurately quantify space sharing. We then describe the characteristics of space use in a wild population of field voles (*Microtus agrestis*), and use this empirical dataset to develop a new method for generating shared space use networks which are generally more strongly correlated with the real network, differ less from the real network and are more powerful to detect effects present in the real network. Our method pools data among individuals to estimate a general home range profile for a given set of individuals. Combining these profiles with the individual-level observation data then allows us to better estimate their overlap in space and requires less data. Our new method provides the potential to generate meaningful space-sharing networks, and in doing so, to address a range of key questions in ecology and evolution, even when point-based observations of individuals are sparse.

INTRODUCTION

Social networks are central to addressing many of the key questions in ecology and evolution (Cantor et al., 2019). However, network construction remains a major challenge in many systems because large numbers of observations are needed to construct meaningful networks (Farine & Whitehead, 2015). Recent technological improvements for collecting proximity, contact, or interaction data allow much more detailed networks to be constructed by improving the temporal resolution at which the data are collected (Douglas, Ji, & Clout, 2006; Rutz et al., 2012; Ryder, Horton, van den Tillaart, Morales, & Moore, 2012; Berkvens, Olivares, Mercelis, Kirkpatrick, & Weyn, 2019). However, for many smaller or more cryptic species, where observation remains difficult, most studies still rely on collecting data by trapping individuals and constructing space-sharing networks from catching different individuals in the same trap across different trapping events (e.g. Perkins, Ferrari, & Hudson, 2008; Porphyre, Stevenson, Jackson, & McKenzie, 2008; Grear, Perkins, & Hudson, 2009; Grear, Luong, & Hudson, 2013; VanderWaal, Atwill, Hooper, Buckle, & McCowan, 2013; Davis et al., 2014). To date, no study has quantified whether sparse observations like this allow us to construct networks that represent the real underlying patterns of shared space use (an issue highlighted in Tompkins, Dunn, Smith, & Telfer, 2011 in relation to disease transmission), and, therefore, whether we can extract meaningful biological relationships from such shared space use networks.

Shared space use networks can be constructed from a range of different data-collection traps, most commonly live-capture traps and camera traps. These methods are characterised by typically catching or observing single individuals at any one time, but having the capability to observe multiple individuals in the same location(s) over time. Individuals that are then

observed (trapped) at the same location are considered to be connected, with binary edges (there or not) between them, or, if constructing a weighted network, with the number of shared locations defining the strength of their connection. However, the definition of these edges, and therefore the interpretation of these networks, can be unclear. For example, both VanderWaal et al. (2013) and Grear et al. (2013) use networks generated from such observed space-sharing events to represent both shared space use and the likelihood of social interaction. The correct definition for these networks is that they represent shared space use, as the edges define the estimated spatial overlap between two individuals. The utility of these networks in representing social interactions comes from whether the assumption that two animals that share the same space are more likely to come into direct contact is accurate. The validity of this assumption will depend on the biology of the system (Farine, 2015).

Beyond the interpretation of the meaning of network connections, there has been no direct quantification of the robustness of shared space use networks to different sampling regimes. It has been suggested that the data-intensive nature of networks may act as a barrier to the more widespread use of networks in the fields of ecology and evolution, with wildlife systems often being data limited (Craft & Caillaud, 2011). Recent investigations into networks based on co-occurrence data (Farine & Strandburg-Peshkin, 2015) and direct observation methods (Davis, Crofoot, & Farine, 2018) have highlighted how networks are data-hungry. Constructing a meaningful network requires sufficient observations to estimate the many relationships that can potentially connect all individuals in a population (specifically: N*(N-1)/2 edges in an undirected network). Thus, the sampling effort required to reach a given accuracy potentially grows exponentially with the number of individuals represented in a given network. Previous studies quantifying the data required to construct meaningful networks have suggested that a good rule-of-thumb is that an average of 15

potential co-observations between all pairs of individuals are needed (i.e. of potential associations or interactions, which could be both individuals together or just one individual in the absence of the other; Farine & Strandburg-Peshkin, 2015; Davis et al., 2018). However, these estimates may not translate well to shared space use networks because estimating space use could require much less data than estimating relationship strengths, and because connections (edges) do not aim to characterise relationship strengths but rather a likelihood of encounter. Addressing this question is also crucially important because many studies that have constructed shared space use networks do not report the average number of observations per individual. Furthermore, the majority of those that do have been based on relatively few average observations of individuals. For example, VanderWaal (2013) had a mean of 11 trapping events per individual, meaning that they would have a mean of 22 potential observations from which to characterise edge weights in their network (observing two individuals apart 11 times each would give the denominator of an association index of 22, see Hoppitt & Farine, 2018).

The choice of what data to include when constructing a shared space use network can also have an impact on how meaningful the resulting network is. Many studies have defined a temporal threshold within which the observation of two individuals in the same place must occur for these observations to be counted as a connection. These are often inspired by biology, commonly the lifetime of a disease vector or pathogen in studies using networks to characterise disease transmission. For example, Porphyre et al. (2008) used 28 days (maximum survival of *Mycobacterium bovis* in the environment) and Perkins et al. (2008) used 14 days (the time needed for an infective L3 larval stage of *Heligmosomoides polygyrus* to develop from the eggs of an infected host). However, such definitions can be at odds with the definition and biological motivation behind applying a network approach. For studies that

rarely observe individuals, a shared space use network is most powerful when used as a proxy defining the potential for individuals to encounter one-another anywhere within their respective home ranges, rather than whether they actually did encounter each other. This definition contrasts with studies that can regularly observe or recapture individuals (such as Smith et al., 2018 who used PIT-tag readers at the entrance of burrows) and can directly relate observed shared space use events at the same spatial and temporal scales as transmission events occur (e.g. within a burrow on a given visit, see Farine, 2018 for further discussion). In a study where few observations of individuals are spaced widely apart, connections between individuals are unlikely to represent real candidate transmission events, but rather describe the system more generally. Greater guidance is needed about when data should consider temporal gaps in observations for constructing shared space use networks.

In this study, we conduct a quantitative evaluation to provide guidance on how to construct shared space use networks. We use simulations to explore the robustness of networks to different sampling regimes, and the inference of biological relationships under different scenarios. We also use empirical data to highlight how characteristics of space use data can help us to establish new ways to model the propensity for individuals to encounter one another. We then use these insights to describe a new method for characterising shared space use network. We show that using this method can generate a network which is generally (1) more strongly correlated with the real generative network, (2) differs less from the real network and, therefore, (3) has greater power to detect effects present in the real network, relative to networks generated from the observed space-sharing events only. Importantly, this method requires much fewer observations to reconstruct meaningful shared space use networks, than what is suggested in the more general guidelines for social networks. Finally,

we discuss the topic of inference from shared space use networks, and how appropriate it is to link network data with biological processes.

METHODS

Our study consists of two core components. First, we estimate the ability for data on shared trap use by individuals to generate networks which are robust to different sampling regimes using simulated data. Second, we describe a new method for defining the edges using the same observed data, and show that this method generates observed networks which are, for a given sampling effort, more strongly correlated with the real generative network, differ less from the real network and are more powerful to detect effects present in the real network based on the same simulation data. Our method is inspired by characteristics of how animals use space, specifically that they have a core and a periphery to their home range, and we further demonstrate that this holds true in a large-scale empirical dataset.

Simulation procedure

In brief, our study used the following procedure (see Fig. 1):

(1) We simulated a set of N individuals with home ranges defined by a centroid and characterized by a negative sigmoidal curve that highlights the declining probability of an individual to use space further away from the centroid of its home range. We defined the real network as the amount of overlap in the home ranges across all combinations of individuals.

(2) We randomly placed individuals in a spatial area containing T traps laid out in a stratified grid. We then simulated observed datasets that contained detections of individuals at traps, where the detection probability for a given individual in a given

trap was determined by the position of the trap relative to the home range of the individual (higher closer to the centroid, lower further away from the centroid).

(3) From the simulated datasets, we first constructed a shared trap network to determine how many trapping events are required to generate meaningful networks.

(4) Finally, we applied our novel method to determine whether it requires fewer data to generate meaningful networks.

1. Simulating real networks

We first drew N sets of x and y coordinates from a uniform distribution, where the boundaries of the distribution correspond to the edges of our study area (in our case, from 0 to 10 in each dimension). For each individual, we also randomly allocated a sex (male or female) and drew home range parameters based on the sex, giving males a larger home range than females. Home range parameters for males and females were based on the empirical data (see below), with added noise drawn from a normal distribution to simulate individual-level variation in home range size and shape (see Fig. S1) .

For each simulation, we generated the real network by calculating the amount of overlap in the home range between each pair of individuals. This was done by overlaying the sigmoidal curves of the two individuals' home range profiles and calculating the area under the two curves (Fig. 1a).

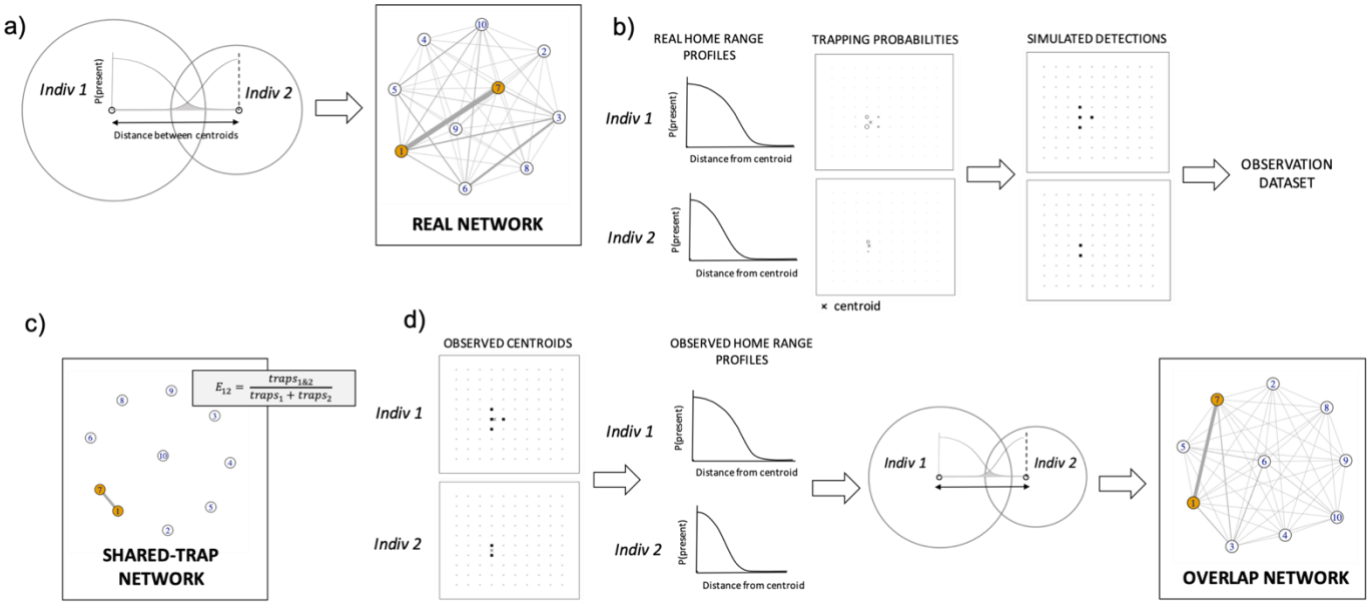

Fig. 1 Schematic showing the simulation process: (a) Simulation of real network with edge weights between two individuals (e.g. Indiv 1 and Indiv 2) equal to the overlap between their respective home range profiles (one reversed), given the distance between their centroids. (b) Generating an observed dataset by calculating the probability for a given individual to be observed in a given trap per unit effort based on its home range profile (the bigger the circle, the higher the probability), then simulating observations by repeatedly drawing from a binomial distribution (0,1) with the probability of getting a 1 for a given individual in a given trap defined by this trapping probability. (c) Generating a shared trap network where edge weights represent the number of traps in which both individuals were observed divided by the total number of traps that either individual was observed in. (d) Generating an overlap network by first calculating observed centroids for individuals, then creating a home range profile by combining data from all individuals (or by class) in a binomial model (proportion of detections of individual in a trap ~ distance to the trap, weights = total no. detections of individual), and finally using these profiles together with the centroid information to generate an overlap network with edge weights calculated as in (a).

2. *Simulating observations of individuals in traps*

We first calculated the probability for a given individual to be observed in a given trap per unit effort. We defined this probability based on the distance of the trap to the center of the individual's home range, and that individual's home range profile. We repeatedly did this for all combinations of individuals and traps (Fig. 1b).

We then simulated observations by repeatedly drawing from a binomial distribution (0,1) with the probability of getting a 1 for a given individual in a given trap defined by the trapping probability described above (Fig. 1b). We incremented the number of draws to generate more observations. In the results, we calculate effort as the mean number of observations per individual (rather than the number of draws we made in the simulation) to make our results more directly interpretable.

3. *Generating shared trap networks*

Each observation dataset contained the number of detections of each individual in each trap. We generated a shared trap network for each observation dataset by calculating the number of traps in which both individuals were observed divided by the total number of traps that either individual was observed in (Fig. 1c).

4. *Generating networks based on overlapping home range profiles*

We developed a novel method for generating shared space use networks based on estimating a population's home range profile from an observation dataset, and then calculating the estimated overlap in the home range profiles of each pair of individuals based on their observed centroids. First, we calculated each individual's centroid by taking the weighted

mean of the positions where it was observed. Second, we calculated the distance between this centroid and all of the traps where it could have been captured. Third, we recorded the number of times each individual was detected in each trap. Fourth, we fitted a negative sigmoidal curve to the relationship between the log of the distance from each centroid to each trap and the number of detections of each individual in each trap as a proportion of their total detections—this curve represents the average home range profile for those individuals. Fifth, we used the population (or sex-specific) profile to calculate the estimated overlap in space use between each pair of individuals based on the distance between their centroids (Fig. 1d). This method estimates home ranges by pooling the data among all individuals, assuming that the resulting estimation of home ranges will be more accurate in this way than by using the very few observations available at the individual level. It is possible to do this by class of individual if classes are expected to differ in their space use (as was the case in our empirical dataset; see below).

Modelling home ranges in an empirical dataset

There is a large body of literature on how best to model animal space use, and it is widely accepted that animals will have a core and a periphery to their home range. For example, both 'HR-sigma' and the mean squared distance from the center of activity (MSD, Calhoun & Casby, 1958; equivalent to the r_2 of Schoener, 1981) measure dispersion of use around an individual's centroid. We test whether this holds true in a large-scale empirical dataset, and in so doing, the utility of this simple representation of home range for modelling the space-use behaviour of large numbers of animals.

We used data from a study of field voles (*Microtus agrestis*) in Kielder Forest, UK (55˚13' N, 2˚3' W) that involved capturing individuals using live-trapping methods. Access to the

study site was provided by the Forestry Commission. Full details are given in Jackson et al. (2014). Briefly, a trapping field design was replicated at two spatially separate sites in 2008–2009, and a further two in 2009–2010. Each site was monitored by monthly trapping sessions between February (2008–2009) or April (2009–2010) and November, and contained a live-trapping grid (~0.375 ha) of 150 (10x15) regularly spaced traps (3–5 m intervals) placed in optimal habitat. Animals were marked with passive radio frequency transponders (AVID) and monitored over time, thus providing sequences of capture and recaptures. This dataset is comprised of 921 individuals and 1851 trapping events.

M. agrestis is a polygynous species, with strictly territorial males. Outside the breeding season, home ranges of *M. agrestis* vary in size from 400m2–800m2, with those of males being larger than those of females (Niethammer & Krapp, 1982). During the breeding season, home ranges of males increase to 1000–1, 500m2, with large males having the largest home ranges (Borowski, 2003). In our own study population, there is evidence for differences in the degree to which large males, small males and females are discouraged by distance (Davis et al., 2014). Therefore, we estimated home range characteristics for females and males (mean weight ≥ 25 g) separately.

Estimating the robustness of shared space use networks to different sampling regimes

We generated 1000 simulated real networks. For each real network, we produced simulated observation datasets that varied in sampling effort (number of draws from the probability of observing an individual in a trap). We designed this such that the sampling effort corresponded to a mean number of observations per individual ranging between 1 and 50 (regardless of trapping grid density), thus capturing the spectrum of what has been reported in the literature. For each observed dataset, we generated a shared trap network, and a network

using the overlap approach by reconstructing separate sigmoidal curves (or home range profiles) for males and females.

In order to assess the performance of each of these observed networks, we calculated three metrics of quality. First, we calculated the correlation between the observed network and the real network using a Mantel test. Second, we calculated the error in the estimated edge weights of the observed network by taking the sum of the absolute differences between the observed and real network edge weights (a measure of accuracy). Third, we calculated the proportion of observed networks in which we could detect a significant difference in the resulting degree (sum of edge weights) between males and females (a measure of power; males were given a larger home range than females, see point 1 of the simulation procedure). We estimated significance for each simulated dataset by comparing the observed difference in the degree between males and females to the distribution of differences in 100 permuted networks. We used node permutations, which involved randomizing the assignment of sex to the identities of each individual. We deemed the effect from an observed network to be significant if fewer than 3 of the randomised networks generated a difference that was larger than the observed one (two-tailed at $p = 0.05$, see Farine, 2017).

Variants

We repeated the procedure described above for simulated real networks with varying effect sizes: (a) an effect size equal to that in our empirical dataset, and, as this was relatively low (Fig. 2), (b) an effect size 4 times greater, and (c) an effect size 9 times greater. We also repeated the procedure using trapping grids of differing densities: (a) a 10x10 grid, and (b) a 19x19 grid within the same area.

RESULTS

Fitting model to empirical data

Consistent with space-use theory, we found evidence for a declining probability of an individual field vole to use space further away from the centroid of its home range. Furthermore, we characterised this empirical relationship, between probability of detection and distance from centroid, using a negative sigmoidal curve (Fig. 2). Our empirical data was also consistent with males having a larger home range than females (sex × log of distance interaction; effect size = 1.28; $p < 0.0001$; Fig. 2). Therefore, we fitted separate sigmoidal curves when generating overlap networks (see below).

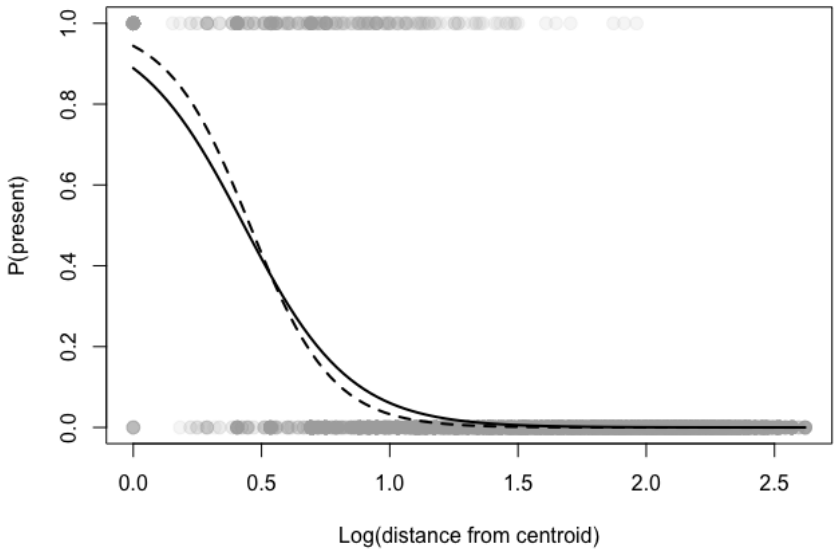

Fig. 2 Sex-specific negative sigmoidal curves for *M. agrestis* describing the change in probability of detection with increasing distance from the centroid of an individual. Line shows the fitted home range profile (solid = large males; dashed = females). Points show the raw data (whether or not an individual was detected at a location).

Performance of simulated observed networks with varying sampling effort

The number of individuals detected at least once at a trap increases with sampling effort, with a mean of 31 individuals (out of a total population of 100 individuals) present at a mean of 1–2 captures per individual, and the whole population present from 20 captures per individual (Fig. 3c).

Correlation

As the mean captures per individual increases, the shared trap network becomes more strongly correlated with the real network. At a mean of 4–5 captures per individual, the Mantel correlation coefficient between the shared trap network and the real network is 0.58. The correlation coefficient plateaus from approx. 20 captures per individual, reaching a maximum of 0.87. The overlap network shows broadly the same pattern, but is, for a given sampling effort, typically more strongly correlated with the real network than the shared trap network. At a mean of 4–5 captures per individual, the correlation coefficient between the overlap network and the real network is 0.80 (1.4 times higher than the correlation between shared trap and real networks). The correlation coefficient also plateaus earlier, from approx. 10 captures per individual, and reaches a slightly higher maximum of 0.93 (Fig. 3a).

Accuracy

As the mean number of captures per individual increases, the shared trap network becomes more accurate. At a mean of 4–5 captures per individual, the sum of the absolute differences in edge weights between the real and shared trap networks is 857, which reaches a minimum of 568 after 37 captures per individual. The overlap network shows broadly the same pattern, but is typically more accurate for a given sampling effort. For example, at a mean of 4–5 captures per individual, the absolute difference in edge weights to the real network is 665 (0.8

times that of the shared trap network). The sum of differences for the overlap network also reaches a slightly lower minimum of 552 (Fig. 3b). Only above approx. 40 captures per individual does the shared trap network become slightly more accurate than the overlap network.

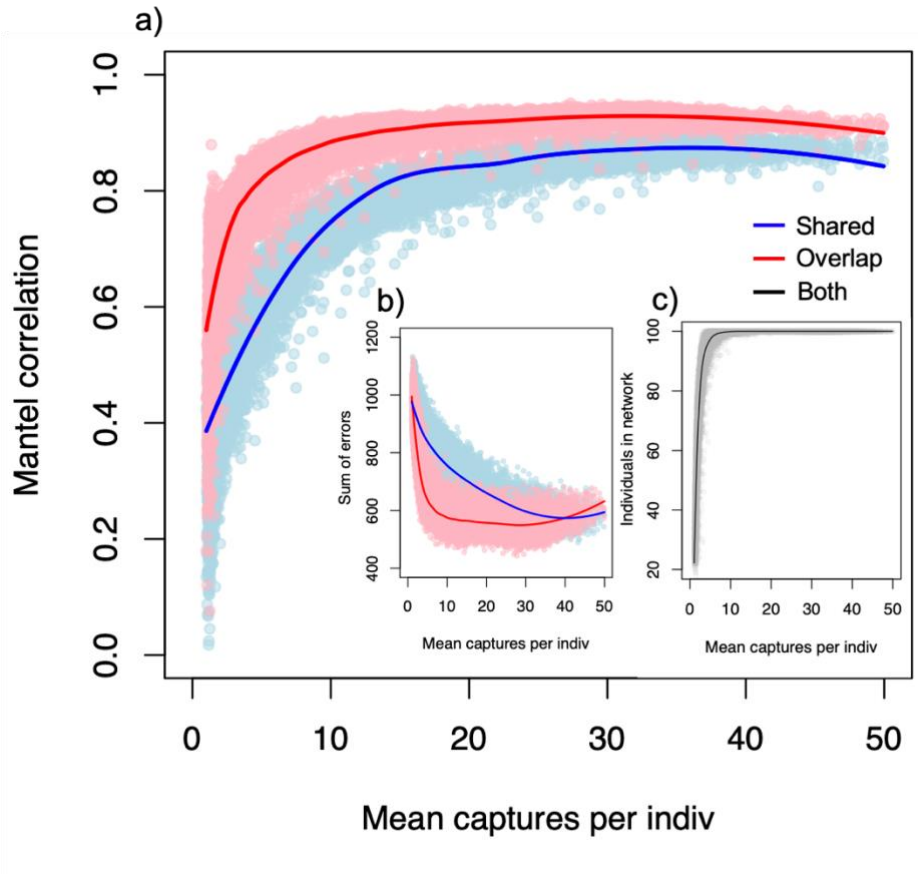

Fig. 3 Performance of observed networks with varying sampling effort on a 10x10 trapping grid, as measured by (a) Mantel correlation between observed and real networks, (b) Absolute difference in edge weights between observed and real networks, (c) Number of individuals in observed networks. LOESS smoother added to aid visual interpretation.

Power

As the mean captures per individual increases, the power of the shared trap network also increases. For example, there is nearly triple the chance of detecting a true positive at a mean of 9–10 captures (15%) compared to 4–5 captures (6%). However, the power remains low for

small effect sizes even after large numbers of captures. The overlap network shows broadly the same pattern, but has greater power to detect an effect for a given sampling effort, below approx. 30 captures per individual. For example, at 4–5 captures per individual there is nearly a 6 times higher chance of detecting a true positive in the overlap network (33%) compared to the shared trap network. Above approx. 30 captures per individual, the shared trap network has slightly greater power to detect a true positive, reaching a maximum power of approximately 52%. The power of the overlap network plateaus at approx. 10 captures per individual, reaching a slightly lower maximum of 39%; Fig. 4a).

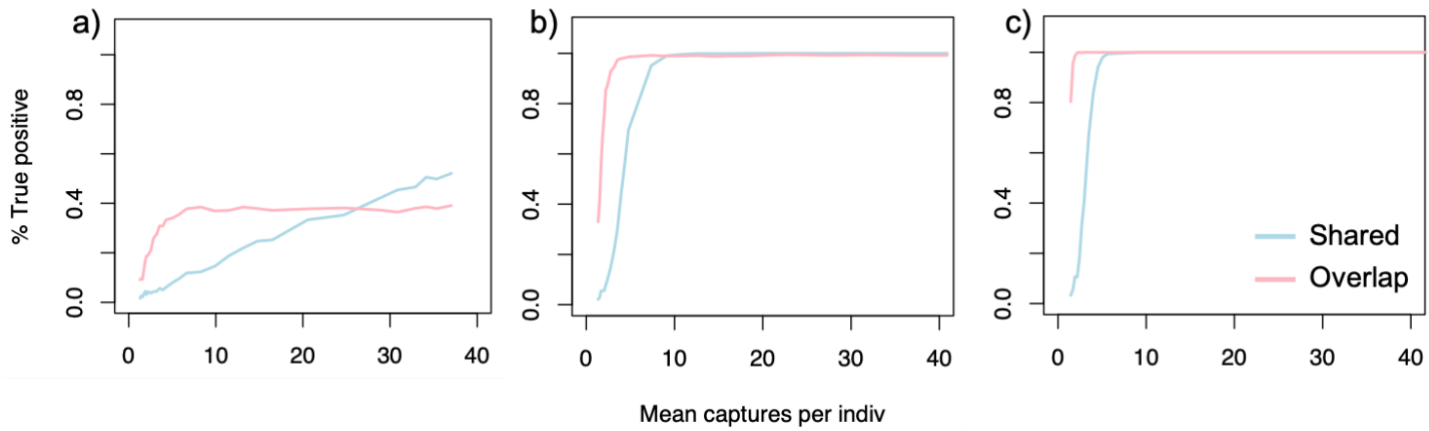

Fig. 4 Performance of observed networks with varying sampling effort on a 10x10 trapping grid, as measured by the power of observed networks to detect an effect present in real networks. Percentage of true positives shown on y-axis, and mean of mean captures per individual shown on x-axis. Repeated for simulated real networks with varying effect sizes (a) an effect size equal to that in our empirical dataset, (b) effect size 4 times greater, (c) effect size 9 times greater.

Performance of observed networks with varying effect size

Only the power of the observed networks changes as a result of varying effect size.

Power

As the effect size increases, there is a corresponding increase in the power of the shared trap network, for a given sampling effort. For example, at 4–5 captures per individual, there is only a 6% chance of detecting a true positive if the effect size is equivalent to that found in our empirical data, but a 56% chance if it is 4 x greater and a 94% chance if it is 9 times greater. The overlap network shows broadly the same pattern. At 4–5 captures per individual, there is a 33% chance of detecting a true positive if the effect size is equivalent to that found in our empirical data, 98% chance if it is 4 times greater and 100% chance if it is 9 times greater.

Changing trapping grid density

Observed networks change in all three metrics of quality as a result of varying trapping grid density. Grid density also changes the number of individuals present in both observed networks when sampling effort is low. For example at 1–2 captures per individual, on a 19x19 grid, a mean of 77 individuals (out of a total population of 100 individuals) are present in the observed networks (compared to 31 on a 10x10 grid; see above). However, all individuals in the population are present in the observed networks from 20 captures per individual, regardless of grid density (Fig. 5c).

Correlation

A higher density grid leads to a weaker correlation between the shared trap network and the real network, for a given sampling effort (Fig. 5a). For example, at 4–5 captures per individual, the Mantel correlation coefficient between the shared trap network and the real network is 0.58 on a 10x10 grid, and 0.38 on a 19x19 grid. The shared trap network reaches a

similar maximum correlation coefficient of 0.85 on a 19x19 grid (compared to 0.87 on a 10x10 grid; see above). The correlation between the overlap network and real network differs very little between the 10x10 and 19x19 grid (Fig. 5a). At a mean of 4–5 captures per individual, the correlation coefficient between the overlap network and the real network is 0.80 on a 10x10 grid, and 0.79 on a 19x19 grid. The overlap network also reaches a similar maximum correlation coefficient of 0.90 on a 19x19 grid (compared to 0.93 on a 10x10 grid; see above).

Accuracy

A higher density grid leads to a less accurate shared trap network for a given sampling effort (Fig. 5b). For example, at 4–5 captures per individual, the absolute difference in edge weights to the real network is 857 on a 10x10 grid, and 938 (1.1 times greater) on a 19x19 grid. The minimum sum of errors for the shared trap network is also higher on a 19x19 grid (716) compared to a 10x10 grid (568; see above). The overlap network shows broadly the same pattern (Fig. 5b). For example, at 4–5 captures per individual, the absolute difference in edge weights of the overlap network to the real network is 665 on a 10x10 grid, and 758 (1.1 times greater) on a 19x19 grid. The minimum sum of errors for the overlap network is also higher on a 19x19 grid (722) compared to a 10x10 grid (552; see above).

Power

A higher density grid leads to a slightly less powerful shared trap network for a given sampling effort (Fig. 5d). For example, at 4–5 captures per individual, the chance of detecting a true positive is 6% on a 10x10 grid, and 3% on a 19x19 grid. Conversely, the power of the overlap network increases with grid density (Fig. 5d). At a mean of 4–5 captures per

individual, the chance of detecting a true positive is 33% on a 10x10 grid, and 59% on a 19x19 grid.

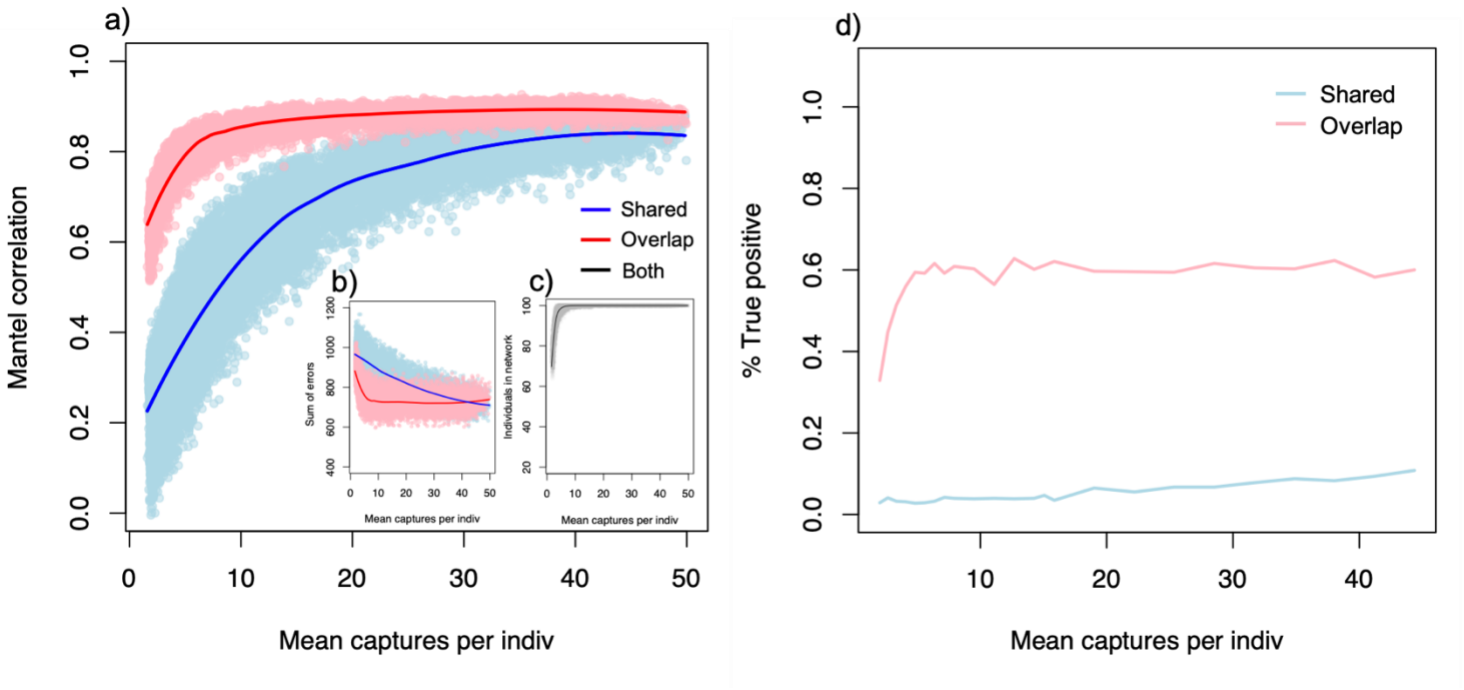

Fig. 5 Performance of observed networks with varying sampling effort on a 19x19 trapping grid, as measured by (a) Mantel correlation between observed and real networks, (b) Absolute difference in edge weights between observed and real networks, (c) Number of individuals in observed networks, (d) Percentage of true positives. LOESS smoother added to (a–c) to aid visual interpretation.

DISCUSSION

In this study, we quantify the robustness of shared space use networks to different sampling regimes. In doing so, we provide much needed guidance for informing the choice of sampling regime when designing studies to accurately quantify space sharing. Using an empirial dataset, we also demonstrate the utility of modelling the space-use behaviour of large numbers of animals on the basis that they have a core and a periphery to their home range.

Finaly, we use these insights to develop a new method for generating shared space use networks. We show that networks generated using this overlap method are generally more strongly correlated with the real network, differ less from the real network, and are more powerful to detect effects present in the real network than networks generated from observed space-sharing events only.

Our overlap method works particularly well when sampling effort is low and provides the potential to generate meaningful networks even from sparse point-based observations of individuals. While standard methods are restrictive, by relying only on joint observations at a trap, and sometimes imposing a temporal threshold within which the observation of two individuals in the same place must occur, our method pools data among individuals to arrive at a more general estimate of home range profile. In doing so, our method accounts for imperfect and heterogeneous observations (as in e.g. Gimenez et al., 2019). Using these general profiles, we then calculate the extent of two individuals' home range overlap, as a function of their observed centroids, to estimate their overlap in space. Our results confirm that this approach results in more accurate and more representative networks than existing methods. We also model sex-difference in home range, based on our empirical data. However, classes of individuals which differ in their space use will vary between systems (e.g. Godsall, Coulson, & Malo, 2014), and prior knowledge will be necessary to identify these classes (if any). Our results further show that our overlap method is more powerful at extracting simulated biological effects.

An important, and perhaps unexpected, finding is that denser grids of data-collection traps can make networks based on joint observations at traps (at least when sampling effort is below 50 captures per individual) less strongly correlated with the real network, less accurate

and less powerful. However, this result makes sense when considering sampling stochasticity. If there are more available traps, then it is less likely that two individuals, which overlap in space, will be trapped in *exactly* the same trap (unless the number of captures is very high and in no way limiting). Subsequently, constructing networks dependent on these space-sharing events, as is the case for the standard method, reduces the correlation between home range overlap and shared detections at traps. By contrast, we show that, in terms of correlation and power, our overlap method performs equally well, or better, on a denser grid. This is because finer-scaled grids provide better estimations of overall space use. Both observed networks are less accurate on a denser grid, despite including a higher proportion of the total population, suggesting that further exploration of sampling design decisions might be needed for studies aimed at accurately estimating edge weights (as opposed to investigating general patterns, such as differences in connectedness among two classes of individuals).

It is important to note that networks generated using the overlap method do not always perform better than those generated using the standard method. We show here that a shared trap network is more accurate and more powerful at detecting true positives when the effect size is low and the sampling effort is high. In other words, if many space-sharing events are observed, then a network based on these alone deviates less from the real network and is more likely to detect a subtle effect present in the real network, relative to the process of pooling data and generating population-wide home range profiles. This pattern is likely to be driven by noise in the process of estimating the centroids of individuals, which can be susceptible to outlying observations. However, we note that the performance of shared trap networks may also depend on the biological system, or the ecological conditions it experiences (as in Perkins, Cagnacci, Stradiotto, Arnoldi, & Hudson, 2009). Our study is

based on empirical data on *M. agrestis* sampled during the breeding season. During this time field voles maintain relatively fixed home ranges which can be estimated with some certainty (Myllymaki, 1977; Niethammer & Krapp, 1982). However, this is likely to be problematic if individuals are highly mobile, resulting in constantly-shifting home ranges. We would expect a shared trap network to be better suited to such a dynamic system. Our study population also inhabits a relatively homogenous landscape, in the form of grassy clear-cuts within a coniferous forest. As a result, individuals are not expected to vary a great deal in the size of their home range. However, landscape features, such as hills, in a more heterogeneous landscape could result in more variability in home range size among individuals, making it difficult to quantify an 'average' home range. A shared trap network could again be better suited to such a system, though this warrants some further investigation (e.g. if both classes are equally likely to be present in all parts of the habitat, then local heterogeneity may not matter). Finally, as previously mentioned, our method is unnecessary for studies in which individuals can be regularly observed or recaptured, so that all (or most) shared space use events are recorded (e.g. using automatic PIT-tag readers at the entrance of burrows).

Shared space use networks are, and will continue to be, widely used to shed light on various biological processes. For example, individuals who share more space may be more likely to compete for resources. Many parasites and pathogens are also transmitted through the environment, and so knowing who shares space with whom, can tell us something about who is likely to transmit infection to whom via this indirect route (VanderWaal et al., 2013). It is also true that shared space use, or proximity, is a prerequisite for interaction (Farine, 2015), but whether or not individuals that share space do indeed associate, or interact, and thus how far point-based observations can be used to draw meaningful inferences, will depend on the biology of the system (Fig. 6). For example, "behaviour [can] play an important role and

some animals may actively avoid each other" (Davis et al., 2014). It is therefore important to take care when making biological inferences from network data.

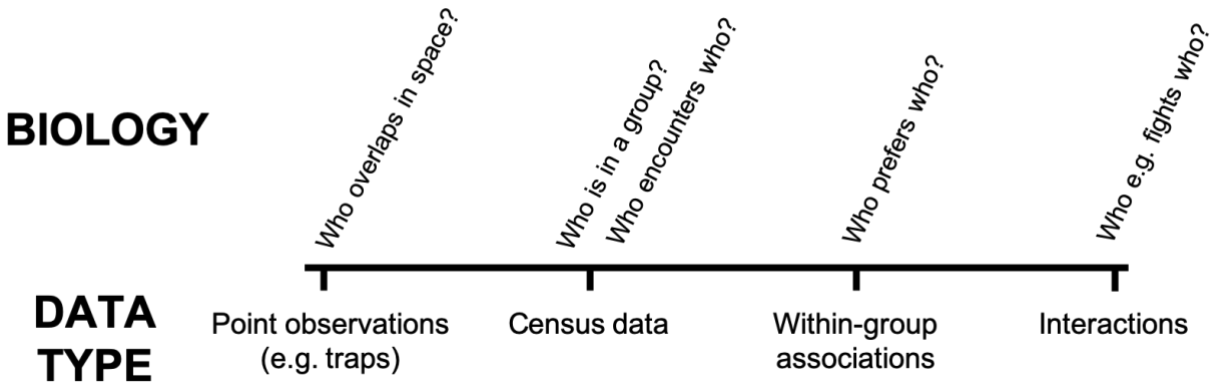

Fig. 6. Schematic summarising link between network data and biological processes.

Our method provides a novel opportunity to generate meaningful space-sharing networks, and if appropriate, to make inferences from shared space use for higher-level types of questions, even from sparse point-based observations of individuals. It therefore unlocks the potential of these data, still the most common form of data available for many smaller or more cryptic species, to address a range of key questions in ecology and evolution.

Acknowledgements

KMW was supported by a Natural Environment Research Council (NERC) research grant NE/L013452/1 (awarded to Steve Paterson, University of Liverpool) and a Johnston Postdoctoral Development Award from the University of Liverpool. DRF received funding from the Max Planck Society, the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 850859), and the DFG Centre of Excellence 2117 'Centre for the Advanced Study of Collective Behaviour' (ID: 422037984). The empirical data used in this study were collected as part of

the NERC research grant NE/E015131/1 (awarded to Mike Begon, University of Liverpool). We would like to thank the many individuals involved in collecting this data and the Forestry Commission for access to the study sites.

References

Berkvens, R., Olivares, I. H., Mercelis, S., Kirkpatrick, L., & Weyn, M. (2019). Contact detection for social networking of small animals. In *Advances on P2P, Parallel, Grid, Cloud and Internet Computing*. Taichung, Taiwan.

Borowski, Z. (2003). Habitat selection and home range size of field voles *Microtus agrestis* in Słowiński National Park, Poland. *Acta Theriologica*, *48*, 325–333.

Calhoun, J. B., & Casby, J. U. (1958). Calculation of home range and a density of small mammals. *United States Public Health Monograph*, *55*, 1–24.

Cantor, M., Maldonado-Chaparro, A. A., Beck, K., Carter, G. G., He, P., Hillemann, F., … Farine, D. (2019). Animal social networks: revealing the causes and implications of social structure in ecology and evolution. *EcoEvoRxiv*. doi:10.32942/osf.io/m62gb

Craft, M. E., & Caillaud, D. (2011). Network models: an underutilized tool in wildlife epidemiology? *Interdisciplinary Perspectives on Infectious Diseases*, *2011*, 676949.

Davis, G. H., Crofoot, M. C., & Farine, D. R. (2018). Estimating the robustness and uncertainty of animal social networks using different observational methods. *Animal Behaviour*, *141*, 29–44.

Davis, S., Abbasi, B., Shah, S., Telfer, S., Begon, M., & Davis, S. (2014). Spatial analyses of wildlife contact networks. *Journal of the Royal Society Interface, 12*, 20141004.

Douglas, M., Ji, W., & Clout, M. N. (2006). MateID: design and testing of a novel device for recording contacts between free-ranging animals. *Wildlife Society Bulletin*, *34*, 203–207.

Farine, D. R. (2015). Proximity as a proxy for interactions: issues of scale in social network

analysis. *Animal Behaviour*, *104*, e1–e5.

Farine, D. R. (2017). A guide to null models for animal social network analysis. *Methods in Ecology and Evolution*, *8*, 1309–1320.

Farine, D. R. (2018). When to choose dynamic vs. static social network analysis. *Journal of Animal Ecology*, *87*, 128–138.

Farine, D. R., & Strandburg-Peshkin, A. (2015). Estimating uncertainty and reliability of social network data using Bayesian inference. *Royal Society Open Science*, *2*, 150367.

Farine, D. R., & Whitehead, H. (2015). Constructing, conducting, and interpreting animal social network analysis. *The Journal of Animal Ecology*, *84*, 1144–1163.

Gimenez, O., Mansilla, L., Klaich, M. J., Coscarella, M. A., Pedraza, S. N., & Crespo, E. A. (2019). Inferring animal social networks with imperfect detection. *Ecological Modelling*, *401*, 69–74.

Godsall, B., Coulson, T., & Malo, A. F. (2014). From physiology to space use: energy reserves and androgenization explain home-range size variation in a woodland rodent. *Journal of Animal Ecology*, *83*, 126–135.

Grear, D. A., Luong, L. T., & Hudson, P. J. (2013). Network transmission inference: host behavior and parasite life cycle make social networks meaningful in disease ecology. *Ecological Applications*, *23*, 1906–1914.

Grear, D. A., Perkins, S. E., & Hudson, P. J. (2009). Does elevated testosterone result in increased exposure and transmission of parasites? *Ecology Letters*, *12*, 528–537.

Hoppitt, W. J. E., & Farine, D. R. (2018). Association indices for quantifying social relationships: how to deal with missing observations of individuals or groups. *Animal Behaviour*, *136*, 227–238.

Jackson, J. A., Hall, A. J., Friberg, I. M., Ralli, C., Lowe, A., Zawadzka, M., … Begon, M. (2014). An immunological marker of tolerance to infection in wild rodents. *PLoS*

Biology, *12*, e1001901.

Myllymaki, A. (1977). Intraspecific competition and home range dynamics in the field vole *Microtus agrestis*. *Oikos*, *29*, 553–569.

Niethammer, J., & Krapp, F. (1982). *Microtus agrestis* (Linnaeus, 1761). In J. Niethammer & F. Krapp (Eds.), *Handbuch der saugetierkunde Europas, Volume 2/I, Nagetiere II.* (pp. 350–373). Akademische Verlag, Wiesbaden, Germany.

Perkins, S. E., Ferrari, M. F., & Hudson, P. J. (2008). The effects of social structure and sex-biased transmission on macroparasite infection. *Parasitology*, *135*, 1561–1569.

Perkins, Sarah E., Cagnacci, F., Stradiotto, A., Arnoldi, D., & Hudson, P. J. (2009). Comparison of social networks derived from ecological data: implications for inferring infectious disease dynamics. *Journal of Animal Ecology*, *78*, 1015–1022.

Porphyre, T., Stevenson, M., Jackson, R., & McKenzie, J. (2008). Influence of contact heterogeneity on TB reproduction ratio R_0 in a free-living brushtail possum *Trichosurus vulpecula* population. *Veterinary Research*, *39*.

Rutz, C., Burns, Z. T., James, R., Ismar, S. M. H., Burt, J., Otis, B., … St Clair, J. J. H. (2012). Automated mapping of social networks in wild birds. *Current Biology*, *22*, R669–R671.

Ryder, T. B., Horton, B. M., van den Tillaart, M., Morales, J. D. D., & Moore, I. T. (2012). Proximity data-loggers increase the quantity and quality of social network data. *Biology Letters*, *8*, 917–920.

Schoener, T. W. (1981). An empirically based estimate of home range. *Theoretical Population Biology*, *20*, 281–325.

Smith, J. E., Gamboa, D. A., Spencer, J. M., Travenick, S. J., Ortiz, C. A., Hunter, R. D., & Sih, A. (2018). Split between two worlds: automated sensing reveals links between above- and belowground social networks in a free-living mammal. *Philosophical*

Transactions of the Royal Society B: Biological Sciences, *373*, 20170249.

Tompkins, D. M., Dunn, A. M., Smith, M. J., & Telfer, S. (2011). Wildlife diseases: from individuals to ecosystems. *Journal of Animal Ecology*, *80*, 19–38.

VanderWaal, K. L., Atwill, E. R., Hooper, S., Buckle, K., & McCowan, B. (2013). Network structure and prevalence of *Cryptosporidium* in Belding's ground squirrels. *Behavioral Ecology and Sociobiology*, *67*, 1951–1959.