The role of population density and structure in Chiropteran pathogen richness

Tim C. D. Lucas

A thesis submitted in partial fulfilment of the requirements for the degree of:

Doctor of Philosophy of

University College London

2016

Primary supervisor:
Prof. Kate E. Jones

Secondary supervisor:
Dr Hilde Herbots

I, Tim C. D. Lucas, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Pathogens make up a huge proportion of global diversity and their role in disease strongly affects human disease, economics and development as well as having an important ecological role. However, the factors that control the number of pathogen species are poorly understood. In this thesis I examine the role population structure and density in maintaining these high levels of diversity in the face of interpathogen competition. Throughout the thesis I focus on bats as a case study for more general disease processes as they have highly varied social structures and have emerged as important reservoires for zoonotic viruses such as Ebola, SARS, Hendra and Nipah.

In Chapter 2

Acknowledgements

Acknowledge all the things!

Contents

List of Figures

List of Tables

1 Introductory Material

1.1 Pathogen richness and the impacts of zoonotic diseases

The diversity of pathogens is huge and largely unknown (Poulin 2014). Recent large studies have found tens (Anthony et al. 2013) or even hundreds of virus species in a single host species (Anthony et al. 2015). This suggests the number of mammalian viruses globally is of the order of hundreds of thousands of virus species (Anthony et al. 2013) while only 3,000 species of virus, across all host groups, are currently described (King et al. 2011). Recent large databases include nearly 2,000 pathogens from approximately 400 wild animal hosts (Wardeh et al. 2015). Given that there are nearly 4,000 named mammal species (Wilson & Reeder 2005), the undiscovered diversity of pathogens is likely huge.

In general we can expect competition between pathogens. Competition between pathogens can occur by different mechanisms: immunological mechanisms such as cross-immunity or shared immune respone (Fenton & Perkins 2010) and ecological mechanisms such as removal of susceptable hosts by death (Rohani et al. 2003) or competition for internal host resources (Griffiths et al. 2014). Like ecological systems this competition leads us to expect competitive exclusions (Ackleh & Allen 2003, Ackleh & Salceanu 2014, Bremermann & Thieme 1989, Martcheva & Li 2013, Turner & Garnett 2002) so the diversity of parasites needs an explanation.

This diversity of pathogens presents a risk to human health. 60% of newly emerged disease are zoonotic (acquired from animals) with wild animals being the predominant source (Jones et al. 2008, Taylor et al. 2001, Woolhouse & Gowtage-Sequeria 2006). Zoonotic diseases can be extremely virulent: viruses such as Nipah (Luby et al. 2009), Ebola (Lefebvre et al. 2014) having case fatality rates over 50%. Furthermore these pathogens can have large economic costs (e.g. SARS is estimated to have cost $40 billion (Knobler et al. 2004)). In particular these impacts can have huge effects on developing economies. For example the 2014 ebola epidemic caused both Liberia and Guinea to fall from positive to negative per capita growth rates (World Bank 2014, World Bank 2015) while death rates per 1,000 people living with AIDS are up to ten times higher in developing countries than in Europe and North America (Granich et al. 2015).

Surveillance of zoonotic diseases is crucial to mitigate these health impacts. In particular we want to categorise diseases before they spillover into humans (SARS was not identified until months into the pandemic for example (Drosten

et al. 2003)) and we want to anticipate outbreaks. If we know there is an increased number of infections in host species, increased numbers of a species that is a known reservoir of a high risk zoonotic disease, or increased contacts between humans and a pathogen reservoir, we can prepare for a potential outbreak in that area.

However, funds for zoonotic surveillance are limited and so efforts must be optimised. Knowing which species are likely to have many pathogens allows us to sample and identify potentially zoonotic viruses efficiently. Suggested factors that might control pathogen richness include individual traits (body mass (Arneberg 2002, Kamiya et al. 2014, Poulin 1995) and longevity (Ezenwa et al. 2006, Nunn et al. 2003)) as well as environmental factors such as latitude (Kamiya et al. 2014, Poulin 2010). Furthermore, population level traits that affect the dynamics of disease spread have also been studied (animal density (Arneberg 2002, Kamiya et al. 2014, Nunn et al. 2003), sociality (Altizer et al. 2003, Bordes et al. 2007, Ezenwa et al. 2006, Vitone et al. 2004), population structure (Gay et al. 2014, Maganga et al. 2014, Nunes et al. 2006, Turmelle & Olival 2009) and species range size (Kamiya et al. 2014, Nunn et al. 2003)). This understanding provides a basis for predicting which species will have high pathogen richness and should be prioritised for sampling and surveillance. Furthermore, given a good mechanistic understanding of how pathogen richness is created and maintained we can start to predict how pathogen richness, and therefore zoonotic disease risk, will respond to global change.

1.2 Bats as reservoires of zoonotic diseases

In recent decades bats have been implicated in a number of high profile zoonotic outbreaks including Nipah (Field et al. 2001, Halpin et al. 2011), Ebola (Leroy et al. 2005), SARS (Li et al. 2005) and Hendra (Field et al. 2001). This has lead to much research on whether bats are a particular source of zoonotic disease (Luis et al. 2013, Olival et al. 2015, Wang et al. 2011) and examinations of factors, such as flight, social living and longevity, that might predispose them to being reservoires of zoonotic viruses (Calisher et al. 2006, Dobson 2005, Kuzmin et al. 2011, O'Shea et al. 2014, Racey 2015). Given that bats are the second largest order of mammals (Wilson & Reeder 2005), we may expect them to be the source of many viruses simply through weight of numbers (Luis et al. 2013). The broad conclusions are that while bat do host more zoonotic viruses than other groups (Luis et

al. 2013) they do not host more virus species in general (Olival et al. 2015). Questions remain as to why bat viruses have a tendancy to have such high zoonotic potential.

Many factors of bat populations make them epidemiologically interesting. Firstly, the population density of most bat species is largely unknown. As they are small, noctornal and difficult to identify in flight, estimating there density is incredibly difficult without distruptive and time consuming roost surveys. As this parameter is tightly linked to pathogen richness (Kamiya et al. 2014) and central to epidemiological models (Anderson & May 1979, May & Anderson 1979) these leaves large gaps in our understanding of disease processes in this group. Secondly they have highly varied and sometimes complex social structures (Kerth 2008). While some species are largely solitary or live in very small groups (e.g. *Lasiurus borealis* (Shump & Shump 1982)) some species live in colonies of millions of individuals (e.g. *Pteropus scapulatus* (Birt et al. 2008)). These groups can be very stable (Kerth et al. 2011, McCracken & Bradbury 1981). Further complexity arises due to their propensity for seasonal migration (Cryan et al. 2014, Fleming et al. 2003, Richter & Cumming 2008) and seasonally changing social organisation including maternity roosts, hibernation roosts and swarming sites (Kerth 2008). Finally, their ability to fly means populations can be well mixed across large distances (Peel et al. 2013, Petit & Mayer 1999) but this is highly variable with some species having limited dispersal (Wilmer et al. 1994).

1.3 Population density and structure as factors that influence pathogen richness

The role of population density and abundance in disease dynamics is well established (Anderson & May 1979, Heesterbeek 2002, Lloyd-Smith et al. 2005, May & Anderson 1979). Broadly, larger populations can maintain diseases more easily, due to having a larger pool of susceptible individuals (individuals without acquired immunity) and having a greater number of new susceptible individuals enter the population by birth or immigration (Anderson & May 1979, May & Anderson 1979). High density populations are expected to have a greater number of contacts between individuals and so promote disease spread (though there is much discussion on when the number of contacts might scale independantly of density (McCallum et al. 2001)).

There is a large literature on the role of population structure on disease dynamics (see review by Pastor-Satorras et al. (2015)) driven by applications to computer viruses (Pastor-Satorras & Vespignani 2001) and the social spread of information (Goffman & Newill 1964), as well as health applications. In particular work has concentrated on how structure affects the invasion threshold, R_0 (Barthélemy et al. 2010, Colizza & Vespignani 2007, May & Lloyd 2001, Pastor-Satorras & Vespignani 2001, Wu et al. 2013). This value combines relevent parameters to yield a threshold above which a disease is expected to infect a significant proportion of the population (Anderson & May 1979, May & Anderson 1979). Below the threshold, only small outbreaks that quickly die out are expected.

However, the majority of theoretical work considers single pathogens. Single pathogen models suggest whether a pathogen can spread and persist in a population, ignoring all other pathogens. A number of studies examine competing epidemics — when two pathogens spread at the same time, which pathogen infects more individuals? — and have found increasing population structure reduces dominance of one strain (Poletto et al. 2013, Poletto et al. 2015, van de Bovenkamp et al. 2014). However, this again tells us little about how pathogen communities form and what factors control total pathogen richness. Many fewer papers explicitely study the longer term competition between two or more pathogens. Those that do commonly find that competitive exclusion is likely (Ackleh & Allen 2003, Ackleh & Salceanu 2014, Bremermann & Thieme 1989, Castillo-Chavez et al. 1995, Martcheva & Li 2013, Turner & Garnett 2002). Mechanisms that have been shown to allow pathogen coexistence include superinfection (Li et al. 2010, May & Nowak 1994), density-dependant deaths (Ackleh & Allen 2003, Kirupaharan & Allen 2004) and differing transmission routes (Allen et al. 2003).

The specific role of density on the ability of pathogens to coexist has not been theoretically studied though it is commonly foudn to promote pathogen richness in comparative empirical studies (Arneberg 2002, Kamiya et al. 2014, Nunn et al. 2003). The few papers that have directly studied how coexistence of pathogens responds to population structure have found population structure can allow pathogens to coexist when competitive exclusion would occur in a fully mixed population (Allen et al. 2004, Nunes et al. 2006, Qiu et al. 2013). Furthermore, genetic diversity has shown to be maximised at intermediate levels of population structure (Campos & Gordo 2006). The roles of population structure and social group size have been examined in comparative studies (Altizer et al. 2003, Bordes et al. 2007, Ezenwa et al. 2006, Gay et al. 2014, Maganga et al. 2014, Rifkin et al. 2012,

Turmelle & Olival 2009, Vitone et al. 2004). There is much disagreement between these studies with population structure being shown to promote (Maganga et al. 2014, Turmelle & Olival 2009) and inhibit pathogen richness (Gay et al. 2014) and similarly group size being shown to promote (Bordes et al. 2007, Rifkin et al. 2012) and inhibit (Ezenwa et al. 2006) richness. While increased group size should generally decrease population structure the literature is rarely clear on how these variables relate.

1.4 Thesis overview

In this thesis I examine the role of population structure and density on pathogen richness. I use bats as a case study throughout due to their interesting social structure and importance as zoonotic reservoires. I combine empirical, comparative studies with simulation models. This allows me to study specific mechanisms while linking my theoretical insights to real world tests.

First, in Chapter 2, I empirically test the hypothesis that population structure is associated with pathogen richness in wild bat populations. I used two measurements of population structure — the number of subspecies and gene flow — and a larger dataset than previous studies. For both measurements I found that bat species with more structured populations have more known viruses. This relationship is robust to controlling for study bias and phylogenetic nonindependance. I also tested for a relationship between body mass and range size, finding strong support for larger bodied bats carrying more viruses and mixed support for range size promoting pathogen richness.

In Chapter 3 I examine one specific mechanism by which population structure may promote increased pathogen richness. I tested whether increased population structure can allow newly evolved pathogen strains to invade and persist more easily. I modelled bat populations as individual-based, stochastic metapopulations and examine the competition dynamics of two identical pathogen strains. I tested two factors related to population structure: dispersal rate and the number of links between subpopulations. I found that, at a number of different transmission rates, neither of these factors altered the probability of newly evolved pathogens invading and persisting in the population.

Next I examine the relationships between a number of elements of population structure (Chapter 4). I clarified the interdependances between range size, and population size and density. I also noted that population size can be decomposed

into colony size and colony density (the number of colonies per unit area). Using the same model as Chapter 3, I then tested which of these factors are most important in promoting pathogen richness. Specifically I tested which factor most strongly promotes the invasion and establishment of newly evolved pathogens. I found that population size is more important than population density and that colony size is the important component of population size.

Given the importance of population size on pathogen richness it is important to have good estimates of these for wild bat populations. However, there are currently very few measurements of bat abundance due to their small size, nocturnal habit and difficulties in identification. Therefore I aimed to develop a method for estimating bat abundance from acoustic data, specifically data collected by the iBats project (Jones et al. 2011). In Chapter 5 I present a generally applicable method — based on random encounter models (Rowcliffe et al. 2008, Yapp 1956) — for estimating abundance of animal populations using camera traps or acoustic deteectors. I used spatial simulations to test the method for biases and to assess it's precision. I found that the method is unbiased and precise as long as a reasonable amount of data is collected.

2 The role of population structure in pathogen diversity in wild bat populations

This work was conducted in collaboration with Kate Jones and Hilde Wilkinson-Herbots

2.1 Abstract

It is still unclear what factors determine the number of pathogens a wild species carries. But once understood, these factors could provide a way to prioritise surveillance of wild populations for zoonotic disease and make predictions as to how pathogen richness will respond to global change.

The pattern of contacts between individuals (i.e. population structure) has long been known to strongly affect epidemic processes. Theory suggests that population structure can promote pathogen richness while the ecological literature generally assumes it will decrease richness.

Previous studies in wild bat populations have had contradictory results and the different measures of population structure have different shortcomings. There is therefore a need for more robust comparative studies testing for a relationship between pathogen richness and population structure.

Here I used comparative data to test whether population structure influences pathogen richness using bats as a case study as they have been associated with a number of important, recent zoonotic outbreaks. Unlike previous studies I used two measures of population structure: a novel measure, number of subspecies, and a more careful application of genetic measures which have been used previously. I find that both of these measures are positively associated with pathogen richness and are probably in the best supported model.

My results add more robust support to the hypothesis that pathogen structure promotes pathogen richness in bats and lends clarity to the contradictory results previously published. The results support predictions from theory. They contradict the assumption commonly made in the ecological literature that factors that increase R_0 should increase pathogen richness implying that competitive processes amongst pathogens are stronger than previously thought.

Although my analysis implies that population structure does promote pathogen richness in bats, the weakness of the relationship and the difficulty in obtaining some measurements means this is probably not a usefully predictive factor on it's own for optimising zoonotic surveillance. However, the relationship has implications for global change, implying that increased habitat fragmentation might promote greater viral richness in bats.

2.2 Introduction

Zoonotic pathogens make up the majority of newly emerging diseases and have profound consequences for public health, economics and international development (Jones et al. 2008, Smith et al. 2014, World Bank 2014). A better understanding of the factors that control which wild host species are potential reservoires of zoonotic diseases would allow us to effectively optimize zoonotic disease surveillance and anticipate how the risks of disease spillover might change with global change. The chance that a host species will be the source of an outbreak depends on a number of factors including it's proximity and interactions with humans and the prevalence and diversity of pathogens it carries. Our understanding of the factors that control the diversity of pathogens in a wild species is still poor.

Single pathogen models show that increasing population structure simply slows disease spread and makes establishment less likely (Colizza & Vespignani 2007, Vespignani 2008). In the ecological literature this is often taken as predicting that increased population structure will decrease pathogen richness (Altizer et al. 2003, Morand 2000, Nunn et al. 2003, Poulin 2014, Poulin & Morand 2000). However, models of competition between multiple pathogens show that in unstructured populations a competitive exclusion process occurs but that splitting the population into two patches allows coexistence (Allen et al. 2004, Nunes et al. 2006, Qiu et al. 2013).

Three studies have used comparative data to test for an association between population structure and viral richness. A study on 15 African bats found a positive relationship between distribution fragmentation and viral richness (Maganga et al. 2014) while a study on 20 South-East Asian bats found the opposite relationship (Gay et al. 2014). A global study on 33 bats found a positive relationship between F_{ST} — a measure of genetic structure — and viral richness (Turmelle & Olival 2009). However, this study included measures using mtDNA which only measures female dispersal which may have biased the results many bat species show female philopatry (Hulva et al. 2010, Kerth et al. 2002). Furthermore, this study used measures of F_{ST} irrespective of the study scale with studies covering from tens (McCracken & Bradbury 1981) to thousands (Petit & Mayer 1999) of kilometers. As isolation by distance has been shown in a number of bat species (Burland et al. 1999, Hulva et al. 2010, O'Donnell et al. 2015, Vonhof et al. 2015) this could bias results further. Finally, when a global F_{ST} value is not given they used the mean of all pairwise F_{ST} between sites. This is not correct as from global F_{ST} we expect migration rates of $M = (1-F_{ST})/8F_{ST}$ while from F_{ST} between pairs

of populations we expect migration rates of $M = (1 - F_{ST})/16F_{ST}$ where M is the absolute number of diploid inviduals dispersing per generation (Slatkin 1995). To use studies that only present pairwise F_{ST} values the raw data would have to be gathered and global F_{ST} calculated from those. As it is in fact the movement of individuals that is epidemiologically relavent, using these studies is probably not correct without attempting to correct for these difference. Studies on single pathogens have also shown that space can allow persistence where a well mixed population ould experience a single, large epidemic followed by pathogen extinction (Blackwood et al. 2013, Plowright et al. 2011, Pons-Salort et al. 2014).

A number of measurments of population structure have been used in the literature and each has it's own shortcomings. In particular, the better, more direct measurements tend to be very work intensive which consequently means data is available for few species.

The ideal metric of population structure is direct measurement of dispersal rates and distance. These are incredibly difficult to obtain, especially over large scales. Due to white nose syndrom, some very large mark-recapture studies have been conducted, but recapture rates are low (). Further, these large studies have been in species that live in a few large colonies, so recapture rates should be higher than in less social species. In practise, direct measurements are not practical for comparative analysis due to the lack of data and inconsistency in data collection methods.

As direct measurement of dispersal are difficult, genetic data is often used. Measurement such as F_{ST} are used to calculate migration. There are strong model assumptions under the conversion from F_{ST} to migration. However, the main issue with this measure is the effort required for each study and the subsequent low number of measurements. Further, there are differences in the scales of the studies and the genetic regions being sequenced. This differences should not be ignored.

For a population to evolve distinct phenotypic or genetic traits, such that they can be classed as a subspecies, there must be limited migration between populations. The number of subspecies a species has therefore reflects the level of population structure in that species. The value of this measurement is available for every bat species. However, it is likely biased, with well studied species being likely to have more recognised subspecies. Further, this is a very course measure and it is important to consider whether it is measuring migration at a timescale and rate that is epidemiologically relevant.

For both measures from F_{ST} and the number of subspecies it is useful to consider the rates of animal movement that are being measured. Rates of migration estimated from F_{ST} tend to be between 1 and 100 individuals per generation dispersing across all subpopulations.

The final measurement that has been used is derived from the shape of the species' range, typically from IUCN (IUCN 2010) maps. The ratio between the perimeter of the range and the area (or similar values) are calculated. Range maps are very course for many species. Furthermore there is a potential bias with island living species being given sea based edges where continental species might be assumed to live everywhere in between locations where it is known to live, without considering the different terrestrial habitats in these areas.

There is a lack of studies using multiple measures of population structure and larger datasets to robustly estimate the importance of population structure.

Here I have used two measures of population structure — the number of subspecies and gene flow — to robustly test for an association between population structure and pathogen richness in bats. Furthermore, I used a much larger dataset for one of these analyses, further promoting robustness of results. I found that both measures of population structure are positively associated with viral richness and are likely to be in the best models for describing viral richness. Further, I found that the role of phylogeny is very weak in the models and in the distribution of viral richness amongst taxa.

2.3 Methods

To test for an association between pathogen richness I have performed multivariate regression using a model selection framework to establish whether or not two measures of pathogen richness are likely to be in a 'best model' and therefore important. As species cannot be considered independant due to shared evolutionary history, phylogeny was controlled for in all regressions. A number of other factors that have previously been found to be important were included as additional independant variables: body mass (Gay et al. 2014, Kamiya et al. 2014, Maganga et al. 2014, Turmelle & Olival 2009), range size (Kamiya et al. 2014, Maganga et al. 2014, Turmelle & Olival 2009) and study effort (Gay et al. 2014, Maganga et al. 2014, Turmelle & Olival 2009). This was to attempt to avoid spurious positive results occuring simply due to correlation between pathogen richness and a different, causal factor. Despite commonly being associated with

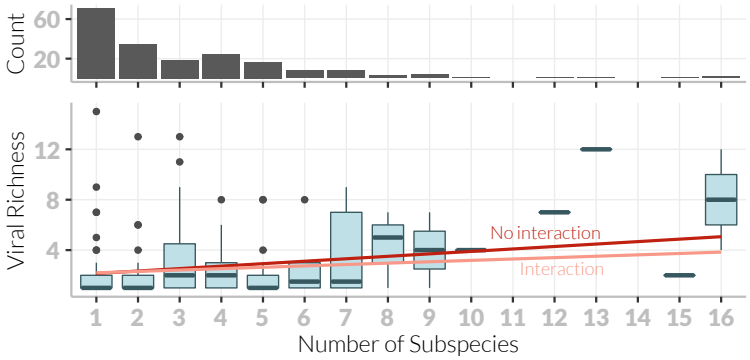


Figure 2.1 Number of virus species against number of subspecies. The top panel shows the univariate distribution of the data with most species having few subspecies. Data within a number of subspecies are plotted as boxplots with the dark bar showing the median, the box showing the interquartile range, vertical lines showing the range and outliers shown as seperate points. Regression lines are from multivariate phylogenetic models with all other independant variables set at their median value. The models shown are those with (pink) and without (red) an interaction between study effort and number of subspecies.

pathogen richness (Arneberg 2002, Kamiya et al. 2014, Nunn et al. 2003), population density is not included in the analysis as there is very little data for bat densities — however Chapter 4 examines the relationship between density and population structure and Chapter 5 presents a method that allows the estimation of density from acoustic surveys. I used both the number of subspecies a bat species has and estimates of gene flow (analysed separately) as measures of population structure. All analyses were run in R (R Development Core Team 2010)

To measure pathogen richness I used data from (Luis et al. 2013). These simply include known infections of a bat species with a pathogen species. Only species with at least one pathogen were included in the analysis. Rows with host species that were not identified to species level were removed. Many viruses were not identified to species level or their specified species names were not in the ICTV virus taxonomy (King et al. 2011). I counted a virus if it was the only virus, for that host species, in the lowest taxonomic level (present in the ICTV taxonomy) identified. That is, if a host carries an unknown Paramyxoviridae virus, then it must carry at least one Paramyxoviridae virus. If a host carries an unknown Paramyxoviridae virus and a known Paramyxoviridae virus, then it is hard to

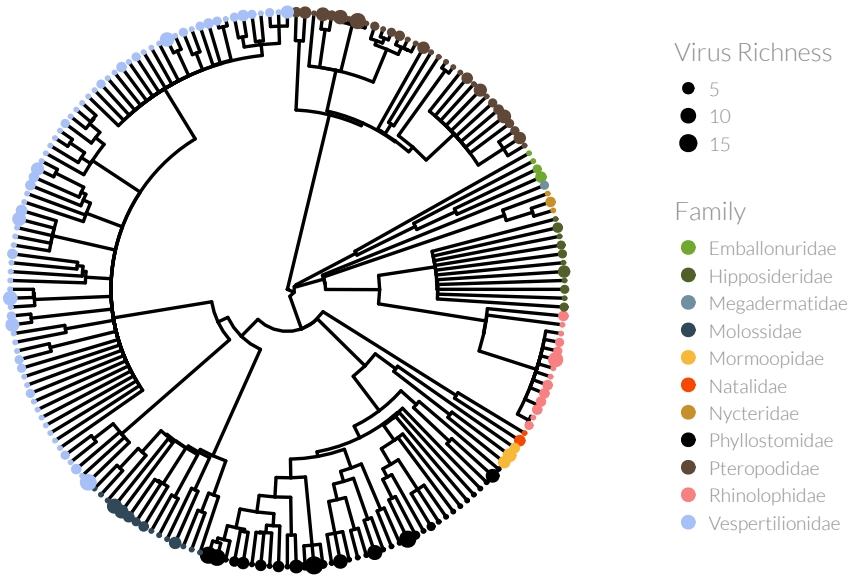


Figure 2.2 Phylogeny from (Bininda-Emonds et al. 2007) pruned to include all species used in either the number of subspecies of gene flow analysis. Dot size shows the number of known viruses for that species and colour shows family.

confirm that the unknown virus is not another record of the known virus. In this case, this would be counted as one virus species.

I used two measures of population structure. Gene flow and the number of subspecies. The number of subspecies was counted using the Wilson and Reeder taxonomy (Wilson & Reeder 2005). Gene flow is calculated from estimates of F_{ST} collated from the literature. Studies are from a wide range of spatial scales, from local (~ 10 km) to continental. As F_{ST} often increases with spatial scale (Burland et al. 1999, Hulva et al. 2010, O'Donnell et al. 2015, Vonhof et al. 2015) I controlled for this by only using data from studies where a large proportion of the species range was studied. I used the ratio of the furthest distance between F_{ST} samples (measured with http://www.distancefromto.net/ if not stated) to the width of the IUCN species range (IUCN 2010) and only used studies if this ratio was greater than 0.2. This is an arbitrary value that was a comprimise between retaining a reasonable number of data points and controlling for the bias in spatial scale. I converted all F_{ST} value to migration using $M = (1 - F_{ST})/8F_{ST}$. This removes the $(0, 1)$ bounds of F_{ST} and is more easily interpretable though the results are unaffected. These two measures of population structure were analysed separately as the number of subspecies has 196 data points while there is only F_{ST} data for 22 bat species. For the subspecies analysis all bat species in Luis et al. (2013) were

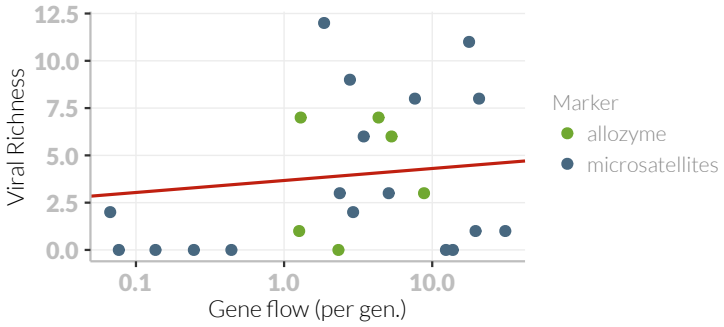


Figure 2.3 Gene flow per generation (on a log scale) against viral richness. The genetic marker used is shown with colour. The red line shows the univariate phylogenetic model.

used (i.e. all species with at least one known virus species). However, for the gene flow analysis, all bat species with suitable F_{ST} estimates were used. As this included some species not present in Luis et al. (2013) this includes some bat species with zero known virus species.

To control for study bias I collected the number of Pubmed and Google Scholar citations for each bat species name including synonyms from ITIS (ITIS 2015) via the taxize package (Chamberlain & Szöcs 2013). The counts were scraped using the rvest package (Wickham 2015). I log transformed these variables as they were strongly right skewed. The log number of citations on Pubmed and Google scholar were highly correlated (pgls: $t = 19.32$, df $= 194$, $p = 0$). As this correlation is strong, the results here are for analyses using only Google Scholar citations.

Measures of body mass are taken from Pantheria (Jones et al. 2009) and primary literature (Aldridge 1987, Arita 1993, Canals et al. 2005, Heaney et al. 2012, Henderson & Broders 2008, Lim & Engstrom 2001, López-Baucells et al. 2014, Ma et al. 2003, Oleksy et al. 2015, Orr & Zuk 2013, Owen et al. 2003, Zhang et al. 2009). *Pipistrellus pygmaeus* was assigned the same mass as *P. pipistrellus* as they are indistinguishable by mass. Body mass measurements were log transformed as they were strongly right skewed. Distribution size was estimated by downloading range maps for all species from IUCN (IUCN 2010) and were also logged due to right skew.

To control for phylogenetic nonindependance I used the best-supported phylogeny from Fritz et al. (2009) (shown in Figure 2.2) which is the supertree from (Bininda-Emonds et al. 2007) with names updated to match the Wilson & Reeder taxonomy (Wilson & Reeder 2005). Phylogenetic manipulation was performed

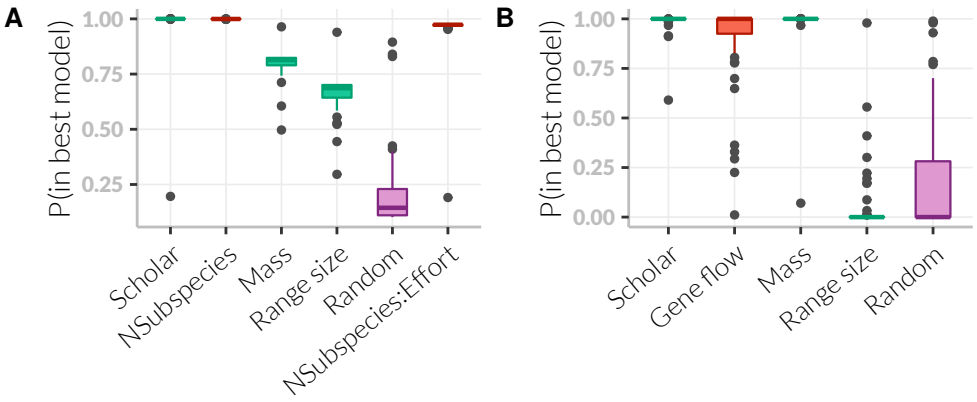


Figure 2.4 Akaika variable weights for both analyses. The probability that each variable will be in the best model if the data were recollected is shown for each of the bootstrap analyses. The purple "Random" box is a uniform random variable used as a null. Population structure (Number of subspecies and Gene flow), shown in red, is likely to be in the best model in both analyses.

using the ape package (Paradis et al. 2004). The importance of the phylogeny on each variable separately (the λ parameter of the variable regressed against an intercept) was estimated and tested against the null of $\lambda = 0$ with log-likelihood ratio tests using caper (Orme et al. 2012). I also performed the analysis using the tree from (Jones et al. 2005) as this has some broad changes with families in different places. However the phylogeny did not affect the analysis.

2.3.1 Statistical analysis

Statistical analysis for both dependant variables was conducted using a information theoretical/model averaging approach (Burnham & Anderson 2002) specifically following (Whittingham et al. 2006, Whittingham et al. 2005). I chose a credible set of models including all combinations of independent variables and a model with just an intercept. In the analysis using the number of subspecies dependant variable I also included an interaction term between study effort and number of subspecies. This interaction was included as I believed *a priori* that this interaction may be present as subspecies in well studied species are more likely to be identified. The interaction was only included in models with both study effort and number of subspecies as individual terms.

I fitted phylogenetic regressions of all models using nlme (Pinheiro et al. 2015). The independant variables were centered and scaled to allow direct comparison of the coefficients (Schielzeth 2010). In each case I simultaneously fitted the λ

parameter as this avoids mispecifying the model (Revell 2010). κ and δ parameters were constrained to one as they are more concerned with when along a branch evolution occurs and because fitting multiple parameters makes interpretation difficult.

To establish the importance of variables I calculated the probability, *Pr*, that each variable would be in the best model if the data were recollected. For each variable the mean of it's coefficient, β, in all models that contained that variable was also calculated to determine the direction and strength of the variables. In the subspecies analysis, this mean of β was calculated for all models, only models with a interaction term and only models without an interaction. As the interaction term greatly affects the estimated value of β, considering these value seperately aids interpretation. Following (Whittingham et al. 2005) I included a uniformally random variable as a null variable as even unimportant variables can have Akaiki weights notably greater than zero. The whole analysis was run 50 times, resampling the random variable each time. I calculated AICc for each model. I calculated the average AICc, $\bar{\text{AIC}}\text{c}$, by averaging AICc scores within models. ΔAICc was calculated as $\min(\bar{\text{AIC}}\text{c}) - \bar{\text{AIC}}\text{c}$, not the mean of the individual ΔAICc scores, to guarantee that the best model has $\Delta\text{AICc} = 0$. From these ΔAICc I calculated Akaiki weights, w. This value can be interpreted as the probability that a model would be the best model if the data were recollected. For each variable, the sum of the Akaiki weights of models containing that variable are summed to give *Pr*. This value can be interpreted as the probability that the given variable is in the best model.

2.4 Results

2.4.1 Number of Subspecies

After data cleaning there was data for 196 bat species in 11 families. There appears to be a positive relationship between the number of subspecies and viral richness (Figure 2.1) though few species have more than four subspecies. The number of described virus species for a bat host ranged up to 15 viruses in *Carollia perspicillata*. Figure 2.2 shows the phylogeny used and the number of viruses for each species. The mean number of viruses across families is fairly constant with a lower range of 1.67 for Nycteridae. The highest mean is Mormoopidae with 5 virus species per bat species, but this is based on a sample size of 3. The

Phyllostomidae have the second highest mean of 3.49 (n = 37).

The small change in mean pathogen richness across families and the lack of clear pattern in Figure 2.2 implies that viral richness is not strongly phylogenetic. This is corroborated by the small estimated size of λ ($\lambda = 0.04$, $p = 0.21$). This fact implies that other factors must control pathogen richness. It also implies that pathogens are not directly inherited down the phylogeny, although this is to be expected by the fast evolution of viruses.

Of the explanatory variables, the number of subspecies has no phylogenetic autocorrelation ($\lambda = 10^{-6}$, $p = 1$), study effort and distribution size have weak but significant autocorrelation (Study Effort: $\lambda = 0.1$, $p = 9.12 \times 10^{-3}$, Distribution size: $\lambda = 0.46$, $p < 10^{-5}$) and mass is strongly phylogenetic ($\lambda = 0.93$, $p < 10^{-5}$). The parameter λ is fitted and governs how important the phylogeny is in the model. Across all models the mean value of λ is 0.08 implying the residuals from the model are weakly phylogenetic. A small number of models (0.4%) had negatively phylogenetically distributed residuals.

The top seven models all had ΔAICc < 4 meaning there was no clear best model (Table 2.2). These top seven models had a combined weight of 0.96 meaning that there was a 0.96% chance that one of these models would be the best model if the data was recollected. However these top seven models all contained study effort, number of subspecies and the interaction between these two variables. log(Mass) and log(Range Size) and the random variable are all in three of the top seven models.

Summing the Akaiki weights of all models that contain a given variable gives a probability, Pr, that the variable would be in the best model (Figure 2.4A) if the data were recollected (Whittingham et al. 2006). The number of subspecies is very likely in the best model ($Pr > 0.99$) as is the interaction between number of species and study effort ($Pr = 0.96$) compared to the benchmark random variable which has $Pr = 0.21$ (see Figure 2.4A and Table 2.1). When models with the interaction term are removed, on average (mean weighted by Akaiki weights) there is a positive relationship between the number of subspecies and viral richness ($\beta = 0.63$, variance = 0.02). Models with an interaction between number of subspecies and study effort have a positive interaction term ($\beta = 0.5$, variance = 5.11×10^{-5}) and linear term ($\beta = 0.31$, variance = 2.13×10^{-4}). This supports the hypothesis that population structure promotes pathogen richness. The strong support for a positive interaction term implies that population structure has a stronger relationship with known pathogen richness in the presence of study effort. only predicts high

Table 2.1 Estimated variable weights (probability that a variable is in the best model) and their estimated coefficients for both number of subspecies and gene flow analyses. The coefficients for number of subspecies variable are also seperated for models with and without the interaction term because this term strongly changes the coefficient and because the coefficient can only be usefully interpreted when estimated without the interaction. However, there are no weights for these seperated terms as they are not directly compared in the model selection framework.

	Number of Subspecies		*Gene flow*	
Variable	*Pr*	Coefficient	*Pr*	Coefficient
Number of subspecies				
Total	1.00	0.32		
Models without interaction term		0.63		
Models with interaction term		0.31		
Number of subspecies:log(Scholar)	0.96	0.50		
Gene flow			0.89	−0.67
log(Scholar)	0.98	0.99	0.99	2.49
log(Mass)	0.80	0.48	0.98	−0.35
log(Range size)	0.66	0.35	0.06	1.57
Random	0.21	0.05	0.18	0.23

known pathogen richness in the presence of high study effort. One interpretation of this is that population structure alone does not predict high known richness; reasonable study effort is also needed to turn the expected high richness into known and recorded viral richness. Another interpretation is that having few subspecies does not predict low viral richness unless the species has been suitable sampley as the low number of subspecies is probably due to a lack of study rather than an accurate measurement.

As seen in Figure 2.4A, study effort is very likely in the best model ($\beta = 0.99$, $Pr > 0.98$). Body mass and range size are also probably in the best model ($\beta = 0.48$, $Pr = 0.8$ and $\beta = 0.35$, $Pr = 0.66$) with positive relationships of slightly lower strength than the number of subspecies in models without an interaction term ($\beta = 0.63$, variance = 0.02).

2.4.2 Gene Flow

Due to the low number of studies and the restrictive requirements imposed on study design, there was only data for 24 bat species in 7 families. The number

Table 2.2 Model selection results for number of subspecies and gene flow analysis. $\bar{AICc}$ is the mean AICc score across 50 resamplings of the null random variable. ΔAICc is the $\bar{AICc}$ score minus the lowest score. w is the Akaike weight and can be interpreted as the probability that the model is the best model (of those in the plausible set). $\sum w$ is the cumulative sum of the Akaike weights. log(Scholar)*NSubspecies implies the interaction term between study effort and number of subspecies as well as both of the individual linear terms. In the number of subspecies analysis there are many models with low ΔAICc scores suggesting there there is no single 'best model'. In the gene flow analysis, only the top model is supported.

Model	$\bar{AICc}$	ΔAICc	w	$\sum w$
Number of Subspecies				
log(Scholar)*NSubspecies + log(Mass) + log(RangeSize)	882	0.00	0.38	0.38
log(Scholar)*NSubspecies + log(Mass)	884	1.39	0.19	0.57
log(Scholar)*NSubspecies + rand + log(Mass)	885	2.24	0.12	0.70
log(Scholar)*NSubspecies	885	3.14	0.08	0.78
log(Scholar)*NSubspecies + log(RangeSize)	886	3.18	0.08	0.86
log(Scholar)*NSubspecies + rand + log(RangeSize)	886	3.94	0.05	0.91
log(Scholar)*NSubspecies + rand	886	3.95	0.05	0.96
log(Scholar) + NSubspecies + log(Mass) + rand	889	6.93	0.01	0.97
log(Scholar) + NSubspecies + log(Mass) + log(RangeSize) + rand	890	7.80	0.01	0.98
Gene flow				
log(Scholar) + log(Gene flow) + log(Mass)	71	0.00	1.00	1.00
log(Range size)	105	34.09	0.00	1.00
log(Mass)	106	35.06	0.00	1.00

of described virus species for a bat host ranged up to 12 viruses in *Miniopterus schreibersii*. Figure 2.3 shows the raw data for the relationship between pathogen richness and log gene flow.

As with the Number of Subspecies dataset, there was no phylogenetic signal in the number of virus species ($\lambda = 10^{-6}$, $p > 0.999$) Gene flow also had no phylogenetic autocorrelation ($\lambda = 10^{-6}$, $p > 0.999$). Due to the low sample size, significance tests are unlikely to have much power. However, study effort had some phylogenetic autocorrelation ($\lambda = 0.15$, $p = 0.56$) while distribution size and mass seemed to show phylogenetic signal (Distribution size: $\lambda = 0.67$, $p = 0.53$, Mass: $\lambda = 0.79$, $p = 2.69 \times 10^{-3}$).

Only the model with study effort, gene flow and mass was well supported with the second model having an ΔAICc of 34 (Table 2.2). While less strongly

supported than the number of subspecies, gene flow was likely in the best model ($Pr = 0.89$) compared to the benchmark random variable which has $Pr = 0.18$ (Figure 2.4B and Table 2.1). On average (mean weighted by Akaiki weights) there was a negative relationship between gene flow and viral richness ($\beta = -0.67$, variance $= 5.48 \times 10^{-3}$) despite the apparent positive relationship (see Figure 2.3) weakly suggested by the bivariate model (pgls: $\beta = 0.63$, $t = 1.16$, df = 13, $p = 0.27$). This supports the hypothesis that population structure promotes viral richness. Possibly due to the smaller sample size, or a weaker relationship, this coefficient was much more varied than the number of subspecies coeficient with 22% of models estimating a positive relationship.

As in the number of subspecies analysis, study effort was very likely in the best model ($Pr = 0.99$) as was body mass ($Pr = 0.98$). However, body mass has a negative average coefficient ($\beta = -0.35$, variance = 0.04) which is in contrast to number of subspecies analysis, many studies in the literature (Gay et al. 2014, Kamiya et al. 2014, Maganga et al. 2014, Turmelle & Olival 2009) and the bivariate model (pgls: $\beta = 0.6$, $t = 0.38$, df = 13, $p = 0.71$). In contrast to the number of subspecies analysis, range size was almost certainly not in the best model with $Pr = 0.06$ being much less than the random variable. This variable being less supported than the random variable was probably because range size is closely correlated with study effort (pgls: $\beta = 0.6$, $t = 4.45$, df = 13, $p = 6.58 \times 10^{-4}$). Of the three independant variables in the best model, study effort had the largest effect ($\beta = 2.49$, variance = 0.08). The effect size of gene flow ($\beta = -0.67$, variance $= 5.48 \times 10^{-3}$) was approximately twice the size of that of body mass ($\beta = -0.35$, variance = 0.04)

Across all models the mean value of λ is -1.64 and a large number of individual models (58%) had negatively phylogenetically distributed residuals implying the residuals from the model are strongly negatively phylogenetic. Due to the small sample size this was probably due to a small number of data points with large residuals being distant on the tree.

2.5 Discussion

I have tested the hypothesis that population structure promotes pathogen richness in bats. By analysing data on two measures of population structure, and using larger datasets than previous studies, it is hoped that any conclusions may be more robust than the conflicting results in the literature (Gay et al. 2014, Maganga

et al. 2014, Turmelle & Olival 2009). I have found that a positive affect of population structure (a positive effect of the number of subspecies and a negative effect of gene flow) are likely to be in the best models for explaining viral richness. Study effort is also clearly supported confirming the expectation that additional study of a bat species yields more known viruses infecting that species and highlighting again that this bias cannot be ignored in studies using known pathogen richness as a proxy for total pathogen richness.

Although I have used measures of study effort to try to control for biases in the viral richness data, this bias could still make the results here unreliable — this is especially true as study effort is by far the strongest predictor of viral richness in both datasets. It is hoped that as untargeted sequencing of viral genetic material (e.g. Anthony et al. (2013)) becomes cheaper and more common this bias can be reduced. The strength of the relationship between study effort and known viral richness also highlights the number of virus species and bat-virus host-pathogen relationships yet to be discovered.

I have used two measures of pathogen richness and the number of subspecies dataset is larger than those used in previous studies. However it is clear that the gene flow dataset is small (n = 24). This leads to some unexpected results. While the model averaging approach has given a negative model averaged coefficient for gene flow, the univariate model of gene flow against viral richness gave a positive coefficient. Furthermore body mass has a negative average coefficient. This is in contrast to number of subspecies analysis, many studies in the literature (Gay et al. 2014, Kamiya et al. 2014, Maganga et al. 2014, Turmelle & Olival 2009) and the bivariate model. It is not easy to interpret these contradictions but it is clear that the results from the gene flow analysis alone should not be considered strong evidence for a relationship between pathogen richness and population structure. The sensitivity of this analysis reiterates the need to use large datasets where possible and use mutliple measures of population structure to promote robust conclusions.

The results here suggest that there is a positive relationship between population structure and pathogen richness in bats. This is in agreement with (Maganga et al. 2014, Turmelle & Olival 2009) but in disagreement with (Gay et al. 2014). Furthermore it contradicts the assumption that factors that promote high R_0 will automatically promote high pathogen richness (Morand 2000, Nunn et al. 2003).

This relationship implies that direct or indirect competitive mechanisms are acting such that population structure is needed in order to allow escape from

competition.

The relationship between population structure and pathogen richness suggests that population structure has a least some potential as being predictive of high pathogen richness and therefore of a species likelihood of being a reservoir of a potentially zoonotic pathogen. However given that it is difficult to measure population structure and given that the relationship appears to be weak at best, this trait on it's own is unlikely to be useful in predicting zoonotic risk. However, as a number of other factors are also associated with pathogen richness (body mass and to a lesser extent range size here but also other traits elsewhere), using a combination traits in a predictive (i.e. machine learning) framework has potential to be used in prioritising zoonotic disease surveillance. The main hurdle in this approach is finding a way to validate models — due to the study effort bias in current data, predictive models will also be biased.

The relationship between pathogen richness and population structure also has implications for habitat fragmentation and range shifts due to global change. In short habitat fragmentation and range shifts that reduce movement between populations would be predicted to increase pathogen richness. However, depending on the mechanisms by which population structure increase pathogen richness this may not be a cause for concern. I the main mechanism is one that reduces pathogen extinction rates, a newly fragmented population is unlikely to increase it's pathogen richness over any appreciable timescale. If however population structure actively promotes the evolution of new pathogen strains or allows the persistence of more virulent strains this could have important public health implications. Therefore further study on the exact mechanisms by which population structure affects pathogen richness is needed.

In conclusion, this study adds to the evidence that population structure may promote pathogen richness. It does not support the view that factors that increase R_0 will increase pathogen richness. Using larger datasets and multiple measurements makes the weight of the evidence here stronger than in previous studies. However, caution must still be taken in interpreting these results as the data is biased and sparse and the analyses show that small changes in model and data choice can give opposite results.

3 Population structure does not increase pathogen diversity in theoretical bat populations

This work was conducted in collaboration with Kate Jones and Hilde Wilkinson-Herbots

3.1 Abstract

An increasingly large fraction of emerging diseases come from wild animals and these diseases have a huge impact on human health, healthcare systems and economic development. The chance that a new zoonosis will come from any particular wild host species increases with the diversity of pathogens in that species. However, the factors that control pathogen diversity in wild populations are still unclear.

Host species traits such as population density, longevity, body size and population structure have been shown to correlate with pathogen diversity. Typically it is assumed that well-connected, unstructured populations promote the invasion of new pathogens and therefore increase pathogen diversity. However, this assumption is largely untested; in particular our mechanistic understanding of how population structure affects pathogen diversity, in the presence of inter-pathogen competition, is poor. Greater understanding is needed to clarify exactly by which mechanisms population structure affects pathogen diversity.

It is unknown whether population structure allows invading pathogens to escape from competition by stochastically creating areas of low pathogen prevalence. I hypothesised that low dispersal rates and a low number of connections in a metapopulation network will allow invading pathogens to establish more readily, thus increasing pathogen diversity. I tested these hypothesese using population networks parameterised to mimic wild bat populations as bats have highly varied social structures and have recently been implicated in a number of high profile diseases such as Ebola, SARS, Hendra and Nipah.

I find that neither population connectedness nor dispersal rate affect the probability that a new pathogen will invade into a population. The assumption that factors causing high R_0 allow new pathogens to invade and therefore increase pathogen diversity is not supported. Instead we find that changes in population structure do not affect the probability of invasion of a new pathogen.

This result implies that population structure does not control pathogen diversity. Instead it seems that local dynamics, that affect the short term branching process at the very start of the epidemic control the chance of pathogen invasion. This also implies that population structure is not a useful proxy for pathogen diversity with respect to zoonotic disease surveillance, for example in bats.

3.2 Introduction

3.2.1 General Intro

Over 50% of emerging infectious diseases have an animal source (Jones et al. 2008, Smith et al. 2014). Zoonotic pathogens can be highly virulent (Lefebvre et al. 2014, Luby et al. 2009) and potentially have huge public health impacts (Granich et al. 2015), economic costs (Knobler et al. 2004) and slow down international development (World Bank 2014). Therefore understanding and predicting changes in the process of zoonotic spillover is a global health priority (Taylor et al. 2001). The number of pathogen species hosted by a wild animal species affects the chance that a disease from that species will infect humans. However, the factors that control the diversity of pathogens in a wild animal population are still unclear (Metcalf et al. 2015); in particular our mechanistic understanding of how population processes inhibit or promote pathogen richness is poor.

A number of host traits have been shown to correlate with pathogen richness including body size (Arneberg 2002, Kamiya et al. 2014), population density (Arneberg 2002, Nunn et al. 2003) and range size (Bordes & Morand 2011, Kamiya et al. 2014). There are few studies that study population structure in a comparative framework. Maganga *et al.* found that distribution fragmentation predicts viral richness (Maganga et al. 2014), but (Gay et al. 2014) finds the opposite relationship. While the data set in (Gay et al. 2014) is larger, the analysis in (Maganga et al. 2014) is much more focused on fragmentation. Genetic correlates of population structure have also been used. Turmelle *et al.* (Turmelle & Olival 2009), in a small analysis, find that high F_{st} (i.e. a structured population) correlates with high richness. However, they do not account for the widely different spatial scales found in population structure studies, nor do they deal with the differences between F_{ST}, ϕ_{ST} and other measures appropriately.

Empirical, correlative studies are often contradictory due to small sample sizes, noisy data and because empirical relationships often do not extrapolate well to other taxa (though see (Kamiya et al. 2014) for a meta-analysis). The correlation between many traits (e.g. Nunn et al. (2015)) also makes it hard to clearly distinguish which factors are important. Furthermore, knowing which factors correlate with pathogen richness does not tell us how they control richness. Mechanisms by which a trait could increase pathogen diversity include promoting the evolution of new strains within a species (Buckee et al. 2004), reduction of the rate of pathogen extinction (Rand et al. 1995) and an increased probability of

pathogen invasion from other host species (Nunes et al. 2006). These separate mechanisms have not been examined and it is difficult to see how they could be approached through comparative methods. We need explicit mechanistic models in order to tease apart these factors. Furthermore, a solid mechanistic understanding of these processes will allow us to predict how wild populations will respond to increased human pressure and global change. As habitats fragment we expect wild populations to become smaller and less well connected. Understanding how these population changes interact reqiures a mechanistic understanding of disease dynamics. Different specific mechanisms predict responses to climate change at different rates — the extinction of parasites could occur relatively rapidly, whereas processes whereby invasive pathogens can persist more easily are likely to change pathogen community competition relatively slowly.

Studies of the role of population structure on pathogen diversity have been in very simple systems. These have been so simple that empirical data cannot easily be applied to them to predict pathogen diversity of real wild animal populations. There is a need for models that can be carefully and fully explored, while still capturing the complexities of the real world.

In analytical models of well mixed populations competitive exclusion has been predicted (Ackleh & Allen 2003, Allen et al. 2004, Bremermann & Thieme 1989, Martcheva & Li 2013, Qiu et al. 2013). When competitive exclusion occurs, population structure has sometimes been shown to allow coexistence (Allen et al. 2004, Nunes et al. 2006, Qiu et al. 2013). Competing epidemics, or two pathogens spreading at the same time in a population, is a well studied area (Karrer & Newman 2011, Poletto et al. 2013, Poletto et al. 2015). This area is related to the study of pathogen richness in that they indicate that dynamics of multiple pathogens in a population do depend on population structure. However, the results for short term epidemic competition do not directly transfer to the study of long term disease persistence which is the relative time scale when looking at zoonotic potential of a host species.

One commonly taken assumption in the ecological literature is that factors that promote high disease spread automatically promotes high diversity (Altizer et al. 2003, Morand 2000, Nunn et al. 2003, Poulin 2014, Poulin & Morand 2000); this is contrary to the predictions from analytical models (Allen et al. 2004, Nunes et al. 2006, Qiu et al. 2013). Furthermore, this assumption ignores competitive mechanisms such as cross-immunity and depletion of susceptible hosts. If competitive mechanisms are strong, pathogens in populations structured such that R_0 will be

high will be able to easily out-compete invading pathogens. Only if competitive mechanisms are weak will high R_0 enable the invasion of new pathogens and allow higher pathogen diversity.

How structured a population is can be defined in many ways on many scales. Social group size is often examined as an indicator of population in large because the data are readily available. However, group size is only one aspect of population structure. Dispersal between is also a vital component of population structure. Furthermore, many species exhibit fission-fusion social structures or other less clearly defined groupings.

The most relevant scale to study is that of an epidemiological population. This is the population within which a pathogen can spread in an epidemiologically relevant time period (years or decades). It is therefore closely related to a population as defined by population genetics, but with movement defined on a shorter time scale.

The epidemiological contacts within the population can be examined at the individual level (as in contact network epidemiology) or larger scales. I consider the metapopulation network the most appropriate. Ignoring the metapopulation assumes a fully mixed population which is unlikely. Trying to study the contact network relies detailed individual level detail which is not available. Metapopulation models consider a network of small subpopulations. Within subpopulations, epidemiological contacts are fully mixed and relatively fast. Between subpopulations, epidemiological contacts are dependant on an underlying network structure and relatively slow. The network underlying the metapopulation is made up of nodes representing the subpopulations, and edges which represent movement between subpopulations. Animals, and therefore infection, can only move between two subpopulations if they are connected by an edge.

There are two factors that affect how structured a population is, given this model framework. Firstly, dispersal is the rate at which individuals move between subpopulations. Secondly, the metapopulation network structure controls population structure. The simplest measure of how structured the network is the average number of edges each node has. In the extremes, all subpopulations could be either connected to all other subpopulations or only connected to one or two other subpopulations. However, other measures that take into account second-order structure in the network are also often used.

Currently the literature contains very abstract, simplified models (Allen et al. 2004, Nunes et al. 2006, Qiu et al. 2013). These cannot be easily applied to real data.

They also do not easily quantitative predict pathogen pathogen; typically they predict either no pathogen coexistence (Bremermann & Thieme 1989, Martcheva & Li 2013) or infinite pathogen diversity (May & Nowak 1994). We need models that can give quantative predictions of pathogen richness in wild populations. While predicting an absolute value of pathogen richness is likely to be impossible, we should aim for models that attempt to rank species from highest to lowest pathogen richness. This requires a middle ground of model complexity.

To this end I have run metapopulation, multipathogen, epidemiological simulations based broadly on real-world bat populations. Although still simplified, the model is complex enough that if good measurements of bat populations could be found, simulations of the real world bat population could be run. Specifically I used these models to test the affects of population structure on the ability of a new pathogen to invade a population. I tested two aspects of structure, dispersal rate and connectedness of the metapopulation network.

I have used bats as a case study as they have an unusually large variety of social structures. Colony sizes range from 10 to 1 million (Jones et al. 2009). Many bats also have interesting seasonal behaviour such as migration (Fleming et al. 2003, Richter & Cumming 2008), hibernation, birth puleses and swarming behaviours. A number of traits have been suggested as predisposing bats towards being reservoirs of zoonotic diseases: high sympatry (Luis et al. 2013), flight (Wang et al. 2011), longevity (Wang et al. 2011) as well as population structure (Gay et al. 2014, Maganga et al. 2014, Turmelle & Olival 2009). Furthermore, bats have, over the last decade, become a focus for disease research (Calisher et al. 2006, Hughes et al. 2007). Recently they have been implicated in a number of high profile diseases such as Ebola, SARS, Hendra and Nipah (Calisher et al. 2006, Li et al. 2005).

I have studied the invasion of new pathogens as a mechanisms for increasing pathogen richness. I have studied two metrics for population structure, dispersal rate and metapopulation network topology, to test for effects of population structure on pathogen richness. In particular I have focussed on studying the invasion of a newly evolved pathogen that is therefore identical in epidemiological parameters to the endemic pathogen. Furthermore, this close evolutionary relationship means that cross immunity is strong.

Here I show that, given the assumptions of a metapopulation, population structure does not affect the rate of invasion of new pathogens. This is true whether population structure is altered by changing the dispersal rate or the topology of the metapopulation network.

3.3 Methods

3.3.1 Metapopulation model

3.3.1.1 Two pathogen SIR model

I examine a multipathogen SIR model with vital dynamics. This is a compartment model with individuals being classed as susceptible, infected or recovered with immunity (Figure 3.1). Susceptible individuals are counted in class S. There

Table 3.1 All symbols used and their units and values.

Symbol	Explanation	Units	Value
S	Number of susceptible individuals		
I_q	N. individuals infectious with diseases q		
I_p^+	Sum of classes infected with pathogen p		
$\bar{n}$	Colony starting size		3,000
β	Transmission rate	Transmission events per year per individual	2, 5, 10
γ	Recovery rate	Recovery events per year per individual.	1
λ	Dispersal	Dispersal events per day per individual	0.001–0.1
b	Birth rate	Births per year per individual	0.05
d	Death rate	Deaths per year per individual	0.05
ρ	No. pathogens		2
p	Pathogen index i.e. $p \in \{1, 2\}$ for pathogens 1 and 2		
q	Disease class i.e., $q \in \{1, 2, 12\}$		
t, t'	Time and time plus waiting time i.e., $t + \delta$	Days	
k_i	Degree of node i		
δ	Waiting time until next event	Days	
α	Cross immunity	Proportion	0.1
n, m	Colony index		
σ	Invading pathogen seed size		10
r_i	The rate that event i occurs.	Days^{-1}	

are three infected classes, I_1, I_2 and I_{12}, being individuals infected with pathogen 1, pathogen 2 or both respectively. Recovered individuals, R, are immune to both pathogens, even if they have only been infected with one. Furthermore, recovery from one pathogen moves an individual straight into the recovered class, even if the individual is infected with both pathogen. This modelling choice allows the model to be easily expanded to include more than two pathogens. The assumption of immediate recovery from all other diseases is likely to be quite accurate for very closely related pathogens as is being studied here as once an acquired immune response is activated, all infections are likely to be cleared quickly.

The coinfection rate (the rate at which an infected individual is infected with a second pathogen) is adjusted compared to the infection rate by a factor α (here I used $\alpha = 0.1$ which means coinfection happens at a tenth the rate of first infections). In the application of long term existence of pathogens it is vital to include vital dynamics (births and deaths) as the SIR without vital dynamics has no endemic state. Birth and death rates are assumed to be equal, $b = d$ (see Table 3.1 for a list of symbols and values used). The time scale of the simulations are set by setting b = 0.05 per year, yielding an average host generation time of 20 years. Infection is assumed to cause no extra mortality as for a number of viruses, bats show no clinical signs of infection (Halpin et al. 2011).

3.3.1.2 Metapopulation

The population is divided into a number of subpopulations. This metapopulation is modelled as a network with subpopulations being nodes and dispersal between subpopulations being indicated by edges (Figure 3.2). Individuals within a subpopulation interact randomly so that the subpopulation is fully mixed. Dispersal between subpopulations occurs at a rate λ. Individuals can only disperse to subpopulations connected to theirs in the network. The rate of dispersal is not affected by the number of edges a subpopulation has (the degree of the subpopulation). So the dispersal rate from a subpopulation m with degree k_m to subpopulation n is $\frac{\lambda}{k_m}$. Note this rate is independent of the degree of subpopulation n.

3.3.1.3 Stochastic simulations

I examine this model using stochastic, continuous time simulations (using the Gillespie algorithm) implemented in R (R Development Core Team 2010). At each step in the simulation we calculate the rate that each possible event might

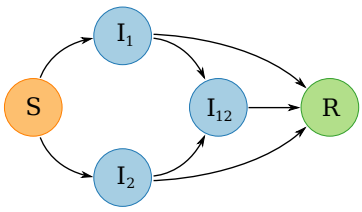


Figure 3.1 Schematic of the SIR model used. Individuals are in one five classes, susceptible (orange, S), infectious with pathogen 1, pathogen 2 or both (blue, I_1, I_2, I_{12}) or recovered and immune from further infection (green, R). Transitions between classes occurs only as indicated by arrows. Note that individuals in I_{12} move into R, not back to I_1 or I_2. That is, recovery from one pathogen causes immediate recovery from the other pathogen.

occur. One event is then randomly chosen, weighted by it's rate

$$p(\text{event } i) = \frac{r_i}{\sum_i r_i} \tag{3.1}$$

where r_i is the rate that event i occurs. Finally, the length of the time step, δ, is drawn from an exponential distribution

$$\delta \sim \text{Exp}\left(\sum_i r_i\right). \tag{3.2}$$

This means that the total length of each simulation is stochastic. We define the number of events we wish to simulate instead.

We can now write down the rates of all events. I defined I_p^+ to be the sum of all classes that are infectious with pathogen p, for example $I_1^+ = I_1 + I_{12}$. Assuming asexual reproduction, that all classes reproduce at the same rate and that individuals are born into the susceptible class we get

$$P\left(S_{nt'} = S_{nt} + 1\right) = b\left(S_{nt} + \sum_q I_{qnt} + R_{nt}\right) \tag{3.3}$$

where $P\left(S_{nt'} = S_{nt} + 1\right)$ is the probability that the number of susceptibles in subpopulation n will increase by 1 (a single birth) in the time interval t to t' and $\sum_q I_{qnt}$ is the sum of all infection classes $q \in 1, 2, 12$. The rates of death, given a death rate d are given by

$$P\left(S_{nt'} = S_{nt} - 1\right) = dS_{nt}, \tag{3.4}$$

$$P\left(I_{qnt'} = I_{qnt} - 1\right) = dI_{qnt}, \tag{3.5}$$

$$P\left(R_{nt'} = R_{nt} - 1\right) = dR_{nt}. \tag{3.6}$$

I modelled transmission as density-dependant. This assumption is more suitable than frequency-dependant transmission as I am modelling a disease transmitted by saliva or urine in highly dense populations confined to caves, building or potentially a small number of tree roosts. I am notably not modelling an STD as these diseases are not expected to be commonly zoonotic. Infection of a susceptible with either pathogen 1 or 2, $S \rightarrow I_p$ where $p \in \{1, 2\}$, is therefore given by

$$P\left(I_{pnt'} = I_{pnt} + 1, S_{nt'} = S_{nt} - 1\right) = \beta S_{nt} I^{+}_{pnt}, \tag{3.7}$$

while coinfection, given a crossimmunity factor α, is given by

$$P\left(I_{12,nt'} = I_{12,nt} + 1,\, I_{1nt'} = I_{1nt} - 1\right) = \alpha\beta I_{1nt} I^{+}_{2nt}, \tag{3.8}$$

$$P\left(I_{12,nt'} = I_{12,nt} + 1,\, I_{2nt'} = I_{2nt} - 1\right) = \alpha\beta I_{2nt} I^{+}_{1nt}. \tag{3.9}$$

The probability of migration from colony m (with degree k_m) to colony n, given a dispersal rate λ is given by

$$P\left(S_{nt'} = S_{nt} + 1,\, S_{mt'} = S_{mt} - 1\right) = \frac{\lambda S_{mt}}{k_m - 1}, \tag{3.10}$$

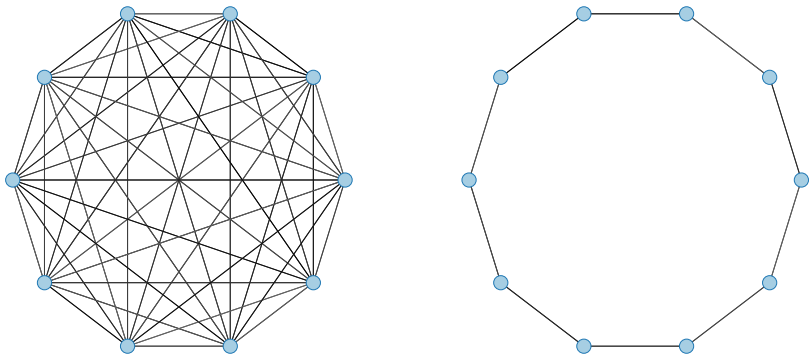
$$P\left(I_{qnt'} = I_{qnt} + 1,\, I_{qmt'} = I_{qmt} - 1\right) = \frac{\lambda I_{qmt}}{k_m - 1}, \tag{3.11}$$

$$P\left(R_{nt'} = S_{nt} + 1,\, R_{mt'} = R_{mt} - 1\right) = \frac{\lambda R_{mt}}{k_m - 1}. \tag{3.12}$$

Finally, recovery from any infectious class occurs at a rate γ

$$P\left(I_{qnt'} = I_{qnt} - 1,\, R_{nt'} = R_{nt} + 1\right) = \gamma I_{qnt}. \tag{3.13}$$

In each simulation the population was seeded with 10 sets of 200 infected individuals of pathogen 1. These groups were seeded into randomly selected colonies with replacement. For each 200 infected individuals added, 200 susceptible individuals were removed to keep starting colony sizes constant. Pathogen 1 was then allowed to spread and reach equilibrium. After 3×10^5 events, 5 individuals infected with pathogen 2 were added to one randomly selected colony. Visual inspection of preliminary simulations was used to decide on 3×10^5 as being long enough for the epidemic to reach an equilibrium state. After another 5×10^5 events the invasion of pathgen 2 was considered successful if any individuals are still infected with pathogen 2. Again visual inspection of preliminary simulations was used to determine that after 5×10^5 events, if an invading pathogen was still present, it was well established.



A Fully connected **B** Minimally connected

Figure 3.2 The two network topologies used to test whether network connectedness influences a pathogens ability to invade. Blue circles are subpopulations of 3,000 individuals. Dispersal only occurs between subpopulations connected by an edge (black line). The dispersal rate is held constant between the two topologies. A) Dispersal can occur between any subpopulation. B) Animals can only disperse to neighbouring subpopulations.

3.3.2 Dispersal

The values used for the independant variables are chosen to highlight the affects of these variables. Dispersal values are $\lambda = 0.1, 0.01$ and 0.001 dispersals per individual per year. $\lambda = 0.1$ relates to individuals moving between colonies on average twice per lifetime. Therefore exclusively juvenile dispersal would have dispersal rates similar to this. Otherwise it relates to dispersal being a rare event with animals often staying in a colony for many years. $\lambda = 0.01$ relates to 20% of individuals dispersing once in their lifetime. This value is therefore close to male-biased dispersal, with female philopatry. Finally, $\lambda = 0.001$ relates to 2% of individuals dispersing in their lifetime. This therefore relates to a population that does not habitually disperse.

3.3.3 Network structure

The network structure is synthetically created to be either fully or minimally connected (See Figure 3.2). 10 subpopulations was selected as a trade off between computation time and a network complicated enough that structure might have an effect. This value is artificially small compared to wildlife populations.

3.3.4 Parameter selection

The fixed parameters used are chosen to roughly reflect realistic wild bat populations. The death rate d is set as 0.05 per year giving a generation time of 20 years. The birth rate b is set to be equal to d so that the population size is stable. The recovery rate γ is set to 1 giving a average infection duration of 1 years. This is therefore a long lasting infection but not a chronic infection. It is very difficult to directly estimate infection durations in wild populations but it seems that these infections might sometimes be long lasting (Peel et al. 2012, Plowright et al. 2015). However, other studies have found much shorter infectious periods (Amengual et al. 2007). These shorter lived in infections are studied further here as preliminary simulations found that they could not persist in the relatively small populations being modelled here.

Cross immunity is set to 0.1 so that an individual infected with one pathogen is 90% less likely to be infected with another. This is a rather arbitrary value. However, the rationale of the model is that the invading species might be a newly speciated strain of the endemic species. Furthermore, the model assumes complete cross immunity after recovery from infection. Therefore cross immunity to coinfection is likely to be very strong as well.

The population size of each subpopulation is set to 3,000. This is appropriate for many bat species (Jones et al. 2009), especially the large, frugivorous *Pteropodidae* that have been particularly associated with recent zoonotic diseases.

Four values of the transmission rate β are used, 0.1, 0.2, 0.3 and 0.4. All simulations are run under all four transmission rates as this is such a fundamental parameter.

For the simulations where an invading pathogen is added to the populations the number of invading pathogens added is set to 10. This is a trade off between getting a reasonable proportion of invasions, while still retaining the stochastic nature of invasion.

3.4 Results

3.4.1 Invasion

3.4.1.1 Dispersal

The proportion of invasions was not different across dispersal rates across 1200 simulations run (Figure 3.3A). This was true at all transmission levels (χ^2 test.

$\beta = 0.1$: $\chi^2 = 1.98$, df $= 2$, $p = 0.37$. $\beta = 0.2$: $\chi^2 = 9.44$, df $= 2$, $p = 8.9 \times 10^{-3}$. $\beta = 0.3$: $\chi^2 = 0.29$, df $= 2$, $p = 0.87$, $\beta = 0.4$: $\chi^2 = 0.83$, df $= 2$, $p = 0.66$).

3.4.1.2 Network structure

I ran 800 simulations over 4 transmission values ($\beta = 0.1, 0.2, 0.3, 0.4$). The proportion of invasions was not different between highly connected and largely unconnected metapopulations (Figure 3.3B). This was true at all transmission levels (χ^2 test. $\beta = 0.1$: $\chi^2 < 10^{-5}$, df $= 1$, $p > 0.9999$. $\beta = 0.2$: $\chi^2 = 0.02$, df $= 1$, $p = 0.88$. $\beta = 0.3$: $\chi^2 = 1.53$, df $= 1$, $p = 0.22$. $\beta = 0.4$: $\chi^2 = 0.26$, df $= 1$, $p = 0.61$).

3.4.2 Transmission

Inline with theory, increasing the transmission rate increased the probability of invasion (Figure 3.3). This is true for all three dispersal values (χ^2 test. $\lambda = 0.001$: $\chi^2 = 243.56$, df $= 3$, $p < 10^{-5}$. $\lambda = 0.01$: $\chi^2 = 265.4$, df $= 3$, $p < 10^{-5}$. $\lambda = 0.1$: $\chi^2 = 244.95$, df $= 3$, $p < 10^{-5}$) and both network structures (χ^2 test. Fully connected: $\chi^2 = 267.3$, df $= 3$, $p < 10^{-5}$. Minimally connected: $\chi^2 = 267.3$, df $= 3$, $p < 10^{-5}$).

3.5 Discussion

Empirical studies on the role of population structure on pathogen richness are equivocal and cannot examine the specific mechanisms by which pathogen communities are created and maintained. I have used mechanistic, metapopulation models to test whether increased population structure can promote pathogen richness by facilitating invasion of new pathogens. I find that population structure does not affect the ability of a new pathogen to invade and persist in a population. Instead I find that only transmission rate affects the chance of pathogen invasion. That population structure does not affect pathogen richness goes against many predictions that increasing R_0 increases pathogen richness (Altizer et al. 2003, Morand 2000, Nunn et al. 2003, Poulin 2014, Poulin & Morand 2000). However, simple analytical models suggest that population structure should increase pathogen richness (Allen et al. 2004, Nunes et al. 2006, Qiu et al. 2013) and I find no evidence of this either.

I have found no evidence that population structure aids the invasion and establishment of newly evolved pathogen species. Instead it seems that invasion relies solely on the local dynamics of the disease. If the transmission rate is high,

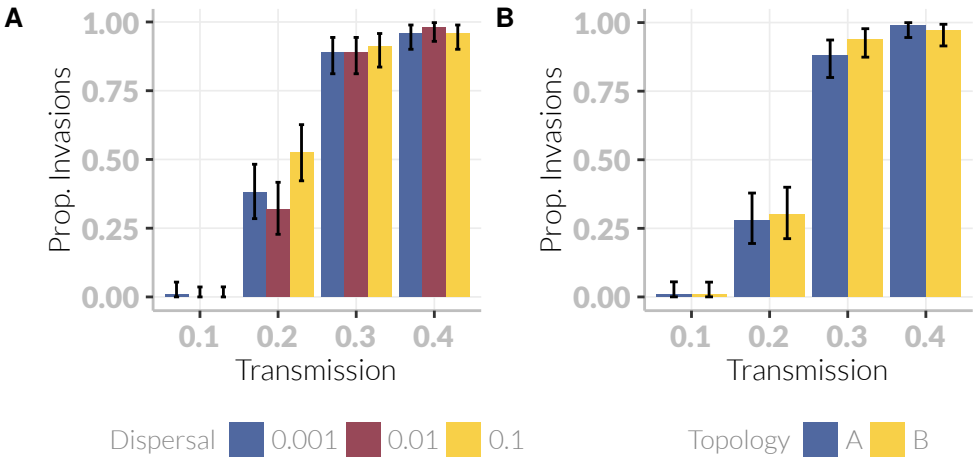


Figure 3.3 The probability of successful invasion. For three different transmission rates, the probability of invasion does not change between different a) dispersal rates or b) network topologies (with network topologies A and B refering to the fully and minimally connected topologies in Figure 3.2). Error bars are 95% confidence intervals. 100 simulations were run for each treatment. Other parameters are kept constant at: $N = 10$, $b = d = 0.05$, $\gamma = 0.1$, $\alpha = 0.1$. When dispersal is varied, the population structure is fully connected. When population structure is varied, $\lambda = 0.01$.

a pathogen has a good chance of spreading immediately after its introduction when it is at low density. Once the pathogen has infected a large number of individuals in the local metapopulation, it is bound to spread throughout the metapopulation without going extinct.

These results imply that if population structure does in fact affect pathogen richness (Gay et al. 2014, Maganga et al. 2014, Turmelle & Olival 2009) it must occur by a mechanism other than the one studied here. Therefore it is not the spread and persistance of a newly evolved pathogen that is facilitated by population structure. Other mechanisms that should be examined include reduced competitive exclusion of already established pathogens or increased invasion of less closely and less strongly competiting pathogens, perhaps mediated by ecological competition of pathogens (i.e. reducuction of the susceptible pool by disease induced mortality). Furthermore, single pathogen dynamics could have an important role such as population structure causing a much slower, asynchronous epidemic preventing acquired herd immunity (Plowright et al. 2011).

Given that we find no affect of population structure on the ability of a new

pathogen to invade, this host trait is not useful for predicting the probability that a wild species has many pathogens. This implies it should not be used as a proxy for the probability that a species carries a zoonotic pathogen.

However, my simulations do highlight the importance of competition for the spread of a new pathogen. All parameters used correspond to pathogens with $R_0 > 1$. However, the competition with the endemic pathogen means that for some transmission rates the chance of epidemic spread and persistence is close to zero. This has implications for human epidemics as well — if there is strong competiton between a newly evolved strain and an endemic strain, we are unlikely to see the new strain spread, irregardless of population structure.

3.5.1 Model assumptions

3.5.1.1 Complete cross-immunity

I have assumed that once recovered, individuals are immune to both pathogens. Furthermore, when a coinfected individual recovers from one pathogen, it immediately recovers from the other as well. This is probably a fairly reasonable assumption given that I am modelling a newly evolved strain. However, further work could relax this assumption using a model similar to (Poletto et al. 2015) which contains additional classes for 'infected with pathogen one, immune to pathogen two' and 'infected with pathogen two, immune to pathogen one'. The model here was formulated such that the study of systems with greater than two pathogens is still computationally feasible while a model such as used in (Poletto et al. 2015) contains 3^ρ classes for a system with ρ pathogen species. This quickly becomes computationally restrictive.

3.5.1.2 Identical strains

Many papers on pathogen richness have focussed on the evolution of pathogen traits and have considered a trade off between transmission rate and virulence (Nowak & May 1994) or infectious period (Poletto et al. 2013). However, here I am interested in host traits. Therefore we have assumed that pathogen strains are identical. It is clear however that there are a number of factors that affect pathogen richness and our focus on host population structure does not imply that pathogen traits are not important.

3.5.1.3 Complex social structure and behaviour

With the models here I have aimed to tread a middle ground between the overly simple models employed in analytical studies (Allen et al. 2004) and the full complexity and variety of true bat social systems (Kerth 2008). Omissions include seasonal migration, maternity roosts, hibernation roosts and swarming sites (Cryan et al. 2014, Fleming et al. 2003, Kerth 2008, Richter & Cumming 2008). While future models might aim to model this complexity more fully the number of parameters that are required to be estimated and varied becomes very large. Furthermore, not all of these social complexities exist in all bat species, so in limiting my analysis to the simpler end of bat social systems it is hoped that the results are more broadly representative of the order.

Furthermore, I have considered a single host species in isolation. In some cases, treating multiple host species as identical could be appropriate. However, more generally, it seems likely that sympatry in bats is epidemiologically important (Brierley et al. 2016, Luis et al. 2013) but was beyond the scope of this study. There is potential for this to be effectively modelled as a multilayered network (Funk & Jansen 2010, Wang et al. 2016) and this would be expected to act to reduce population structure.

Finally, many species of bat exhibit strong seasonal birth pulses which are known to affect disease dynamics (Amman et al. 2012, Hayman 2015, Peel et al. 2014). This would be expected to facilitate the invasion of new pathogen species; if a new strain evolved or entered the population by migration during a period of low population immunity, it would have a higher chance of invading and establishing in the population.

3.5.2 Conclusions

In conclusion I have found no evidence that population structure aids the invasion and establishment of newly evolved pathogen species. This suggests that if population structure does have a role in shaping pathogen communities, it is not by this specific mechanism. Practically, this implies that population structure is not a useful metric for predicting pathogen richness in wild bats. Furthermore it implies that population fragmentation by global change will not promote an increase in pathogen richness.

4 The interactions between population structure and density in pathogen diversity

This work was conducted in collaboration with Kate Jones and Hilde Wilkinson-Herbots

4.1 Abstract

An increasingly large fraction of emerging diseases come from animals (Jones et al. 2008, Taylor et al. 2001) and these diseases have a huge impact on human health. The chance that a new disease will come from any particularly wild host species increases with the diversity of pathogens in that species. However, the factors that control pathogen diversity in wild populations are still unknown.

Population density is known to increase pathogen richness while theory suggests that population structure may also play a role. However, these factors, along with population abundance, are intrinsically linked; reducing density reduces contacts between individuals. In group living species, this is particularly true, with group size and the number of groups and distribution size contributing to total animal density. As these factors are completely interdependant, it is very difficult to study them empirically e.g. in a comparative frame work.

It is unknown whether it is specifically density that controls pathogen diversity or whether density merely correlates with population structure, group size or population size (abundance).

Here I use metapopulation SIR models to test whether it is density *per se* that increases the ability of a new pathogen to invade as apposed to colony size, population abundance or population structure.

4.2 Introduction

4.2.1 General Intro

4.2.1.1 Why is pathogen diversity important?

4.2.1.2 We know some factors that correlate with pathogen diversity

4.2.1.3 But we do not understand the mechanistic processes

4.2.2 Specific Intro

4.2.3 The gap

4.2.4 What I did

4.2.5 What I found

4.3 Methods

4.3.1 Metapopulation model

4.3.1.1 Two pathogen SIR model

We examine a multpathogen SIR model. This is a compartment model with individuals being classed as susceptible, infected or recovered with immunity (Figure 3.1). Susceptible individuals are counted in class S. There are three infected classes, I_1, I_2 and I_{12}, being individuals infected with pathogen 1, pathogen 2 or both respectively. Recovered individuals, R, are immune to both pathogens, even if they have only been infected with one. Furthermore, recovery from a pathogen moves an individual straight into the recovered class, even if the individual is infected with both pathogen. This modelling choice allows the model to be easily expanded to included more than two pathogens. The assumption of immediate recovery from all other diseases is likely to be quite accurate for very closely related pathogens as is being studied here as once an acquired immune response is activated, all infections are likely to be cleared quickly.

The coinfection rate is adjusted compared to the first infection rate by a factor α. Birth and death rates are assumed to be equal, $b = d$.

4.3.1.2 Metapopulation

The population is divided into a number of subpopulations. This metapopulation is modelled as a network with subpopulations being nodes and dispersal between subpopulations being indicated by edges (Figure 3.2) Individuals with a subpopulation interact randomly so that the subpopulation is fully mixed. However, dispersal between subpopulations occurs at a rate λ. Individuals can only disperse to subpopulations connected to theirs in the network. The rate of dispersal is not affected by the number of edges a subpopulation has (the degree of the subpopulation). So the dispersal rate from a subpopulation m with degree k_m to subpopulation n is $\frac{\lambda}{k_m}$. Note this rate is independent of the degree of subpopulation n.

4.3.1.3 Stochastic simulations

We examine this model using stochastic, continuous time simulations using the Gillespie algorithm. At each step in the simulation we calculate the rate that each possible event might occur. One event is then randomly chosen, weighted by it's

rate

$$p(\text{event } i) = \frac{r_i}{\sum_i r_i} \tag{4.1}$$

where r_i is the rate that event i occurs. Finally, the length of the time step, δ, is drawn from an exponential distribution

$$\delta \sim \text{Exp}\left(\sum_i r_i\right). \tag{4.2}$$

This means that the length of each simulation is stochastic. We define the number of events we wish to simulate instead.

We can now write down the rates of all events. I define I_p^+ to be the sum of all classes that are infectious with pathogen p, for example $I_1^+ = I_1 + I_{12}$. Assuming asexual reproduction, that all classes reproduce at the same rate and that individuals are born into the susceptible class we get

$$P\left(S_{nt'} = S_{nt} + 1\right) = b\left(S_{nt} + \sum_q I_{qnt} + R_{nt}\right) \tag{4.3}$$

where $P\left(S_{nt'} = S_{nt} + 1\right)$ is the probability that the number of susceptibles in subpopulation n will increase by 1 (a single birth) the short time interval t to t' and $\sum_q I_{qnt}$ is the sum of all infection classes $q \in 1, 2, 12$. The rates of death, given a death rate d are given by

$$P\left(S_{nt'} = S_{nt} - 1\right) = dS_{nt} \tag{4.4}$$

$$P\left(I_{qnt'} = I_{qnt} - 1\right) = dI_{qnt} \tag{4.5}$$

$$P\left(R_{nt'} = R_{nt} - 1\right) = dR_{nt}. \tag{4.6}$$

Infection of a susceptible with either pathogen 1 or 2, $S \to I_p$ where $p \in \{1, 2\}$, is given by

$$P\left(I_{pnt'} = I_{pnt} + 1, S_{nt'} = S_{nt} - 1\right) = \beta S_{nt} I_{pnt}^+, \tag{4.7}$$

while coinfection, given a crossimmunity factor α, is given by

$$P\left(I_{12,nt'} = I_{12,nt} + 1,\, I_{pnt'} = I_{pnt} - 1\right) = \alpha\beta I_{nt} I_{pnt}^+. \tag{4.8}$$

The probability of migration from colony m (with degree k_m) to colony n, given a dispersal rate λ is given by

$$P\left(S_{nt'} = S_{nt} + 1,\, S_{mt'} = S_{mt} - 1\right) = \frac{\lambda S_{mt}}{k_m - 1} \tag{4.9}$$

$$P\left(I_{qnt'} = I_{qnt} + 1,\, I_{qmt'} = I_{qmt} - 1\right) = \frac{\lambda I_{qmt}}{k_m} \tag{4.10}$$

$$P\left(R_{nt'} = S_{nt} + 1,\, R_{mt'} = R_{mt} - 1\right) = \frac{\lambda R_{mt}}{k_m}. \tag{4.11}$$

Finally, recovery from any infectious class occurs at a rate γ

$$P\left(I_{qnt'} = I_{qnt} - 1,\, R_{nt'} = R_{nt} + 1\right) = \gamma I_{qnt}. \tag{4.12}$$

In each simulation the population is seeded with 20 infected individuals of disease 1 in each colony. Disease 1 is then allowed to spread and reach equilibrium. After 7×10^5 events, 5 individuals infected with disease 2 are added to one colony. After another 3×10^5 events the invasion of disease 2 is considered successful if any individuals with disease 2 still remain.

4.3.2 Dispersal

The values used for the independant variables are chosen to highlight the affects of these variables. Dispersal values are $\lambda = 0.1, 0.01$ and 0.001 dispersals per individual per year. $\lambda = 0.1$ relates to individuals moving between colonies on average twice per lifetime. Therefore exclusively juvenile dispersal would have dispersal rates similar to this. Otherwise it relates to dispersal being a rare event with animals often staying in a colony for many years. $\lambda = 0.01$ relates to 20% of individuals dispersing once in their lifetime. This value is therefore close to male-biased dispersal, with female philopatry. Finally, $\lambda = 0.001$ relates to 2% of individuals dispersing in their lifetime. This therefore relates to a population that does not habitually disperse.

4.3.3 Network structure

The network structure is synthetically created to be either fully or minimally connected (See Figure 3.2). 10 subpopulations was selected as a trade off between computation time and a network complicated enough that structure might have an effect. This value is artificially small compared to wildlife populations.

4.3.4 Parameter selection

The fixed parameters used are chosen to roughly reflect realistic wild bat populations. The death rate d is set as 0.05 per year giving a generation time of 20 years. The birth rate b is set to be equal to d so that the population size is stable. The

recovery rate γ is set to 0.1 giving a average infection duration of 10 years. This is therefore a chronic infection. It is very difficult to directly estimate infection durations in wild populations. But it seems that these infections might be long lasting ().

Cross immunity is set to 0.1 so that an individual infected with one disease is 90% less likely to be infected with another. This is a rather arbitrary value. However, the model assumes complete cross immunity after infection. Furthermore, the rationale of the model is that the invading species might be a newly speciated strain of the endemic species. Therefore cross immunity is likely to be very strong.

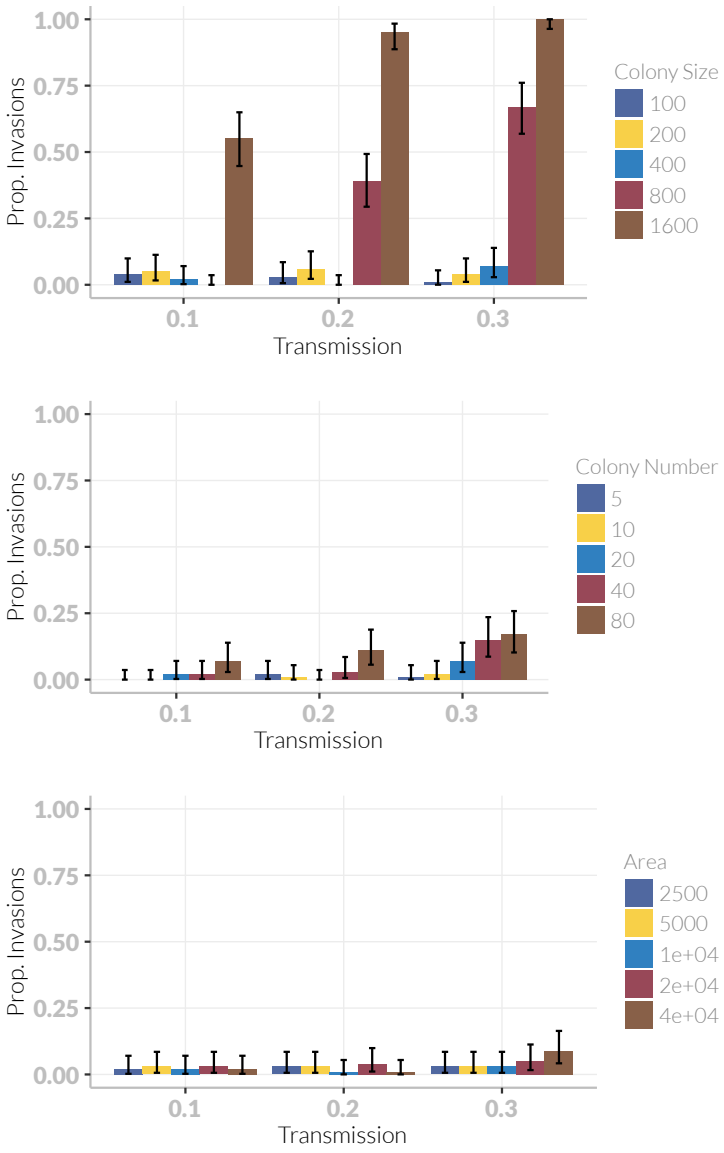
The population size of each subpopulation is set to 3000. This is appropriate for many bat species (Jones et al. 2009), especially the large, frugivorous *Pteropodidae* that have been particularly associated with recent zoonotic diseases.

Three values of the transmission rate β are used, 2, 5 and 10. All simulations are run under all three transmission rates as this is such a fundamental parameter. Given the recovery, birth and death rates we can calculate an approximation of R_0 that ignores spatial structure. That is, this is R_0 for the local, within-subpopulation dynamics. Furthermore, it is R_0 for the first pathogen; R_0 of the invading pathogen will be lower due to competition. We can calculate that $R_0 \approx \frac{\beta d}{d(d+\gamma)}$. For our three values of $\beta = 2, 5, 10$ we therefore get $R_0 \approx 13.3, 33.3, 66.6$. These values are very high in part to again find a reasonable trade off between the number of simulations and the reasonableness of the parameters. $R_0 \approx 13.3$ is similar to a highly contagious disease such as measles or pertussis.

For the simulations where an invading pathogen is added to the populations the number of invading pathogens added is set to 10. This is a trade off between getting a reasonable proportion of invasions, while still retaining the stochastic nature of invasion.

For simulations studying extinction rates, half of the subpopulation were seeded with each pathogen. 2000 susceptibles and 1000 infected individuals were placed in each subpopulation in order to quickly reach equilibrium.

4.4 Results


Prop. Invasions
1.00
0.75
0.50
0.25
0.00
0.1
0.2
0.3
Transmission
Colony Size
100
200
400
800
1600
Colony Number
5
10
20
40
80
Area
2500
5000
1e+04
2e+04
4e+04

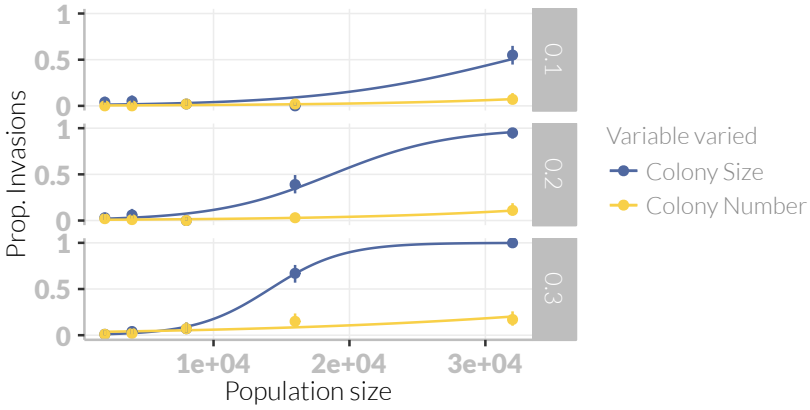


Figure 4.1 Comparison of the effect of population size on probability of invasion when population size is altered by changing colony size or colony number. Relationship is shown seperately for each transmission value. It can be seen that changes in colony size give a much greater increase in invasion probability thatn changes in colony number.

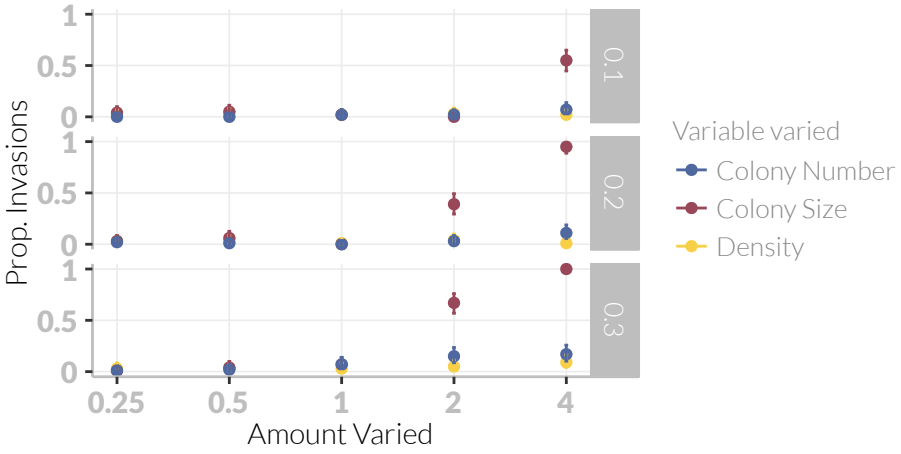


Figure 4.2 Relationship is shown seperately for each transmission value.

4.5 Discussion

4.5.1 Restate the gap and the main result

Empirical studies on the role of population structure on the are equivocal and cannot examine the specific mechanisms by which pathogen communities are created and maintained. I have used mechanistic, metapopulation models to test whether increased population structure can promote pathogen richness by facilitating invasion of new pathogens.

4.5.2 Link results to consequences

4.5.2.1 Population structure does not affect pathogen richness

Probably because dynamics are dominated by local processes. This goes against many predictions that increasing R_0 increases pathogen richness. Further work could examine reduced col ony sizes to test when global structure become more important. Measures of population structure should not be used to predict zoonotic potential.

4.5.2.2 Dispersal does not affect pathogen richness

4.5.2.3 Network connectedness does not affect pathogen richness

This is in direct contrast to (Campos & Gordo 2006). However, the model in (Campos & Gordo 2006) is a contact network, so increasing the connectedness increases the chance of succesful transmission events for the first few transmission generations. This lends support to the idea that I found now affect of connectedness due to the dominance of local dynamics.

Network connectedness can be seen as a function of average dispersal distance, density and colony size. A high density species with small colony sizes must have colonies relatively close together. Therefore colonies would be more likely to be connected for a given dispersal distance.

4.5.3 Discuss assumptions

4.5.3.1 Complete cross-immunity

I have assumed that once recovered, individuals are immune to both pathogens. Furthermore, when a coinfected individual recovers from one pathogen, it im-

mediately recovers from the other as well. This is probably a fairly reasonable assumption given that I am modelling a newly evolved strain. However, further work could relax this assumption using a model similar to (Poletto et al. 2015) which contains additional classes for 'infected with pathogen one, immune to pathogen two' and 'infected with pathogen two, immune to pathogen one'. The model here was formulated such that the study of systems with greater than two pathogens is still computationally feasible while a model such as used in (Poletto et al. 2015) contains 3^{ρ} classes for a system with ρ pathogen species. This quickly becomes computationally restrictive.

4.5.3.2 Identical strains

Many papers on pathogen richness have focussed on the evolution of pathogen traits and have considered a trade off between transmission rate and virulence (Nowak & May 1994) or infectious period (Poletto et al. 2013). However, here we are interested in host traits. Therefore we have assumed that pathogen strains are identical. It is clear however that there are a number of factors that affect pathogen richness and our focus on host population structure does not imply that pathogen traits are not important.

4.6 Appendix

	Explanation	Units	Value
S	Susceptible individuals		
I_q	Infectious with diseases q		
I_p^+	Sum of classes infected with pathogen p		
N	Number of colonies		10
$\bar{n}$	Mean colony starting size		3000
β	Transmission rate	Transmission events per year per individual	2, 5, 10
γ	Recovery rate	Recovery events per year.	0.1
λ	Dispersal	Dispersal events per day per individual	0.001–0.1
b	Birth rate	Births per year per individual	0.05
d	Death rate	Deaths per year per individual	0.05
d_I	Infectious death rate	Additional deaths per day per individual	
ρ	No. pathogens		2
p	Pathogen index i.e. $p \in \{1, 2\}$ for pathogens 1 and 2		
q	Disease class i.e., $q \in \{1, 2, 12\}$		
$\mathcal{U}$	Neighbourhood of a node		
t, t'	Time and time plus waiting time i.e., $t + \delta$	Days	
k_i	Degree of node i		
δ	Waiting time until next event	Days	
α	Cross immunity	Proportion	0.1
n, m	Colony index		
μ	Maximum distance for edge to exist	km	40, 100
σ	Invading pathogen seed size		10
r_i	The rate that event i occurs.	Days^{-1}	

Table 4.1 All symbols used.

5 A generalised random encounter model for estimating animal density with remote sensor data

This work was conducted in collaboration with Elizabeth Moorcroft, Robin Freeman, Marcus Rowcliffe and Kate Jones and is now published in Methods in Ecology and Evolution (Lucas et al. 2015). I formulated and analysed the analytical model. Elizabeth Moorcroft wrote the simulations. I lead the writing of the manuscript, but much of it was written by Elizabeth Moorcroft and other coauthors.

5.1 Abstract

Wildlife monitoring technology is advancing rapidly and the use of remote sensors such as camera traps and acoustic detectors is becoming common in both the terrestrial and marine environments. Current methods to estimate abundance or density require individual recognition of animals or knowing the distance of the animal from the sensor, which is often difficult. A method without these requirements, the random encounter model (REM), has been successfully applied to estimate animal densities from count data generated from camera traps. However, count data from acoustic detectors do not fit the assumptions of the REM due to the directionality of animal signals.

We developed a generalised REM (gREM), to estimate absolute animal density from count data from both camera traps and acoustic detectors. We derived the gREM for different combinations of sensor detection widths and animal signal widths (a measure of directionality). We tested the accuracy and precision of this model using simulations of different combinations of sensor detection widths and animal signal widths, number of captures, and models of animal movement.

We find that the gREM produces accurate estimates of absolute animal density for all combinations of sensor detection widths and animal signal widths. However, larger sensor detection and animal signal widths were found to be more precise. While the model is accurate for all capture efforts tested, the precision of the estimate increases with the number of captures. We found no effect of different animal movement models on the accuracy and precision of the gREM.

We conclude that the gREM provides an effective method to estimate absolute animal densities from remote sensor count data over a range of sensor and animal signal widths. The gREM is applicable for count data obtained in both marine and terrestrial environments, visually or acoustically (e.g., big cats, sharks, birds, echolocating bats and cetaceans). As sensors such as camera traps and acoustic detectors become more ubiquitous, the gREM will be increasingly useful for monitoring unmarked animal populations across broad spatial, temporal and taxonomic scales.

5.2 Introduction

The density of animal populations is one of the fundamental measures in ecology and conservation and has important implications for a range of issues, such

as sensitivity to stochastic fluctuations (Wright & Hubbell 1983) and extinction risk (Purvis et al. 2000). Monitoring animal population changes in response to anthropogenic pressure is becoming increasingly important as humans rapidly modify habitats and change climates (Everatt et al. 2014). Sensor technology, such as camera traps (Karanth 1995, Rowcliffe & Carbone 2008) and acoustic detectors (Acevedo & Villanueva-Rivera 2006, Walters et al. 2012) are widely used to monitor changes in animal populations as they are efficient, relativity cheap and non-invasive, allowing for surveys over large areas and long periods (Kessel et al. 2014, Rowcliffe & Carbone 2008, Walters et al. 2013). However, converting sampled count data into estimates of density is problematic as detectability of animals needs to be accounted for (Anderson 2001).

Existing methods for estimating animal density often require additional information that is often unavailable. For example, capture-mark-recapture methods (Borchers et al. 2014, Karanth 1995) require recognition of individuals, and distance methods (Harris et al. 2013) require estimates of how far away individuals are from the sensor (Barlow & Taylor 2005, Marques et al. 2011). When individuals cannot be told apart, an extension of occupancy modelling can be used to estimate absolute abundance (Royle & Nichols 2003). However, as the model is originally formulated to estimate occupancy, count information is simplified to presence–absence data. Assumptions about the distribution of individuals (e.g. a poisson distribution) must also be made (Royle & Nichols 2003) which may be a poor assumption for nonrandomly distributed species. Furthermore repeat, independent surveys must be performed and the definition of a site can be difficult, especially for wide-ranging species (MacKenzie & Royle 2005).

The REM method has been successfully applied to estimate animal densities from camera trap surveys (Zero et al. 2013). However, extending the REM method to other types of sensors (e.g., acoustic detectors) is more problematic, because the original derivation assumes a relatively narrow sensor width (up to $\pi/2$ radians) and that the animal is equally detectable irrespective of its heading (Rowcliffe et al. 2008).

Whilst these restrictions are not problematic for most camera trap makes (e.g., Reconyx, Cuddeback), the REM cannot be used to estimate densities from camera traps with a wider sensor width (e.g. canopy monitoring with fish eye lenses, (Brusa & Bunker 2014)). Additionally, the REM method is not useful in estimating densities from acoustic survey data as acoustic detector angles are often wider than $\pi/2$ radians. Acoustic detectors are designed for a range of diverse tasks

and environments (Kessel et al. 2014), which naturally leads to a wide range of sensor detection widths and detection distances. In addition to this, calls emitted by many animals are directional (Blumstein et al. 2011), breaking the assumption of the REM method.

There has been a sharp rise in interest around passive acoustic detectors in recent years, with a 10 fold increase in publications in the decade between 2000 and 2010 (Kessel et al. 2014). Acoustic monitoring is being developed to study many aspects of ecology, including the interactions of animals and their environments (Blumstein et al. 2011, Rogers et al. 2013), the presence and relative abundances of species (Marcoux et al. 2011), biodiversity of an area (Depraetere et al. 2012), and monitoring population trends (Walters et al. 2013).

Acoustic data suffers from many of the problems associated with data from camera trap surveys in that individuals are often unmarked, making capture-mark-recapture methods more difficult to use (Marques et al. 2013). In some cases the distance between the animal and the sensor is known, for example when an array of sensors is deployed and the position of the animal is estimated by triangulation (Lewis et al. 2007). In these situations distance-sampling methods can be applied (Buckland et al. 2008). However, in many cases distance estimation is not possible, for example when single sensors are deployed, a situation typical in the majority of terrestrial acoustic surveys (Buckland et al. 2008). In these cases, only relative measures of local abundance can be calculated, and not absolute densities. This means that comparison of populations between species and sites is problematic without assuming equal detectability (Schmidt 2003, Walters et al. 2013). Equal detectability is unlikely because of differences in environmental conditions, sensor type, habitat, and species biology.

In this study, we create a generalised REM (gREM) as an extension to the camera trap model of (Rowcliffe et al. 2008), to estimate absolute density from count data from acoustic detectors, or camera traps, where the sensor width can vary from 0 to 2π radians, and the signal given from the animal can be directional. We assessed the accuracy and precision of the gREM within a simulated environment, by varying the sensor detection widths, animal signal widths, number of captures and models of animal movement. We use the simulation results to recommend best survey practice for estimating animal densities from remote sensors.

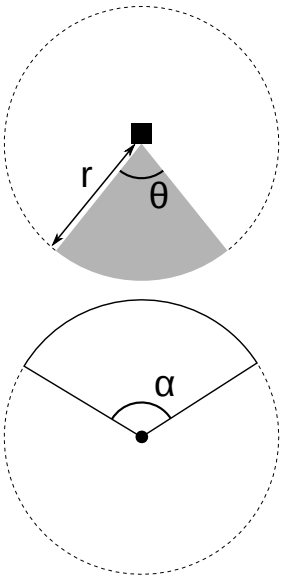


Figure 5.1 Representation of sensor detection width and animal signal width. The filled square and circle represent a sensor and an animal, respectively; θ, sensor detection width (radians); r, sensor detection distance; dark grey shaded area, sensor detection zone; α, animal signal width (radians). Dashed lines around the filled square and circle represents the maximum extent of θ and α, respectively.

5.3 Methods

5.3.1 Analytical Model

The REM presented by (Rowcliffe et al. 2008) adapts the gas model to count data collected from camera trap surveys. The REM is derived assuming a stationary sensor with a detection width less than $\pi/2$ radians. However, in order to apply this approach more generally, and in particular to stationary acoustic detectors, we need both to relax the constraint on sensor detection width, and allow for animals with directional signals. Consequently, we derive the gREM for any detection width, θ, between 0 and 2π with a detection distance r giving a circular sector within which animals can be captured (the detection zone) (Figure 5.1). Additionally, we model the animal as having an associated signal width α between 0 and 2π (Figure 5.1, see Appendix S1 for a list of symbols). We start deriving the gREM with the simplest situation, the gas model where $\theta = 2\pi$ and $\alpha = 2\pi$.

5.3.1.1 Gas Model

Following (Yapp 1956), we derive the gas model where sensors can capture animals in any direction and animal signals are detectable from any direction ($\theta = 2\pi$ and $\alpha = 2\pi$). We assume that animals are in a homogeneous environment, and move in straight lines of random direction with velocity v. We allow that our stationary sensor can capture animals at a detection distance r and that if an animal moves within this detection zone they are captured with a probability of one; while outside this zone, animals are never captured.

In order to derive animal density, we need to consider relative velocity from the reference frame of the animals. Conceptually, this requires us to imagine that all animals are stationary and randomly distributed in space, while the sensor moves with velocity v. If we calculate the area covered by the sensor during the survey period, we can estimate the number of animals the sensor should capture. As a circle moving across a plane, the area covered by the sensor per unit time is $2rv$. The expected number of captures, z, for a survey period of t, with an animal density of D is $z = 2rvtD$. To estimate the density we rearrange to get $D = z/2rvt$. Note that as z is the number of encounters, not individuals, the possibility of repeated detections of the same individual is accounted for (Hutchinson & Waser 2007).

5.3.1.2 gREM derivations for different detection and signal widths

Different combinations of θ and α would be expected to occur (e.g., sensors have different detection widths and animals have different signal widths). For different combinations θ and α, the area covered per unit time is no longer given by $2rv$. Instead of the size of the sensor detection zone having a diameter of $2r$, the size changes with the approach angle between the sensor and the animal. The width of the area within which an animal can be detected is called the profile, p. The size of p depends on the signal width, detector width and the angle that the animal approaches the sensor. The size of the profile (averaged across all approach angles) is defined as the average profile $\bar{p}$. However, different combinations of θ and α need different equations to calculate $\bar{p}$.

We have identified the parameter space for the combinations of θ and α for which the derivation of the equations are the same (defined as sub-models in the gREM) (Figure 5.2). For example, the gas model becomes the simplest gREM sub-model (upper right in Figure 5.2) and the REM from (Rowcliffe et al. 2008) is another gREM sub-model where $\theta < \pi/2$ and $\alpha = 2\pi$. We derive one gREM

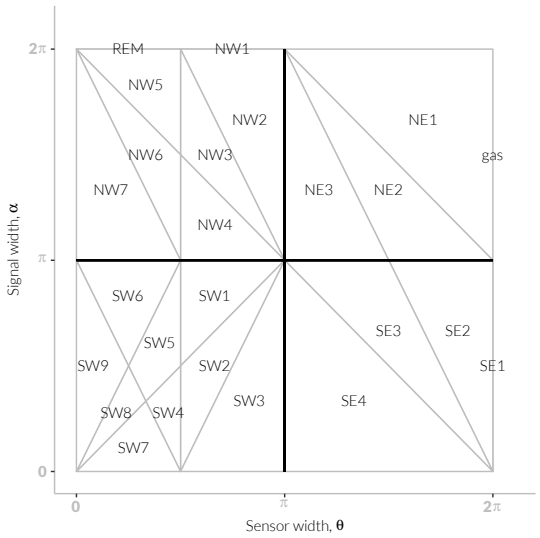


Figure 5.2 Locations where derivation of the average profile $\bar{p}$ is the same for different combinations of sensor detection and animal signal widths. Symbols within each polygon refer to each gREM submodel named after their compass point, except for Gas and REM which highlight the position of these previously derived models within the gREM. Symbols on the edge of the plot are for sub-models where $\alpha, \theta = 2\pi$

sub-model SE2 as an example below, where $2\pi - \alpha/2 < \theta < 2\pi,\ 0 < \alpha < \pi$ (see Appendix S2 for derivations of all gREM sub-models). Any estimate of density would require prior knowledge of animal velocity, v and animal signal width, α taken from other sources, for example existing literature (Brinkløv et al. 2011, Carbone et al. 2005). Sensor width, θ, and detection distance, r would also need to be measured or obtained from manufacturer specifications (Adams et al. 2012, Holderied & Von Helversen 2003).

5.3.1.3 Example derivation of SE2

In order to calculate $\bar{p}$, we have to integrate over the focal angle, x_1 (Figure 5.3a). This is the angle taken from the centre line of the sensor. Other focal angles are possible (x_2, x_3, x_4) and are used in other gREM sub-models (see Appendix S2). As the size of the profile depends on the approach angle, we present the derivation across all approach angles. When the sensor is directly approaching the animal $x_1 = \pi/2$.

Starting from $x_1 = \pi/2$ until $\theta/2 + \pi/2 - \alpha/2$, the size of the profile is $2r \sin \alpha/2$

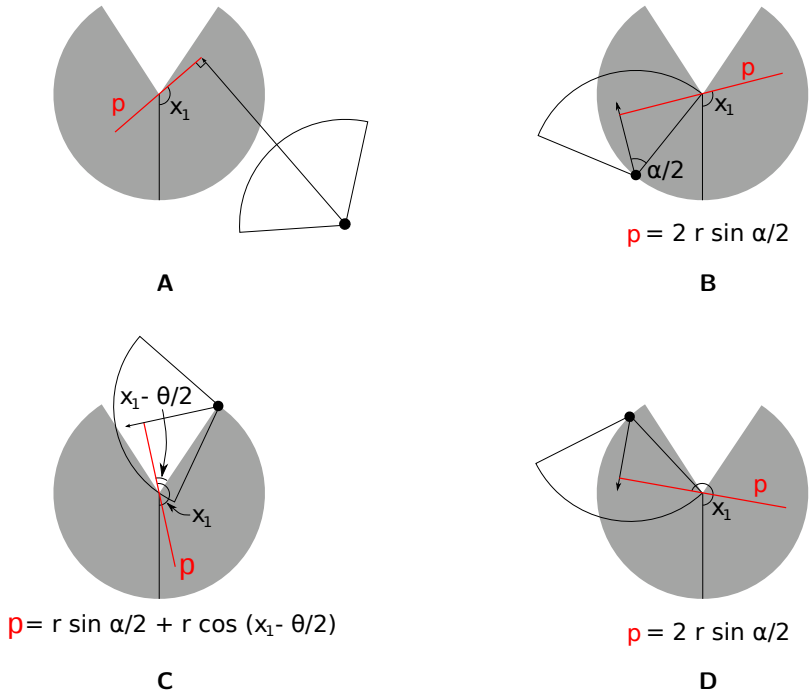


Figure 5.3 An overview of the derivation of the average profile $\bar{p}$ for the gREM submodel SE2, where (a) shows the location of the profile p (the line an animal must pass through in order to be captured) in red and the focal angle, x_1, for an animal (filled circle), its signal (unfilled sector), and direction of movement (shown as an arrow). The detection zone of the sensor is shown as a filled grey sector with a detection distance of r. The vertical black line within the circle shows the direction the sensor is facing. The derivation of p changes as the animal approaches the sensor from different directions (shown in b-d), where (b) is the derivation of p when x_1 is in the interval $[\frac{\pi}{2}, \frac{\pi}{2} + \frac{\theta}{2} - \frac{\alpha}{2}]$, (c) p when x_1 is in the interval $[\frac{\pi}{2} + \frac{\theta}{2} - \frac{\alpha}{2}, \frac{5\pi}{2} - \frac{\theta}{2} - \frac{\alpha}{2}]$ and (d) p when x_1 is in the interval $[\frac{5\pi}{2} - \frac{\theta}{2} - \frac{\alpha}{2}, \frac{3\pi}{2}]$, where θ, sensor detection width; α, animal signal width. The resultant equation for p is shown beneath b-d. The average profile $\bar{p}$ is the size of the profile averaged across all approach angles.

(Figure 5.3b). During this first interval, the size of α limits the width of the profile. When the animal reaches $x_1 = \theta/2 + \pi/2 - \alpha/2$ (Figure 5.3c), the size of the profile is $r\sin(\alpha/2) + r\cos(x_1 - \theta/2)$ and the size of θ and α both limit the width of the profile (Figure 5.3c). Finally, at $x_1 = 5\pi/2 - \theta/2 - \alpha/2$ until $x_1 = 3\pi/2$, the width of the profile is again $2r\sin\alpha/2$ (Figure 5.3d) and the size of α again limits the width of the profile.

The profile width p for π radians of rotation (from directly towards the sensor to directly behind the sensor) is completely characterised by the three intervals (Figure 5.3b–d). Average profile width $\bar{p}$ is calculated by integrating these profiles over their appropriate intervals of x_1 and dividing by π which gives

$$\bar{p} = \frac{1}{\pi}\left(\int_{\frac{\pi}{2}}^{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}} 2r\sin\frac{\alpha}{2}\,\mathrm{d}x_1 + \int_{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{5\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}} r\sin\frac{\alpha}{2} + r\cos\left(x_1-\frac{\theta}{2}\right)\mathrm{d}x_1 + \int_{\frac{5\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{3\pi}{2}} 2r\sin\frac{\alpha}{2}\,\mathrm{d}x_1\right) \tag{5.1}$$

$$= \frac{r}{\pi}\left(\theta\sin\frac{\alpha}{2} - \cos\frac{\alpha}{2} + \cos\left(\frac{\alpha}{2}+\theta\right)\right) \tag{5.2}$$

We then use this expression to calculate density

$$D = z/vt\bar{p}. \tag{5.3}$$

Rather than having one equation that describes $\bar{p}$ globally, the gREM must be split into submodels due to discontinuous changes in p as α and β change. These discontinuities can occur for a number of reasons such as a profile switching between being limited by α and θ, the difference between very small profiles and profiles of size zero, and the fact that the width of a sector stops increasing once the central angle reaches π radians (i.e., a semi-circle is just as wide as a full circle). As an example, if α is small, there is an interval between Figure 5.3c and 5.3d where the 'blind spot' would prevent animals being detected giving $p = 0$. This would require an extra integral in our equation, as simply putting our small value of α into 5.1 would not give us this integral of $p = 0$.

gREM submodel specifications were done by hand, and the integration was done using SymPy (SymPy Development Team 2014) in Python (Appendix S3). The gREM submodels were checked by confirming that: (1) submodels adjacent in parameter space were equal at the boundary between them; (2) submodels that border $\alpha = 0$ had $p = 0$ when $\alpha = 0$; (3) average profile widths $\bar{p}$ were between 0 and $2r$ and; (4) each integral, divided by the range of angles that it was integrated over, was between 0 and $2r$. The scripts for these tests are included in Appendix S3 and the R (R Development Core Team 2010) implementation of the gREM is given in Appendix S4.

5.3.2 Simulation Model

We tested the accuracy and precision of the gREM by developing a spatially explicit simulation of the interaction of sensors and animals using different combinations of sensor detection widths, animal signal widths, number of captures, and models of animal movement. One hundred simulations were run where each consisted of a 7.5 km by 7.5 km square with periodic boundaries. A stationary sensor of radius r, 10 m, was set up in the exact centre of each simulated study area, covering seven sensor detection widths θ, between 0 and 2π ($2/9\pi$, $4/9\pi$, $6/9\pi$, $8/9\pi$, $10/9\pi$, $14/9\pi$, and 2π). Each sensor was set to record continuously and to capture animal signals instantaneously from emission. Each simulation was populated with a density of 70 animals km^{-2}, calculated from the equation in (Damuth 1981) as the expected density of mammals weighing 1 g. This density therefore represents a reasonable estimate of density of individuals, given that the smallest mammal is around 2 g (Jones et al. 2009). A total of 3937 individuals per simulation were created which were placed randomly at the start of the simulation. 11 signal widths α between 0 and π were used ($1/11\pi$, $2/11\pi$, $3/11\pi$, $4/11\pi$, $5/11\pi$, $6/11\pi$, $7/11\pi$, $8/11\pi$, $9/11\pi$, $10/11\pi$, π).

Each simulation lasted for N steps (14400) of duration T (15 minutes) giving a total duration of 150 days. The individuals moved within each step with a distance d, with an average speed, v. The distance, d, was sampled from a normal distribution with mean distance, $\mu_d = vT$, and standard deviation, $\sigma_d = vT/10$, where the standard deviation was chosen to scale with the average distance travelled. An average speed, $v = 40\,\mathrm{km\,day}^{-1}$, was chosen based on the largest day range of terrestrial animals (Carbone et al. 2005), and represents the upper limit of realistic speeds. At the end of each step, individuals were allowed to either remain stationary for a time step (with a given probability, S), or change direction where the change in direction has a uniform distribution in the interval $[-A, A]$. This resulted in seven different movement models where: (1) simple movement, where S and $A = 0$; (2) stop-start movement, where (i) $S = 0.25$, $A = 0$, (ii) $S = 0.5$, $A = 0$, (iii) $S = 0.75$, $A = 0$; (3) correlated random walk movement, where (i) $S = 0$, $A = \pi/3$, (ii) $S = 0$, $A = 2\pi/3$, iii) $S = 0$, $A = \pi$. Individuals were counted as they moved into the detection zone of the sensor per simulation.

We calculated the estimated animal density from the gREM by summing the number of captures per simulation and inputting these values into the correct gREM submodel. The accuracy of the gREM was determined by comparing the true simulation density with the estimated density. Precision of the gREM

was determined by the standard deviation of estimated densities. We used this method to compare the accuracy and precision of all the gREM submodels. As these submodels are derived for different combinations of α and θ, the accuracy and precision of the submodels was used to determine the impact of different values of α and θ.

The influence of the number of captures and animal movement models on accuracy and precision was investigated using four different gREM submodels representative of the range α and θ values (submodels NW1, SW1, NE1, and SE3, Figure 5.2). From a random starting point we ran the simulation until a range of different capture numbers were recorded (from 10 to 100 captures), recorded the length of time this took, and estimated the animal density for each of the four sub-models. These estimated densities were compared to the true density to assess the impact on the accuracy and precision of the gREM. We calculated the coefficient of variation in order to compare the precision of the density estimates from simulations with different expected numbers of captures. The gREM also assumes that individuals move continuously with straight-line movement (simple movement model) and we therefore assessed the impact of breaking the gREM assumptions. We used the four submodels to compare the accuracy and precision of a simple movement model, stop-start movement models (using different average amounts of time spent stationary), and random walk movement models. Finally, as the parameters (α, β, r and v) are likely to be measured with error, we compared true simulation densities to densities estimated with parameters with errors of 0%, $\pm$5% and $\pm$10%, for all gREM submodels.

5.4 Results

5.4.1 Analytical model

The equation for $\bar{p}$ has been newly derived for each submodel in the gREM, except for the gas model and REM which have been calculated previously. However, many models, although derived separately, have the same expression for $\bar{p}$. Figure 5.4 shows the expression for $\bar{p}$ in each case. The general equation for density, 5.3, is used with the correct value of $\bar{p}$ substituted. Although more thorough checks are performed in Appendix S3, it can be seen that all adjacent expressions in Figure 5.4 are equal when expressions for the boundaries between them are substituted in.

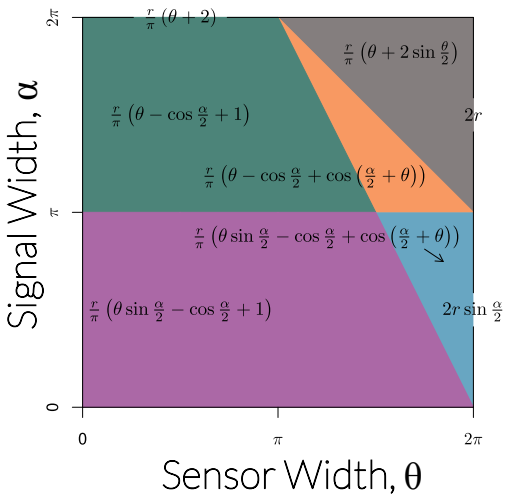


Figure 5.4 Expressions for the average profile width, $\bar{p}$, given a range of sensor and signal widths. Despite independent derivation within each block, many models result in the same expression. These are collected together and presented as one block of colour. Expressions on the edge of the plot are for submodels with $\alpha, \theta = 2\pi$.

5.4.2 Simulation model

5.4.2.1 gREM submodels

All gREM submodels showed a high accuracy, i.e., the median difference between the estimated and true values was less than 2% across all models (Figure 5.5). However, the precision of the submodels do vary, where the gas model is the most precise and the SW7 sub model the least precise, having the smallest and the largest interquartile range, respectively (Figure 5.5). The standard deviation of the error between the estimated and true densities is strongly related to both the sensor and signal widths (Appendix S5), such that larger widths have lower standard deviations (greater precision) due to the increased capture rate of these models.

5.4.2.2 Number of captures

Within the four gREM submodels tested (NW1, SW1, SE3, NE1), the accuracy was not strongly affected by the number of captures. The median difference between the estimated and true values was less than 15% across all capture rates (Figure 5.6). However, the precision was dependent on the number of captures

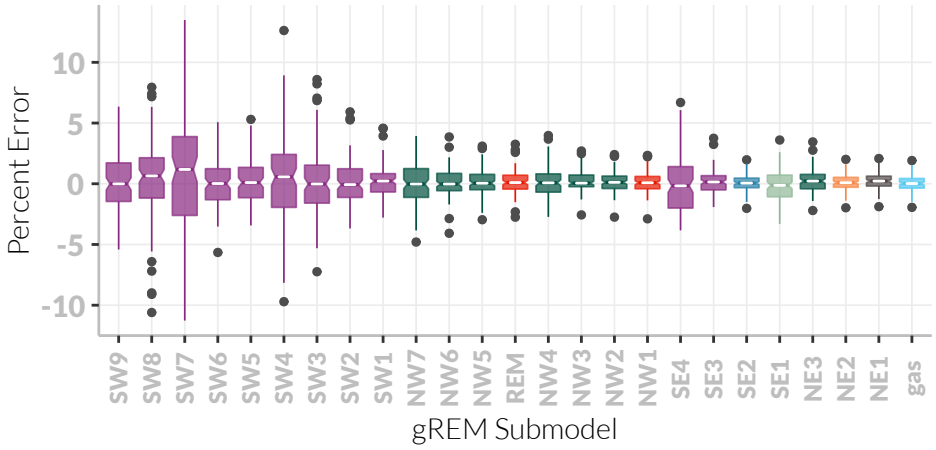


Figure 5.5 Simulation model results of the accuracy and precision for gREM submodels. The percentage error between estimated and true density for each gREM sub model is shown within each box plot, where the white line represents the median percentage error across all simulations, boxes represent the middle 50% of the data, whiskers represent variability outside the upper and lower quartiles with outliers plotted as individual points. Notches indicate 95% confidence intervals. Box colours correspond to the expressions for average profile width $\bar{p}$ given in Figure 5.4.

across all four of the gREM submodels, where precision increases as number of captures increases, as would be expected for any statistical estimate (Figure 5.6). For all gREM submodels, the the coefficient of variation falls to 10% at 100 captures.

5.4.2.3 Movement models

Within the four gREM submodels tested (NW1, SW1, SE3, NE1), neither the accuracy or precision was affected by the average amount of time spent stationary. The median difference between the estimated and true values was less than 2% for each category of stationary time (0, 0.25, 0.5 and 0.75) (Figure 5.7). Altering the maximum change in direction in each step (0, $\pi/3$, $2\pi/3$, and π) did not affect the accuracy or precision of the four gREM submodels (Figure 5.7).

5.4.2.4 Impact of parameter error

The percentage error in the density estimates across all parameters and gREM submodels shows a similar response for under and over estimated parameters,

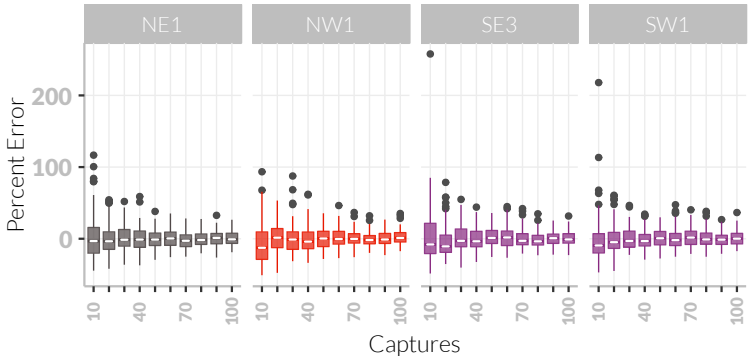


Figure 5.6 Simulation model results of the accuracy and precision of four gREM submodels (NW1, SW1, SE3 and NE1) given different numbers of captures. The percentage error between estimated and true density within each gREM sub model for capture rate is shown within each box plot, where the white line represents the median percentage error across all simulations, boxes represent the middle 50% of the data, whiskers represent variability outside the upper and lower quartiles with outliers plotted as individual points. Notches show the 95% confidence interval. Sensor and signal widths vary between submodels. The numbers beneath each plot represent the coefficient of variation. The colour of each box plot corresponds to the expressions for average profile width $\bar{p}$ given in Figure 5.4.

suggesting the accuracy is reasonable with respect to parameter error (Appendix S6). The impact of parameter error on the precision of the density estimate varies across gREM submodels and parameters, where α shows the largest variation including the largest values. However, in all cases the percentage error in the density estimate is not more than 5% greater than the error in the parameter estimate (Appendix S6).

5.5 Discussion

5.5.1 Analytical model

We have developed the gREM such that it can be used to estimate density from acoustic sensors and camera traps. This has entailed a generalisation of the gas model and the REM in (Rowcliffe et al. 2008) to be applicable to any combination of sensor width θ and signal directionality α. We emphasise that the approach is robust to multiple detections of the same individual. We have used simulations to show, as a proof of principle, that these models are accurate and precise.

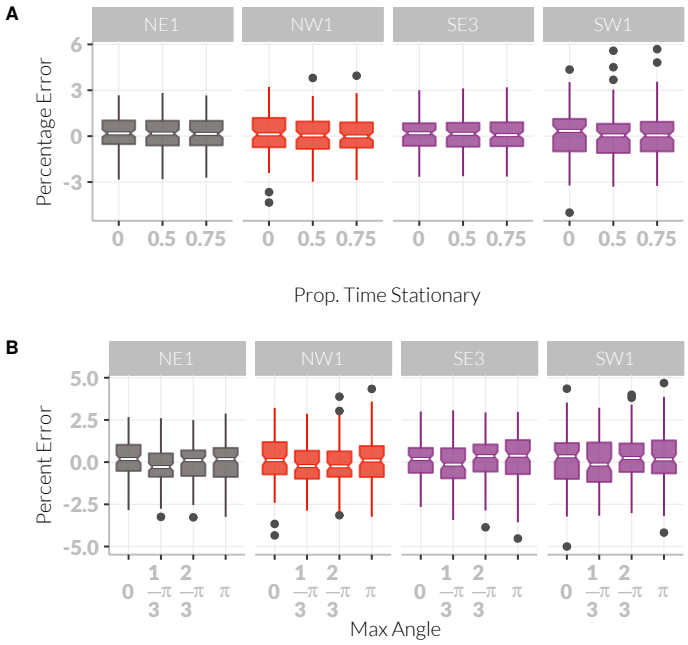


Figure 5.7 Simulation model results of the accuracy and precision of four gREM submodels (NW1, SW1, SE3 and NE1) given different movement models where (a) average amount of time spent stationary (stop-start movement) and (b) maximum change in direction at each step (correlated random walk model). The percentage error between estimated and true density within each gREM sub model for the different movement models is shown within each box plot, where the white line represents the median percentage error across all simulations, boxes represent the middle 50% of the data, whiskers represent variability outside the upper and lower quartiles with outliers plotted as individual points. Notches in boxplots show the 95% confidence for the median. The simple model is represented where time and maximum change in direction equals 0. The colour of each box plot corresponds to the expressions for average profile width $\bar{p}$ given in Figure 4.

There are a number of possible extensions to the gREM which could be developed in the future. The original gas model was formulated for the case where both animals and sensor are moving (Hutchinson & Waser 2007). Indeed any of the models which have animals that are equally detectable in all directions ($\alpha = 2\pi$) can be trivially expanded by replacing animal speed v with $v + v_s$ where v_s is the speed of the sensor. However, when the animal has a directional call the extension becomes less simple. The approach would be to calculate again the mean profile width. However, for each angle of approach, one would have to average the profile width for an animal facing in any direction (i.e., not ne-

cessarily moving towards the sensor) weighted by the relative velocity of that direction. There are a number of situations where a moving detector and animal could occur, e.g. an acoustic detector towed from a boat when studying porpoises (Kimura et al. 2014) or surveying echolocating bats from a moving car (Jones et al. 2011).

Interesting but unstudied problems impacting the gREM are firstly, edge effects caused by sensor trigger delays (the delay between sensing an animal and attempting to record the encounter) (Rovero et al. 2013), and secondly, sensors which repeatedly turn on an off during sampling (Jones et al. 2011). The second problem is particularly relevant to acoustic detectors which record ultrasound by time expansion. Here ultrasound is recorded for a set time period and then slowed down and played back, rendering the sensor 'deaf' periodically during sampling. Both of these problems may cause biases in the gREM, as animals can move through the detection zone without being detected. As the gREM assumes constant surveillance, the error created by switching the sensor on and off quickly will become more important if the sensor is only on for short periods of time. We recommend that the gREM is applied to constantly sampled data, and the impacts of breaking these assumptions on the gREM should be further explored.

5.5.2 Accuracy, Precision and Recommendations for Best Practice

Based on our simulations, we believe that the gREM has the potential to produce accurate estimates for many different species, using either camera traps or acoustic detectors. However, the precision of the gREM differed between submodels. For example, when the sensor and signal width were small, the precision of the model was reduced. Therefore when choosing a sensor for use in a gREM study, the sensor detection width should be maximised. If the study species has a narrow signal directionality, other aspects of the study protocol, such as length of the survey, should be used to compensate.

The precision of the gREM is greatly affected by the number of captures. The coefficient of variation falls dramatically between 10 and 60 captures and then after this continues to slowly reduce. At 100 captures the submodels reach 10% coefficient of variation, considered to be a very good level of precision and better than many previous studies (Foster & Harmsen 2012, O'Brien et al. 2003, Thomas & Marques 2012). The length of surveys in the field will need to be adjusted so that enough data can be collected to reach this precision level. Populations of fast

moving animals or populations with high densities will require less survey effort than those species that are slow moving or have populations with low densities.

We found that the sensitivity of the gREM to inaccurate parameter estimates was both predictable and reasonable (Appendix S6), although this varies between different parameters and gREM submodels. Whilst care should be taken in parameter estimation when analysing both acoustic and camera trap data, acoustic data poses particular problems. For acoustic surveys, estimates of r (detection distance) can be measured directly or calculated using sound attenuation models (Holderied & Von Helversen 2003), while the sensor angle is often easily measured (Adams et al. 2012) or found in the manufacturer's specifications. When estimating animal movement speed v, only the speed of movement during the survey period should be used. The signal width is the most sensitive parameter to inaccurate estimates (Appendix S6) and is also the most difficult to measure. While this parameter will typically be assumed to be 2π for camera trap surveys, fewer estimates exist for acoustic signal widths. Although signal width has been measured for echolocating bats using arrays of microphones (Brinkløv et al. 2011), more work should be done on obtaining estimates for a range of acoustically surveyed species.

5.5.3 Limitations

Although the REM has been found to be effective in field tests (Rowcliffe et al. 2008, Zero et al. 2013), the gREM requires further validation by both field tests and simulations. For example, capture-mark-recapture methods could be used alongside the gREM to test the accuracy under field conditions (Rowcliffe et al. 2008). While we found no effect of the movement model on the accuracy or precision of the gREM, the models we have used in our simulations to validate the gREM are still simple representations of true animal movement. Animal movement may be highly nonlinear and often dependent on multiple factors such as behavioural state and existence of home ranges (Smouse et al. 2010). Therefore testing the gREM against real animal data, or further simulations with more complex movement models, would be beneficial.

The assumptions of our simulations may require further consideration, for example we have assumed an equal density across the study area. However, in a field environment the situation may be more complex, with additional variation coming from local changes in density between sensor sites. Athough unequal densities should theoretically not affect accuracy (Hutchinson & Waser 2007),

it will affect precision and further simulations should be used to quantify this effect. Additionally, we allowed the sensor to be stationary and continuously detecting, negating the triggering, and non-continuous recording issues that could exist with some sensors and reduce precision or accuracy. Finally, in the simulation animals moved at the equivalent of the largest day range of terrestrial animals (Carbone et al. 2005). Slower speed values should not alter the accuracy of the gREM, but precision would be affected since slower speeds produce fewer records. The gREM was both accurate and precise for all the movement models we tested (stop-start movement and correlated random walks).

A feature of the gREM is that it does not fit a statistical model to estimate detection probability as occupancy models and distance sampling do (Barlow & Taylor 2005, Marques et al. 2011, Royle & Nichols 2003). Instead it explicitly models the process, with animals only being detected if they approach the sensor from a suitable direction. Other processes that affect detection probability could be included in the model to improve realism.

5.5.4 Implications for ecology and conservation

The gREM is applicable for count data obtained either visually or acoustically in both marine and terrestrial environments, and is suitable for taxa including echolocating bats (Walters et al. 2012), songbirds (Buckland & Handel 2006), whales (Marques et al. 2011) and forest primates (Hassel-Finnegan et al. 2008). Many of these taxa contain critically endangered species and monitoring their populations is of conservation interest. For example, current methods of density estimation for the threatened Franciscana dolphin (*Pontoporia blainvillei*) may result in underestimation of their numbers (Crespo et al. 2010). In addition, using gREM may be easier than other methods for measuring the density of animals which may be useful in quantifying ecosystem services, such as songbirds with a known positive influence on pest control (Jirinec et al. 2011).

The gREM will aid researchers to study species with non-invasive methods such as remote sensors, which allows for large, continuous monitoring projects with limited human resources (Kelly et al. 2012). The gREM is also suitable for species that are sensitive to human contact or are difficult or dangerous to catch (Thomas & Marques 2012). As sensors such as camera traps and acoustic detectors become more ubiquitous, the gREM will be increasingly useful for monitoring unmarked animal populations across broad spatial, temporal and taxonomic scales.

6 Discussion

6.1 What I did

- In this thesis I aimed to test the importance of population structure and density on pathogen diversity
- With a particular focus on bats
- Combining simulations and empirical studies.
- Identify that population structure does affect pathogen richness in wild bats.
- But found that invasision of new pathogens is probably not the mechanism.
- Clarified the relationships between population size and density, range size, colony size and population structure.
- And found that colony size is more important than density per se.
- Finally, created a method to more easily estimate bat population densities.

6.2 How I did it (Chapters overview)

1.
   - I tested the hypothesis that population structure predicts viral richness in wild bats.
   - I used two measurements of population structure.
     a) A novel measure, number of subspecies. Largest dataset to date.
     b) Gene flow, dealing with issues of marker type and spatial scale
   - I used multivariate regression, appraised with information theory techniques.
   - I controlled for phylogeny.
   - I found that in both analyses, increased population structure predicts increased pathogen richness and is in best model.
2.
   - I modelled a multi-pathogen, metapopulation based on bat populations.
   - Testing the specific mechanism that population structure increases richness by enabling invasion of newly evolved pathogens.

- I found that spatial structure, either by dispersal rate or topology, did not allow invasion.

3.
- I clarified confusion on the relationships between group size, group number, density, population size and range size.
- Using same model as Chapter 3 I tested whether it is in fact density or population size that matters.
- I tested whether the importact factor in increased density is group size or group number.
- I found that it is in fact abundance and group size that matter much more than other factors.

4.
- I aimed to collect data on bat density as lit. says this is important.
- Discovered I needed a model.
- Started with specific model for iBats, ended up writing general model.
- I formulated a model that estimates density from acoustic sensors or camera traps.
- I tested it using simulations.
- I found it to be precise and unbiased.

6.3 What agreed/disagreed with the literature.

- Don't find high R_0 leads to high richness in contrast to eco assumption.
- Don't find structure makes any difference in contrast to poletto and (Nunes et al. 2006) and ackleh.
- Find that colony size is more important which implies density is just proxy for group size. Group size *is* density at small scale.
- Importance of colony size agrees with nunn primate papers.

6.4 What are the implications for research?

- Global change and pop structure.

- Structure is not a strong enough predictor to use for zoonotic surveillance. Perhaps colony size is.
- Studies should more carefully consider density vs structure vs group size.
- Population structure should be studied in other groups other than bats.
- Can more easily estimate bat density.

6.5 Furtherwork and limitations

- Collect data for colony size and test importance against structure.
- Limitations of poor data, escpecially for gene flow.
- Examine other mechanisms for richness
- Examine multi host species more carefully
- Field test gREM
- Use gREM to collect density estimates

6.6 Conclusions

- Population structure does influence pathogen richness but the mechanisms are still unclear.
- Local dynamics (local density) are most important for pathogen invasions not broad scale structure.
- Data on density should be collected using the gREM.

A gREM Appendix

A.1 Table of symbols

Table A.1 List of symbols used to describe the gREM and simulations. '-' means the quantity has no units.

Symbol	Description	Units
θ	Sensor width	rad
α	Animal signal width	rad
x_i	Focal angle, $i \in \{1, 2, 3, 4\}$	rad
r	Detection distance	m
$\bar{p}$	Average profile width	m
p	A specific profile width	m
v	Velocity	$\mathrm{m\,s^{-1}}$
t	Time	s
z	Number of detections	-
D	Animal density	$\mathrm{m^{-2}}$
T	Step length	s
N	Number of steps per simulation	-
d	Distance moved in a time step	m
S	Probability of remaining stationary	-
A	Maximum turning angle	rad

A.2 Supplementary Methods

A.2.1 Introduction

These supplementary methods derive all the models used. For continuity, the gas model derivation is included here as well as in the main text. The calculation of all integrals used in the gREM is included in the Python script S3.

A.2.2 Gas model

Following (Yapp 1956), we derive the gas model where sensors can capture animals in any direction and animal signals are detectable from any direction ($\theta = 2\pi$ and $\alpha = 2\pi$). We assume that animals are in a homogeneous environment, and move in straight lines of random direction with velocity v. We allow that our stationary sensor can capture animals at a detection distance r and that if an animal moves within this detection zone they are captured with a probability of one, while animals outside the zone are never captured.

In order to derive animal density, we need to consider relative velocity from the reference frame of the animals. Conceptually, this requires us to imagine that all animals are stationary and randomly distributed in space, while the sensor moves with velocity v. If we calculate the area covered by the sensor during the survey period we can estimate the number of animals the sensor should capture. As a circle moving across a plane, the area covered by the sensor per unit time is $2rv$. The number of expected captures, z, for a survey period of t, with an animal density of D is $z = 2rvtD$. To estimate the density, we rearrange to get $D = z/2rvt$.

A.2.2.1 gREM derivations for different detection and signal widths

Different combinations of θ and α would be expected to occur (e.g., sensors have different detection widths and animals have different signal widths). For different combinations θ and α, the area covered per unit time is no longer given by $2rv$. Instead of the size of the sensor detection zone having a diameter of $2r$, the size changes with the approach angle between the sensor and the animal. For any given signal width and detector width and depending on the angle that the animal approaches the sensor, the width of the area within which an animal can be detected is called the profile, p. The size of the profile (averaged across all approach angles) is defined as the average profile $\bar{p}$. However, different combin-

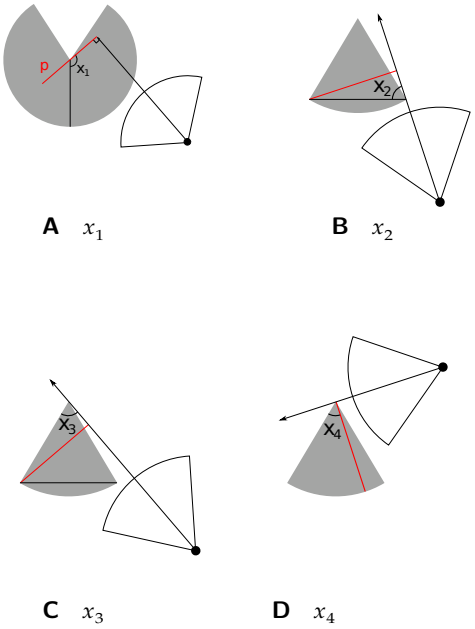


Figure A.1 The location of the focal angles $x_{i\in[1,4]}$. x_1 is used in SE and NE models (including the gas model). $x_2 - x_4$ are used in NW and SW models. The sector shaped detection region is shown in grey. Animals are filled black circles and the animal signal is an unfilled sector. The animals direction of movement is indicated with an arrow. The profile p is shown with a red line. (a) Animal is directly approaching the sensor at $x_1 = \frac{\pi}{2}$. (b) Animal is directly approaching the sensor at $x_2 = \frac{\pi}{2}$. x_2 then decreases until the profile is perpendicular to the edge of the detection region. (c) When the profile is perpendicular to the edge of the detection region, $x_3 = \theta$. (d) x_4 measures the angle between the left side of the detection region and the profile.

ations of θ and α need different equations to calculate $\bar{p}$. This $\bar{p}$ is the only thing that changes

We have identified the parameter space for the combinations of θ and α for which the derivation of the equations are the same (defined as sub-models in the gREM) (Figure 5.2). For example, the gas model becomes the simplest gREM sub-model (upper right in Figure 5.2) and the REM from (Rowcliffe et al. 2008) is another gREM sub-model where $\theta < \pi/2$ and $\alpha = 2\pi$.

Models with $\theta = 2\pi$ are described first (the gas model described above and SE1). Then models with $\theta > \pi$ are described (NE then SE). Finally models with

$\theta < \pi$ (NW then SW) are described.

A.2.3 Model SE1

SE1 is very similar to the gas model except that because $\alpha \leq \pi$ the profile width is no longer $2r$ but is instead limited by the width of the animal signal. We therefore get a profile width of $2r\sin(\alpha/2)$ instead.

$$\bar{p}_{\text{SE1}} = \frac{1}{\pi} \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} 2r \sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_1 \tag{A.1}$$

$$\bar{p}_{\text{SE1}} = 2r \sin\left(\frac{\alpha}{2}\right) \tag{A.2}$$

This profile is integrated over the interval $[\frac{\pi}{2}, \frac{3\pi}{2}]$ which is π radians of rotation starting with the animal moving directly towards the sensor (Figure A.1a).

A.2.4 Models NE1–3

When the detection zone is not a circle, we have more complex profiles and need to explicitly write functions for the width of the profile for every approach angle. We then use these functions to find the average profile width $\bar{p}$ for all approach angles by integrating across all 2π angles of approach and dividing by 2π.

There are three submodels within quadrant NE (Figure 5.2). Note that NE1 covers the area $\alpha = 2\pi$ as well as the triangle below it as these two models are specified exactly the same, rather than happening to have equal results.

These models have up to five profiles.

1. The profile width starts, from $x_1 = \frac{\pi}{2}$ as $2r$.

2. At $x_1 = \theta/2$, the right hand side of the profile cannot be r wide as the corner of the 'blind spot' limits its size to being $r\cos(x_1 - \theta/2)$ wide (Figure A.2A).

3. The third profile is only found in NE3. If $\alpha < 4\pi - 2\theta$, then at $x_1 = \theta/2 + \pi/2$, when the profile is perpendicular to the edge of the blind spot, the whole right side of the profile is invisible to the sensor (Figure A.2B). This gives a profile size of just r.

4. At some point, the sensor can detect animals once they have passed the blind spot giving a profile width of $r + r\cos(x_1 + \theta/2)$ (Figure A.2C). From

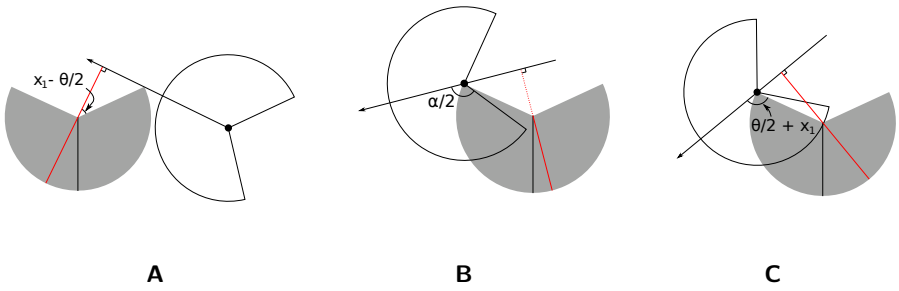


Figure A.2 Three of the integrals in NE models. The sector shaped detection region is shown in grey. Animals are filled black circles and the animal signal is an unfilled sector. The animals direction of movement is indicated with an arrow. The profile p is shown with a red line. Dashed red lines indicate areas where animals cannot be detected. (a) The second integral in NE with width $r + r\cos(x_1 - \theta/2)$. (b) The third integral in NE3. $\alpha/2$ is labelled. As it is small, animals to the right of the detector cannot be detected. (c) After further rotation, $\alpha/2$ is now bigger than the angle shown and animals to the right of the detector can again be detected.

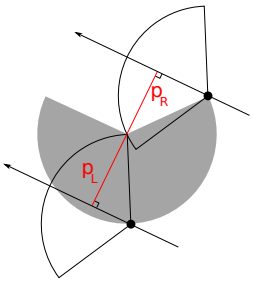


Figure A.3 The second integral in SE. The right side of the profile (p_R) is limited by the size of the sensor region while the left side of the profile (p_L) is limited by the size of the signal width. The full profile has width $p = r\sin(\alpha/2) + r\cos(\theta/2 - x_1)$. The sector shaped detection region is shown in grey. Animals are filled black circles and the animal signal is an unfilled sector. The animals direction of movement is indicated with an arrow. The profile p is shown with a red line.

$x_1 = \pi$, if the animal signal is wide enough to be detected in this area, this is the wider profile. This then defines the split between NE1 and NE2. In NE1, with $\alpha > 3\pi - \theta$, the animal signal is wide enough that at $x_1 = \pi$ the animal can immediately be detected past the blind spot and so this profile is used. In NE2, with $\alpha < 3\pi - \theta$, the latter profile is reached at $5\pi/2 - \theta/2 - \alpha/2$.

5. Finally, common to all three models, at $x_1 = 2\pi - \theta/2$ the profile becomes a full $2r$ once again.

A.2.4.1 Model NE1

Submodel NE1 exists within the area bounded by $\alpha \leq 2\pi$, $\theta \leq 2\pi$ and $\alpha \geq 3\pi - \theta$ (Figure 5.2). It has four profiles; it does not include the r profile at $x_1 = \pi$ (profile described in point (3) in Section A.2.4). Furthermore, θ is wide enough that the $r + r\cos(x_1 + \theta/2)$ profile starts at π. This then gives us

$$\begin{aligned}\bar{p}_{\mathrm{NE1}} =& \frac{1}{\pi}\left(\int_{\frac{\pi}{2}}^{\frac{\theta}{2}} 2r\,\mathrm{d}x_1 + \int_{\frac{\theta}{2}}^{\pi} r\cos\left(\frac{\theta}{2} - x_1\right) + r\,\mathrm{d}x_1 \right. \\ &\left. + \int_{\pi}^{2\pi-\frac{\theta}{2}} r\cos\left(\frac{\theta}{2} + x_1\right) + r\,\mathrm{d}x_1 + \int_{2\pi-\frac{\theta}{2}}^{\frac{3\pi}{2}} 2r\,\mathrm{d}x_1 \right) \end{aligned} \tag{A.3}$$

$$\bar{p}_{\mathrm{NE1}} = \frac{r}{\pi}\left(\theta + 2\sin\left(\frac{\theta}{2}\right)\right) \tag{A.4}$$

A.2.4.2 Model NE2

Model NE2 is bounded by $\alpha \leq 3\pi - \theta$, $\alpha \geq 4\pi - 2\theta$ and $\alpha \geq \pi$ (Figure 5.2). It is the same as NE1 except that the third profile starts at $5\pi/2 - \theta/2 - \alpha/2$ instead of at π which is reflected in the different bounds in the second and third integral.

$$\begin{aligned}\bar{p}_{\mathrm{NE2}} =& \frac{1}{\pi}\left(\int_{\frac{\pi}{2}}^{\frac{\theta}{2}} 2r\,\mathrm{d}x_1 + \int_{\frac{\theta}{2}}^{\frac{5\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}} r\cos\left(\frac{\theta}{2} - x_1\right) + r\,\mathrm{d}x_1 \right. \\ &\left. + \int_{\frac{5\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}}^{2\pi-\frac{\theta}{2}} r\cos\left(\frac{\theta}{2} + x_1\right) + r\,\mathrm{d}x_1 + \int_{2\pi-\frac{\theta}{2}}^{\frac{3\pi}{2}} 2r\,\mathrm{d}x_1 \right) \end{aligned} \tag{A.5}$$

$$\bar{p}_{\mathrm{NE2}} = \frac{r}{\pi}\left(\theta - \cos\left(\frac{\alpha}{2}\right) + \cos\left(\frac{\alpha}{2} + \theta\right)\right) \tag{A.6}$$

A.2.4.3 Model NE3

Model NE3 is bound by $\alpha \leq 4\pi - 2\theta$, $\alpha \geq \pi$ and $\theta \geq \pi$ (Figure 5.2). It is the same as NE2 except that it contains the extra profile with width r (third integral).

$$\bar{p}_{\text{NE3}} = \frac{1}{\pi}\left(\int_{\frac{\pi}{2}}^{\frac{\theta}{2}} 2r\,\mathrm{d}x_1 + \int_{\frac{\theta}{2}}^{\frac{\theta}{2}+\frac{\pi}{2}} r\cos\left(\frac{\theta}{2} - x_1\right) + r\,\mathrm{d}x_1 \right.$$
$$\left. + \int_{\frac{\theta}{2}+\frac{\pi}{2}}^{\frac{5\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}} r\,\mathrm{d}x_1 + \int_{\frac{5\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}}^{2\pi-\frac{\theta}{2}} r\cos\left(\frac{\theta}{2} + x_1\right) + r\,\mathrm{d}x_1 + \int_{2\pi-\frac{\theta}{2}}^{\frac{3\pi}{2}} 2r\,\mathrm{d}x_1 \right) \tag{A.7}$$

$$\bar{p}_{\text{NE3}} = \frac{r}{\pi}\left(\theta - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.8}$$

A.2.5 Models SE2–4

Quadrant SE contains three submodels excluding SE1 (Figure 5.2). The differences between these three models are similar to the differences between the models in NE. There are four possible profiles.

1. As α is less than π the profile is smaller than $2r$, even when the sensor width is a full diameter. The profile width starts as $2r\sin(\alpha/2)$.
2. Similar to NE, at a certain point the blind spot of the sensor area limits the profile width on one side. This gives a profile width of $r\sin(\alpha/2) + r\cos(x_1 - \theta/2)$ (Figure A.3).
3. Also similar to NE, there can be a point where the right side of the profile is 0 giving a profile width of $r\sin(\alpha/2)$.
4. If $\alpha \leq 2\pi - \theta$, then at $x_1 = \theta/2 + \pi/2 + \alpha/2$ the profile width becomes 0. This inequality distinguishes between SE3 and SE4.
5. The third profile $r\sin(\alpha/2)$ starts at $\theta/2 + \pi/2$ while at $5\pi/2 - \alpha/2 - \theta/2$ the profile returns to size $2r\sin(\alpha/2)$. If $\theta/2 + \pi/2 \geq 5\pi/2 - \alpha/2 - \theta/2$ we go straight into the $2r\sin(\alpha/2)$ profile and miss the $r\sin(\alpha/2)$ profile. SE2 and SE3 are separated by this inequality which simplifies to $\alpha \leq 4\pi - 2\theta$.

A.2.5.1 Model SE2

SE2 is bounded by $\alpha \geq 4\pi - 2\theta$, $\alpha \leq \pi$ and $\theta \leq 2\pi$ (Figure 5.2). As $\alpha \geq 4\pi - 2\theta$, there is no $r\sin(\alpha/2)$ profile. As $\alpha \leq 4\pi - 2\theta$, the profile returns to $2r\sin(\alpha/2)$ rather than going to 0. These integrals relate to profiles (1), (2) and (5) in Section A.2.5.

$$\bar{p}_{\mathrm{SE2}} = \frac{1}{\pi}\left(\int_{\frac{\pi}{2}}^{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}} 2r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_1 + \int_{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{5\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}} r\sin\left(\frac{\alpha}{2}\right) + r\cos\left(\frac{\theta}{2}-x_1\right) \mathrm{d}x_1 \right.$$
$$\left. + \int_{\frac{5\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{3\pi}{2}} 2r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_1 \right) \tag{A.9}$$

$$\bar{p}_{\mathrm{SE2}} = \frac{r}{\pi}\left(\theta\sin\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\alpha}{2}\right) + \cos\left(\frac{\alpha}{2}+\theta\right)\right) \tag{A.10}$$

A.2.5.2 Model SE3

SE3 is bounded by $4\pi - 2\theta \leq \alpha \leq 4\pi - 2\theta$ and $\alpha \leq \pi$ (Figure 5.2). Therefore there is a $r\sin(\alpha/2)$ profile but no $0r$ profile. This relates to profiles (1), (2), (3) and (5) above.

$$\bar{p}_{\mathrm{SE3}} = \frac{1}{\pi}\left(\int_{\frac{\pi}{2}}^{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}} 2r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_1 + \int_{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{\theta}{2}+\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right) + r\cos\left(\frac{\theta}{2}-x_1\right) \mathrm{d}x_1 \right.$$
$$\left. + \int_{\frac{\theta}{2}+\frac{\pi}{2}}^{\frac{5\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}} r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_1 + \int_{\frac{5\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{3\pi}{2}} 2r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_1 \right) \tag{A.11}$$

$$\bar{p}_{\mathrm{SE3}} = \frac{r}{\pi}\left(\theta\sin\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.12}$$

A.2.5.3 Model SE4

Finally SE4 is bounded by $\alpha \leq 4\pi - 2\theta$, $\alpha \leq \pi$ and $\theta \leq \pi$ (Figure 5.2). It is the same as SE3 except that the profile becomes 0 rather than returning to $2r\sin(\alpha/2)$. This relates to profiles (1), (2), (3) and (4) above though profile (4) with width 0 is not shown.

$$\bar{p}_{\mathrm{SE4}} = \frac{1}{\pi}\left(\int_{\frac{\pi}{2}}^{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}} 2r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_1 + \int_{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{\theta}{2}+\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right) + r\cos\left(\frac{\theta}{2}-x_1\right) \mathrm{d}x_1 + \int_{\frac{\theta}{2}+\frac{\pi}{2}}^{\frac{\alpha}{2}+\frac{\theta}{2}+\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_1 \right) \tag{A.13}$$

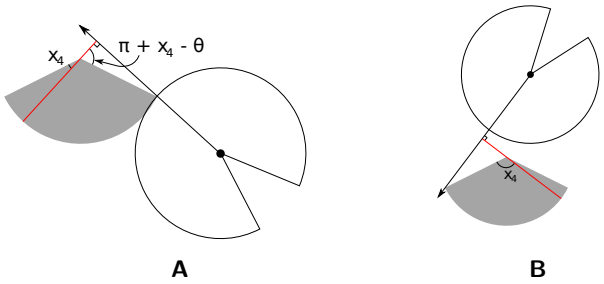


Figure A.4 The second and fourth profiles of NW1. The left side of of both profiles is of width r while the right side differs. (a) The right side of the profile is $r\cos(\pi + x_4 - \theta) = -r\cos(\theta - x_4)$ (b) The right side is $r\cos(\pi - x_4) = -r\cos x_4$ respectively. In both images the sector shaped detection region is shown in grey. Animals are filled black circles and the animal signal is an unfilled sector. The animals direction of movement is indicated with an arrow. The profile p is shown with a red line.

$$\bar{p}_{\text{SE4}} = \frac{r}{\pi}\left(\theta \sin\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.14}$$

A.2.6 Model NW1

NW1 is the first model with $\theta < \pi$. Whereas previously the focal angle has always been x_1, we now use different focal angles. x_2 and x_3 correspond to γ_1 and γ_2 in (Rowcliffe et al. 2008) while x_4 is new. They are described in Figure A.1b–d.

There are five different profiles in NW1.

1. x_2 has an interval of $[\pi/2, \theta/2]$ which is from the angle of approach being directly towards the sensor until the profile is parallel to the left hand radius of the sensor sector (Figure A.1B). During this interval the profile width is $2r\sin(\theta/2)\sin(x_2)$ which is calculated using the equation for the length of a chord . Note that while rotating anti-clockwise (as usual) x_2 decreases in size.

2. From here, we examine focal angle x_4 (note that x_3 is used in later models, but is not relevant here.) The left side of the profile is a full radius while the right side is limited to $-r\cos(x_4 - \theta)$ (Figure A.4A).

3. At $x_4 = \theta - \pi/2$, the profile is perpendicular to the edge of the sensor area. Here, the right side of the profile is $0r$ giving a profile size of r.

4. When $x_4 = \pi/2$ the angle of approach is from behind the sensor, but we can once again be detected on the right side of the sensor (Figure A.4B). Therefore the width of the profile is $r - r\cos(x_4)$.

5. Finally, we have the x_2 profile, but from behind.

$$\begin{aligned}\bar{p}_{\mathrm{NW1}} =& \frac{1}{\pi}\left(\int_{\frac{\theta}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin\left(x_2\right)\,\mathrm{d}x_2 + \int_{0}^{\theta-\frac{\pi}{2}} r - r\cos\left(-x_4+\theta\right)\,\mathrm{d}x_4 \right.\\ &\left. + \int_{\theta-\frac{\pi}{2}}^{\frac{\pi}{2}} r\,\mathrm{d}x_4 + \int_{\frac{\pi}{2}}^{\theta} r - r\cos\left(x_4\right)\,\mathrm{d}x_4 + \int_{\frac{\theta}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin\left(x_2\right)\,\mathrm{d}x_2 \right) \qquad \text{(A.15)}\\ \bar{p}_{\mathrm{NW1}} =& \frac{r}{\pi}\left(\theta + 2\right) \qquad \text{(A.16)}\end{aligned}$$

A.2.7 Models NW2–4

The models NW2–4 have the five potential profiles in NW1 but not all profiles occur in each model, and the angle at which transitions occur are different. Furthermore, there is one extra profile possible.

1. When approaching the sensor from behind, there is a period where the profile is r wide as in NW1 profile (3).

2. At some point after profile (1) animals to the right of the sensor can be detected again. If this occurs in the x_4 region, the profile width becomes $r - r\cos(x_4)$ as in NW1.

3. However, as α is now less than 2π, animals to the right of the sensor may be undetectable until we are in the second x_2 region. In this case, when we first enter the second x_2 region, the profile has a width of $r\cos(x_2 - \theta/2)$. This occurs only if $\alpha \leq 3\pi - 2\theta$. This inequality is found by noting that animals to the right of the sensor can be detected again at $x_4 = 3\pi/2 - \alpha$ but the x_2 region starts at $x_4 = \theta$. The new profile in x_2 will only occur if $\theta < 3\pi/2 - \alpha/2$ which is rearranged to find the inequality above. This defines the boundary between NW2 and NW3.

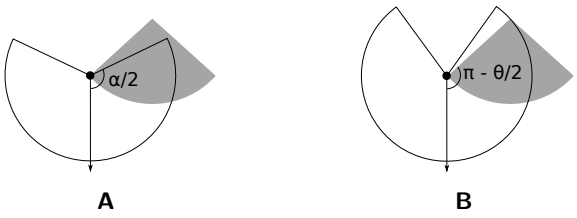


Figure A.5 Profile sizes when an animal approaches from behind in models NW2–4. If α is relatively large, animals can be detected when approaching from behind. Otherwise animals cannot be detected. The sector shaped detection region is shown in grey. Animals are filled black circles and the animal signal is an unfilled sector. The animals direction of movement is indicated with an arrow. (a) If $\alpha/2$ is less than $\pi - \theta/2$, as is the case here, then the width of the profile when an animal approaches directly from behind is zero. (b) If $\alpha/2 > \pi - \theta/2$ the profile width from behind is $2r \sin(\theta/2) \sin(x_2)$.

4. As $\alpha \leq 2\pi$ it is possible that when the angle of approach is from directly behind the sensor the animal will not be detected at all. This is the case if $\alpha/2 \leq \pi - \theta/2$ (Figure A.5). This inequality (simplified as $\alpha \leq 2\pi - \theta$) defines the boundary between NW3 and NW4.

A.2.7.1 Model NW2

NW2 is bounded by $\alpha \geq 3\pi - 2\theta$, $\alpha \leq 2\pi$ and $\theta \leq \pi$ (Figure 5.2).

NW2 has all five profiles as found in NW1. However, the change from the r profile (third integral) to the $r - r\cos(x_4)$ profile (fourth integral) occurs at $x_4 = 3\pi/2 - \alpha/2$ instead of at $x_4 = \theta$.

$$\begin{aligned}\bar{p}_{\mathrm{NW2}} = \frac{1}{\pi}\Bigg(&\int_{\frac{\theta}{2}}^{\frac{\pi}{2}} 2r \sin\left(\frac{\theta}{2}\right) \sin(x_2)\, \mathrm{d}x_2 + \int_{0}^{\theta-\frac{\pi}{2}} r - r\cos(-x_4 + \theta)\, \mathrm{d}x_4 \\ &+ \int_{\theta-\frac{\pi}{2}}^{\frac{3\pi}{2}-\frac{\alpha}{2}} r\, \mathrm{d}x_4 + \int_{\frac{3\pi}{2}-\frac{\alpha}{2}}^{\theta} r - r\cos(x_4)\, \mathrm{d}x_4 + \int_{\frac{\theta}{2}}^{\frac{\pi}{2}} 2r \sin\left(\frac{\theta}{2}\right) \sin(x_2)\, \mathrm{d}x_2 \Bigg)\end{aligned} \tag{A.17}$$

$$\bar{p}_{\mathrm{NW2}} = \frac{r}{\pi}\left(\theta - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.18}$$

A.2.7.2 Model NW3

NW3 is bounded by $\alpha \leq 3\pi - 2\theta$, $\alpha \geq 2\pi - \theta$ and $\theta \geq \pi/2$ (Figure 5.2).

NW3 does not have the fourth integral from NW2 as animals are not detectable to the right of the sensor until after the x_4 region has ended and the x_2 region has begun. Therefore the second x_4 integral has an upper limit of θ and the profile after has a width of $r\cos(x_2 - \theta/2)$ and is integrated with respect to x_2. The final integral starts at $x_4 = 3\pi/2 - \alpha/2 - \theta/2$ and has the full width of $2r\sin(x_2)\sin(\theta/2)$.

$$\begin{aligned}\bar{p}_{\mathrm{NW3}} =& \frac{1}{\pi}\left(\int_{\frac{\theta}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin\left(x_2\right)\,\mathrm{d}x_2 + \int_{0}^{\theta-\frac{\pi}{2}} r - r\cos\left(-x_4+\theta\right)\,\mathrm{d}x_4 + \int_{\theta-\frac{\pi}{2}}^{\theta} r\,\mathrm{d}x_4 \right.\\ &\left. + \int_{\frac{\theta}{2}}^{\frac{3\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}} r\cos\left(\frac{\theta}{2}-x_2\right)\mathrm{d}x_2 + \int_{\frac{3\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin\left(x_2\right)\,\mathrm{d}x_2 \right) \qquad \text{(A.19)}\\ \bar{p}_{\mathrm{NW3}} =& \frac{r}{\pi}\left(\theta - \cos\left(\frac{\alpha}{2}\right) + 1\right) \qquad \text{(A.20)}\end{aligned}$$

A.2.7.3 Model NW4

Finally, NW4 is bounded by $\alpha \geq \pi$, $\theta \geq \pi/2$ and $\alpha \leq 2\pi - \theta$ (Figure 5.2). NW4 is the same as NW3 except that the final profile width is zero and this profile is reached at $\alpha/2 + \theta/2 - \pi/2$.

$$\begin{aligned}\bar{p}_{\mathrm{NW4}} =& \frac{1}{\pi}\left(\int_{\frac{\theta}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin\left(x_2\right)\,\mathrm{d}x_2 + \int_{0}^{\theta-\frac{\pi}{2}} r - r\cos\left(-x_4+\theta\right)\,\mathrm{d}x_4 \right.\\ &\left. + \int_{\theta-\frac{\pi}{2}}^{\theta} r\,\mathrm{d}x_4 + \int_{\frac{\theta}{2}}^{\frac{\alpha}{2}+\frac{\theta}{2}-\frac{\pi}{2}} r\cos\left(\frac{\theta}{2}-x_2\right)\mathrm{d}x_2 \right) \qquad \text{(A.21)}\\ \bar{p}_{\mathrm{NW4}} =& \frac{r}{\pi}\left(\theta - \cos\left(\frac{\alpha}{2}\right) + 1\right) \qquad \text{(A.22)}\end{aligned}$$

A.2.8 Model REM

REM is the model from (Rowcliffe et al. 2008). It has $\alpha = 2\pi$ and $\theta \leq \pi/2$ (Figure 5.2). It has three profile widths, two of which are repeated, once as the animal

approaches from in front of the sensor and once as the animal approaches from behind the sensor.

1. Starting with an approach direction of directly towards the sensor, and examining focal angle x_2, the profile width is $2r\sin(x_2)\sin(\theta/2)$.
2. When the profile is perpendicular to the radius on the right hand of the sector sensor region, we instead examine x_3 where the profile width is $r\sin(x_3)$.
3. At $x_3 = \pi/2$ the profile becomes simply r and this continues for θ radians of x_4.
4. The x_3 profile is then repeated with an approach direction from behind the sensor.
5. Finally the x_2 profile is repeated, again with an approach direction from behind the sensor.

$$\bar{p}_{\text{REM}} = \frac{1}{\pi}\left(\int_{\frac{\pi}{2}-\frac{\theta}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin(x_2)\,\mathrm{d}x_2 + \int_{\theta}^{\frac{\pi}{2}} r\sin(x_3)\,\mathrm{d}x_3 + \int_{0}^{\theta} r\,\mathrm{d}x_4 + \int_{\theta}^{\frac{\pi}{2}} r\sin(x_3)\,\mathrm{d}x_3 + \int_{\frac{\pi}{2}-\frac{\theta}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin(x_2)\,\mathrm{d}x_2 \right) \tag{A.23}$$

$$\bar{p}_{\text{REM}} = \frac{r}{\pi}(\theta + 2) \tag{A.24}$$

A.2.9 Models NW5–7

In the models NW5–7, the sensor has $\theta \leq \pi/2$ as in the REM. As $\alpha \geq \pi$ a lot of the profiles are similar to the REM. Specifically, the first three profiles are always the same as the first three profiles of the REM. This is because when an animal is moving towards the sensor, the $\alpha \geq \pi$ signal is no different to a 2π signal. However, when approaching the sensor from behind, things are slightly different. The animal can only be detected by the sensor if the signal width is large enough that it can be detected once it has passed the sensor.

1. Starting with an approach direction of directly towards the sensor, and examining focal angle x_2, the profile width is $2r\sin(x_2)\sin(\theta/2)$.

2. When the profile is perpendicular to the radius edge of the sector sensor region, we instead examine x_3 where the profile width is $r\sin(x_3)$.

3. At $x_3 = \pi/2$ the profile becomes simply r and this continues for θ radians of x_4.

4. If $\alpha \leq 2\pi + 2\theta$, the animal becomes undetectable during this profile when x_3 has decreased in size to $\pi - \alpha/2$. This inequality marks the boundary between NW7 and NW6.

5. If instead $\alpha \geq 2\pi + 2\theta$ then the animal does not become undetectable during the x_3 focal angle. Instead the profile has width greater than zero for the whole of the x_3 angle. The x_2 profile starts with width $r\cos(x_2 - \theta/2)$ as only animals approaching to the left of the sensor are detectable.

6. During this second x_2 profile the signal width needed for animals to be detected to the left of the detector is increasing while the angle needed for animals to be detected to the right of the detector is decreasing. Therefore, either the left side becomes undetectable, making both sides undetectable (this occurs if $\alpha \leq 2\pi - \theta$ as in NW6)

7. or the right becomes detectable (if $\alpha \geq 2\pi - \theta$ as in NW5), making both sides detectable and giving a profile width of $2r\sin(x_2)\sin(\theta/2)$.

A.2.9.1 Model NW5

NW5 is bounded by $\alpha \geq 2\pi - \theta$, $\alpha \leq 2\pi$ and $\theta \leq \pi/2$ (Figure 5.2).

It is the same as REM except that it includes the extra profile in x_2 (the fifth integral) where only animals approaching to the left of the profile are detected.

$$\bar{p}_{\text{NW5}} = \frac{1}{\pi}\left(\int_{\frac{\pi}{2}-\frac{\theta}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin(x_2)\,\mathrm{d}x_2 + \int_{\theta}^{\frac{\pi}{2}} r\sin(x_3)\,\mathrm{d}x_3 + \int_{0}^{\theta} r\,\mathrm{d}x_4 \right.$$
$$\left. + \int_{\theta}^{\frac{\pi}{2}} r\sin(x_3)\,\mathrm{d}x_3 + \int_{\frac{\pi}{2}-\frac{\theta}{2}}^{\frac{3\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}} r\cos\left(\frac{\theta}{2} - x_2\right)\mathrm{d}x_2 \right.$$

$$+ \int_{\frac{3\pi}{2}-\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{\pi}{2}} 2r \sin\left(\frac{\theta}{2}\right) \sin(x_2)\, \mathrm{d}x_2 \Bigg) \tag{A.25}$$

$$\bar{p}_{\mathrm{NW5}} = \frac{r}{\pi}\left(\theta - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.26}$$

A.2.9.2 Model NW6

NW6 is bounded by $\alpha \leq 2\pi - \theta$, $\alpha \geq 2\pi + 2\theta$ and $\theta \leq \pi/2$ (Figure 5.2).

NW6 is the same NW5 except that as $\alpha \leq 2\pi - \theta$, animals that approach from directly behind the detector are not detected. Therefore at $x_2 = \alpha/2 + \theta/2 - \pi/2$ the profile width goes to zero and therefore the last integral in NW5 is not included.

$$\begin{aligned}\bar{p}_{\mathrm{NW6}} = \frac{1}{\pi}\Bigg(& \int_{\frac{\pi}{2}-\frac{\theta}{2}}^{\frac{\pi}{2}} 2r \sin\left(\frac{\theta}{2}\right) \sin(x_2)\, \mathrm{d}x_2 + \int_{\theta}^{\frac{\pi}{2}} r \sin(x_3)\, \mathrm{d}x_3 \\ & + \int_{0}^{\theta} r\, \mathrm{d}x_4 + \int_{\theta}^{\frac{\pi}{2}} r \sin(x_3)\, \mathrm{d}x_3 + \int_{\frac{\pi}{2}-\frac{\theta}{2}}^{\frac{\alpha}{2}+\frac{\theta}{2}-\frac{\pi}{2}} r \cos\left(\frac{\theta}{2} - x_2\right) \mathrm{d}x_2 \Bigg)\end{aligned} \tag{A.27}$$

$$\bar{p}_{\mathrm{NW6}} = \frac{r}{\pi}\left(\theta - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.28}$$

A.2.9.3 Model NW7

NW7 is bounded by $\alpha \geq 2\pi + 2\theta$, $\alpha \geq \pi$ and $\theta \geq 0$ (Figure 5.2).

It is similar to NW6 but does not include the last integral as during the x_3 profile, at $x_3 = \pi - \alpha/2$ the signal width is too small for any animals to be detected, so the profile width goes to zero.

$$\begin{aligned}\bar{p}_{\mathrm{NW7}} = \frac{1}{\pi}\Bigg(& \int_{\frac{\pi}{2}-\frac{\theta}{2}}^{\frac{\pi}{2}} 2r \sin\left(\frac{\theta}{2}\right) \sin(x_2)\, \mathrm{d}x_2 + \int_{\theta}^{\frac{\pi}{2}} r \sin(x_3)\, \mathrm{d}x_3 \\ & + \int_{0}^{\theta} r\, \mathrm{d}x_4 + \int_{\pi-\frac{\alpha}{2}}^{\frac{\pi}{2}} r \sin(x_3)\, \mathrm{d}x_3 \Bigg)\end{aligned} \tag{A.29}$$

$$\bar{p}_{\mathrm{NW7}} = \frac{r}{\pi}\left(\theta - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.30}$$

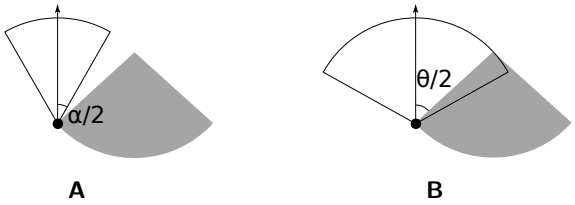


Figure A.6 The first profile in SW models is limited by either α or β depending on whether $\alpha < \beta$. The sector shaped detection region is shown in grey. Animals are filled black circles and the animal signal is an unfilled sector. The animals direction of movement is indicated with an arrow. (a) As $\alpha/2 < \theta/2$ the profile width is limited by the signal width rather than the sensor region. The profile width is $2r \sin(\alpha/2)$ (b) As $\alpha/2 > \theta/2$ the profile width is limited by the sensor region, not the signal width. The profile width is $2r \sin(\theta/2) \sin(x_2)$.

A.2.10 Model SW1–3

The models in SW1–3 are described with the two focal angles used in models NW2–4, x_2 and x_4. As $\alpha \leq \pi$ an animal can never be detected if it is approaching the detector from behind. This makes these models simpler in that they go through the x_2 and x_4 profiles only once each.

There are five potential profile sizes.

1. At the beginning of x_2, with an approach direction directly towards the sensor, the parameter that limits the width of the profile can either be the sensor width, in which case the profile width is $2r \sin(\theta/2) \sin(x_2)$.

2. Or the signal width can be the limiting parameter, in which case the profile width is instead $2r \sin(\alpha/2)$ (Figure A.6)

3. The next potential profile in x_2 has a width of $r \sin(\alpha/2) - r \cos(x_2 + \theta/2)$ as the right side of the profile is limited by the width of the sensor region while the left side is limited by the signal width. However, the angle at which the profile starts depends on whether the first profile was 1) or 2) above. If the first profile is profile 1) then the profile is limited on both sides by the sensor region and then the left side of the profile becomes limited by the signal width. This happens at $x_2 = \pi/2 - \alpha/2 + \theta/2$. If however the first profile was 2) then the first profile is limited by the signal width. We move into the new profile when the right side of the profile becomes limited by the sensor region. This occurs at $x_2 = \pi/2 + \alpha/2 - \theta/2$.

4. In the x_4 region the left side of the profile is always $r\sin(\alpha/2)$ while the right side is either 0, giving a profile of $r\sin(\alpha/2)$.

5. Or limited by the sensor giving a profile of size $r\sin(\alpha/2) - r\cos(x_4 - \theta)$.

A.2.10.1 Model SW1

SW1 is bounded by $\alpha \geq \theta$, $\alpha \leq \pi$ and $\theta \leq \pi$ (Figure 5.2).

As α is large the first profile is limited by the size of the sensor region giving it a width of $2r\sin(\theta/2)\sin(x_2)$. It is the only one of the three SW models to start in this way. Later on, still with x_2 as the focal angle the left side of the profile does become limited by the signal width. So at $x_2 = \pi/2 - \alpha/2 + \theta/2$ the profile width becomes $r\sin(\alpha/2) - r\cos(x_2 + \theta/2)$.

As we enter the x_4 region, the profile remains limited by the signal on the left and by the sensor on the right, giving a profile width of $r\sin(\alpha/2) - r\cos(x_4 - \theta)$. Finally, at $x_4 = \theta - \pi/2$ the right side of the profile becomes zero and the profile is width is $r\sin(\alpha/2)$.

$$\begin{aligned}\bar{p}_{\text{SW1}} =& \frac{1}{\pi}\left(\int_{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin(x_2)\,\mathrm{d}x_2 + \int_{\frac{\theta}{2}}^{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}} r\sin\left(\frac{\alpha}{2}\right) - r\cos\left(\frac{\theta}{2}+x_2\right)\mathrm{d}x_2\right.\\ &\left. + \int_{0}^{\theta-\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right) - r\cos(\theta - x_4)\,\mathrm{d}x_4 + \int_{\theta-\frac{\pi}{2}}^{\frac{\alpha}{2}+\theta-\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right)\mathrm{d}x_4\right) \qquad \text{(A.31)}\\ \bar{p}_{\text{SW1}} =& \frac{r}{\pi}\left(\theta\sin\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\alpha}{2}\right) + 1\right) \qquad \text{(A.32)}\end{aligned}$$

A.2.10.2 Model SW2

SW2 is bounded by $\theta \geq \pi/2$, $\alpha \leq \theta$ and $\alpha \geq 2\theta - \pi$ (Figure 5.2).

SW2 is largely similar to SW1. However, as $\alpha \leq \theta$ the first profile is limited by α and not by the detection region. Therefore the first profile has width $2r\sin(\alpha/2)$. This also means the transition to the second profile occurs at $x_2 = \pi/2 + \alpha/2 - \theta/2$ instead of $x_2 = \pi/2 - \alpha/2 + \theta/2$.

$$\bar{p}_{\text{SW2}} = \frac{1}{\pi}\left(\int_{\frac{\alpha}{2}-\frac{\theta}{2}+\frac{\pi}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\alpha}{2}\right)\mathrm{d}x_2 + \int_{\frac{\theta}{2}}^{\frac{\alpha}{2}-\frac{\theta}{2}+\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right) - r\cos\left(\frac{\theta}{2}+x_2\right)\mathrm{d}x_2\right.$$

$$+ \int_{0}^{\theta-\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right) - r\cos\left(\theta - x_4\right)\,\mathrm{d}x_4 + \int_{\theta-\frac{\pi}{2}}^{\frac{\alpha}{2}+\theta-\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right)\mathrm{d}x_4\Bigg) \tag{A.33}$$

$$\bar{p}_{\mathrm{SW2}} = \frac{r}{\pi}\left(\theta\sin\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.34}$$

A.2.10.3 Model SW3

SW3 is bounded by $\alpha \leq 2\theta - \pi$ and $\theta \leq \pi$ (Figure 5.2).

SW3 is similar to SW2 except that the profile does not become limited by sensor at all during the the x_4 regions. Therefore, at $x_4 = 0$ the profile is still of width $2r\sin(\alpha/2)$. Only at $x_4 = \theta - \pi/2 - \alpha/2$ does the profile become limited on the right by the sensor region.

$$\bar{p}_{\mathrm{SW3}} = \frac{1}{\pi}\Bigg(\int_{\frac{\theta}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\alpha}{2}\right)\mathrm{d}x_2 + \int_{0}^{-\frac{\pi}{2}+\theta-\frac{\alpha}{2}} 2r\sin\left(\frac{\alpha}{2}\right)\mathrm{d}x_4$$
$$+ \int_{-\frac{\pi}{2}+\theta-\frac{\alpha}{2}}^{\theta-\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right) - r\cos\left(\theta - x_4\right)\,\mathrm{d}x_4 + \int_{\theta-\frac{\pi}{2}}^{\frac{\alpha}{2}+\theta-\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right)\mathrm{d}x_4\Bigg) \tag{A.35}$$

$$\bar{p}_{\mathrm{SW3}} = \frac{r}{\pi}\left(\theta\sin\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.36}$$

A.2.11 Model SW4–9

As $\alpha < \pi$, animals approaching the sensor from behind can never be detected, so unlike REM, the second x_2 and x_3 profiles are always zero. The six models are split by three inequalities that relate to the models as follows.

1. Models with $\alpha \leq \pi - 2\theta$ have no x_4 profile. This is because at $x_4 = 0$, the signal width is already too small to be detected as can be seen in Figure A.7A where $\alpha/2 < \pi/2 - \theta$ which simplifies to give the previous inequality.

2. Models with $\alpha \leq \theta$ are limited by α in the first, x_2 region (Figure A.6), rather than being limited by θ. Therefore this first profile is of width $2r\sin(\alpha/2)$ rather than $2r\sin(\theta/2)\sin(x_2)$.

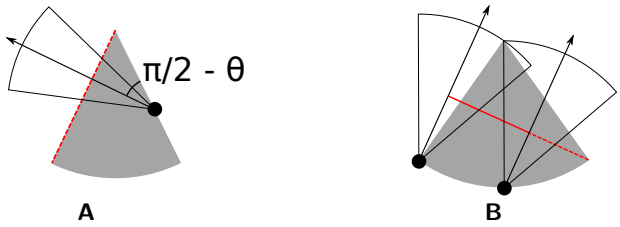


Figure A.7 Description of two profiles in SW models. The sector shaped detection region is shown in grey. Animals are filled black circles and the animal signal is an unfilled sector. The animals direction of movement is indicated with an arrow. The profile p is shown with a red line. Dashed red lines indicate areas where animals cannot be detected. (a) At $x_4 = 0$, if $\alpha/2 < \pi/2 - \theta$ then $\alpha/2$ is too small for an animal to be detected at all during the x_4 profile (shown with dashed red.) This inequality simplifies to $\alpha < \pi - 2\theta$. (b) The right of the profile is limited by the signal width, not the sensor. On the left, the profile is limited by the sensor and not the signal. Overall the profile width is $r\sin(\alpha/2) - r\cos(x_2 + \theta/2)$.

3. Finally, models with $\alpha \leq 2\theta$ have a second profile in x_2 where to one side of the sensor α is the limiting factor of profile width, while on the other side θ is (Figure A.7B). This gives a width of $r\sin(\alpha/2) - r\cos(x_2 + \theta/2)$. This profile does not occur in models with $\alpha \geq 2\theta$.

A.2.11.1 Model SW4

SW4 is bounded by $\alpha \leq \theta$, $\alpha \geq \pi - 2\theta$ and $\theta \leq \pi/2$ (Figure 5.2). Therefore it does contain a x_4 profile, starts with an α limited profile and does contain the $r\sin(\alpha/2) - r\cos(x_2 + \theta/2)$ profile in x_2.

$$\bar{p}_{\text{SE4}} = \frac{1}{\pi}\left(\int_{\frac{\pi}{2}}^{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}} 2r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_1 + \int_{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{\theta}{2}+\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right) + r\cos\left(\frac{\theta}{2} - x_1\right) \mathrm{d}x_1 + \int_{\frac{\theta}{2}+\frac{\pi}{2}}^{\frac{\alpha}{2}+\frac{\theta}{2}+\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_1 \right) \tag{A.37}$$

$$\bar{p}_{\text{SE4}} = \frac{r}{\pi}\left(\theta\sin\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.38}$$

A.2.11.2 Model SW5

SW5 is the only model with a tetrahedral bounding region. It is bounded by $\alpha \geq \theta$, $\alpha \geq \pi - 2\theta$, $\alpha \leq 2\theta$ and $\theta \leq \pi/2$ (Figure 5.2). Therefore it does contain a x_4 profile, but starts with a θ limited profile. It does contain the $r\sin(\alpha/2) - r\cos(x_2 + \theta/2)$ profile in x_2.

$$\begin{aligned}\bar{p}_{\text{SW5}} =& \frac{1}{\pi}\left(\int_{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin\left(x_2\right)\,\mathrm{d}x_2 \right. \\ &+ \int_{\frac{\pi}{2}-\frac{\theta}{2}}^{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}} r\sin\left(\frac{\alpha}{2}\right) - r\cos\left(\frac{\theta}{2}+x_2\right)\mathrm{d}x_2 \\ &\left. + \int_{\theta}^{\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right)\mathrm{d}x_3 + \int_{0}^{\frac{\alpha}{2}+\theta-\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right)\mathrm{d}x_4 \right) \end{aligned} \tag{A.39}$$

$$\bar{p}_{\text{SW5}} = \frac{r}{\pi}\left(\theta\sin\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.40}$$

A.2.11.3 Model SW6

SW6 is bounded by $\alpha \geq \pi - 2\theta$, $\alpha \geq 2\theta$ and $\alpha \leq \pi$ (Figure 5.2). It starts with a θ limited profile and has a x_4 profile. However, it does not contain the $r\sin(\alpha/2) - r\cos(x_2 + \theta/2)$ profile.

$$\begin{aligned}\bar{p}_{\text{SW6}} =& \frac{1}{\pi}\left(\int_{\frac{\pi}{2}-\frac{\theta}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin\left(x_2\right)\,\mathrm{d}x_2 + \int_{\theta}^{\frac{\alpha}{2}} r\sin\left(x_3\right)\,\mathrm{d}x_3 \right. \\ &\left. + \int_{\frac{\alpha}{2}}^{\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right)\mathrm{d}x_3 + \int_{0}^{\frac{\alpha}{2}+\theta-\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right)\mathrm{d}x_4 \right) \end{aligned} \tag{A.41}$$

$$\bar{p}_{\text{SW6}} = \frac{r}{\pi}\left(\theta\sin\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.42}$$

A.2.11.4 Model SW7

SW7 is bounded by $\alpha \leq \pi - 2\theta$, $\alpha \leq \theta$ and $\alpha < 0$ (Figure 5.2). Therefore it does not contain a x_4 profile. It starts with an α limited profile and contains the $r\sin(\alpha/2) - r\cos(x_2 + \theta/2)$ profile in x_2.

$$\bar{p}_{\text{SW7}} = \frac{1}{\pi}\left(\int_{\frac{\alpha}{2}-\frac{\theta}{2}+\frac{\pi}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_2 + \int_{\frac{\pi}{2}-\frac{\theta}{2}}^{\frac{\alpha}{2}-\frac{\theta}{2}+\frac{\pi}{2}} r\sin\left(\frac{\alpha}{2}\right) - r\cos\left(\frac{\theta}{2}+x_2\right) \mathrm{d}x_2 + \int_{\theta}^{\frac{\alpha}{2}+\theta} r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_3 \right) \tag{A.43}$$

$$\bar{p}_{\text{SW7}} = \frac{r}{\pi}\left(\theta\sin\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.44}$$

A.2.11.5 Model SW8

SW8 is bounded by $\alpha \leq \pi - 2\theta$, $\alpha \geq \theta$ and $\alpha \leq 2\theta$ (Figure 5.2). It starts with a θ limited profile. It does contain the $r\sin(\alpha/2) - r\cos(x_2 + \theta/2)$ profile in x_2 but does not have a x_4 profile.

$$\bar{p}_{\text{SW8}} = \frac{1}{\pi}\left(\int_{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin\left(x_2\right) \mathrm{d}x_2 + \int_{\frac{\pi}{2}-\frac{\theta}{2}}^{\frac{\pi}{2}+\frac{\theta}{2}-\frac{\alpha}{2}} r\sin\left(\frac{\alpha}{2}\right) - r\cos\left(\frac{\theta}{2}+x_2\right) \mathrm{d}x_2 + \int_{\theta}^{\frac{\alpha}{2}+\theta} r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_3 \right) \tag{A.45}$$

$$\bar{p}_{\text{SW8}} = \frac{r}{\pi}\left(\theta\sin\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.46}$$

A.2.11.6 Model SW9

Finally, SW9, the last model, is bounded by y $\alpha \leq \pi - 2\theta$, $\alpha \geq 2\theta$ and $\theta \geq 0$ (Figure 5.2). Therefore it starts with a θ limited profile. However it does not contain the extra x_2 profile nor a x_4 profile.

$$\bar{p}_{\text{SW9}} = \frac{1}{\pi}\left(\int_{\frac{\pi}{2}-\frac{\theta}{2}}^{\frac{\pi}{2}} 2r\sin\left(\frac{\theta}{2}\right)\sin\left(x_2\right) \mathrm{d}x_2 + \int_{\theta}^{\frac{\alpha}{2}} r\sin\left(x_3\right) \mathrm{d}x_3 + \int_{\frac{\alpha}{2}}^{\frac{\alpha}{2}+\theta} r\sin\left(\frac{\alpha}{2}\right) \mathrm{d}x_3 \right) \tag{A.47}$$

$$\bar{p}_{\text{SW9}} = \frac{r}{\pi}\left(\theta\sin\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\alpha}{2}\right) + 1\right) \tag{A.48}$$

A.3 Supplementary Information: Simulation model results of the gREM precision

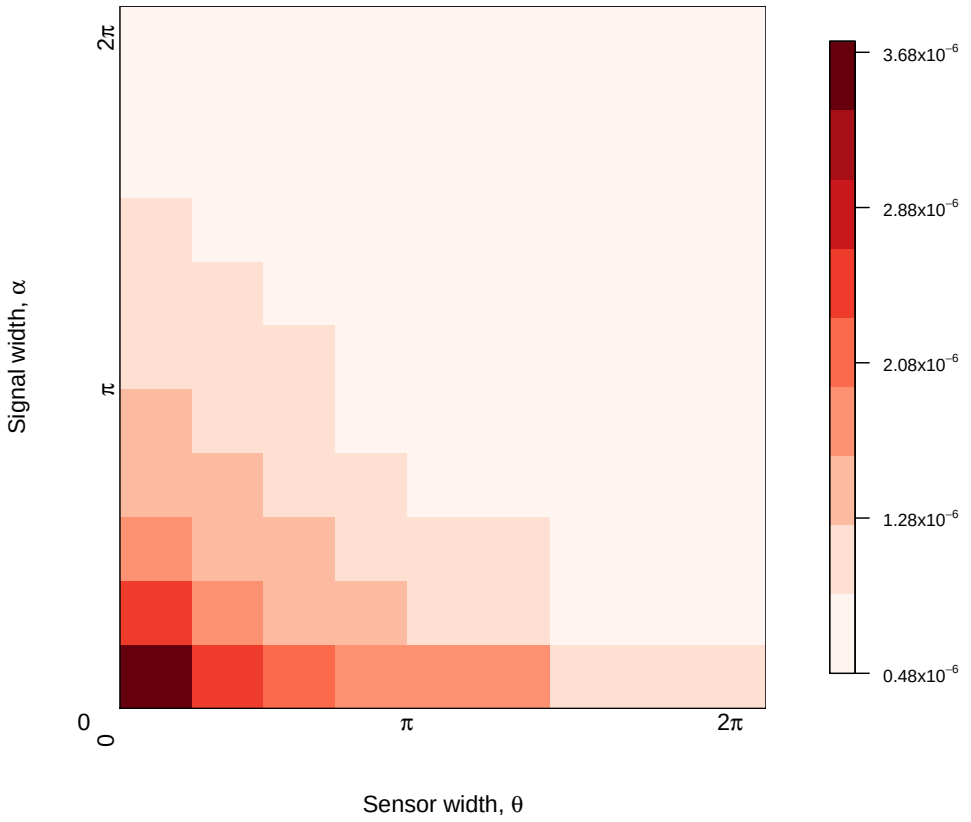
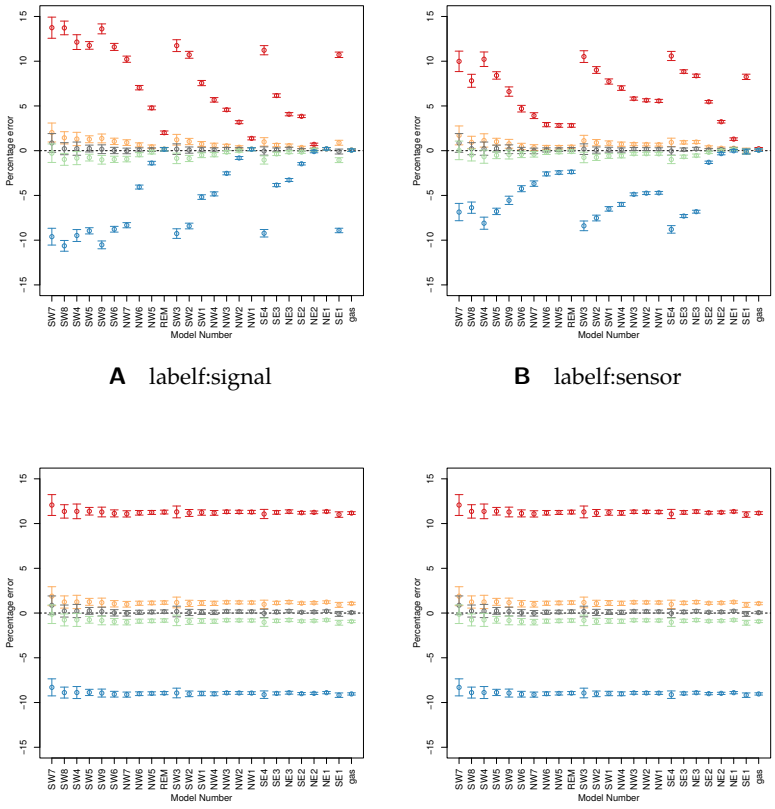


Figure A.1 Simulation model results of the gREM precision given a range of sensor and signal widths, shown by the standard deviation of the error between the estimated and true densities. Standard deviations are shown from deep red to pink, representing high to low values between 0.483×10^{-6} to 3.74×10^{-6}.

A.4 Supplementary Information: Impact of parameter error



A labelf:signal **B** labelf:sensor

C labelf:radius **D** labelf:speed

Figure A.2 Model sensitivity (for all gREM submodels) to error in estimates of a) signal width α, b) sensor width θ, c) detection distance r and d) animal movement speed v. Estimates are -10% (red), -1% (orange), 0% (grey), +1% (green) and +10% (blue) of the true parameter value. The black dashed line indicates zero error in density estimates. The error bars 95% confidence intervals across all simulations.

B Colophon

This document was set using LaTeX, XeLaTeX and BibLaTeX. The formatting is defined by the phdthesis class by Robert Stanley. The TeX Gyre Pagella typeface is used throughout. Chapters 2, 3 and 4 are entirely reproducible knitr documents (Xie 2015). Code for the simulations in Chapter 5 is not combined into a knitr document but code for creating figures is. Plots were created with a combination of Inkscape, Lucas (2015), Wickham (2009) and Yu (2015). References were handled with JabRef (JabRef Development Team 2015).

Bibliography

Acevedo MA & Villanueva-Rivera LJ (2006). 'Using automated digital recording systems as effective tools for the monitoring of birds and amphibians'. *Wildlife Society Bulletin* **34**(1), 211–214. DOI: 10.1016/s1769-4493(08)70101-8

Ackleh AS & Allen LJ (2003). 'Competitive exclusion and coexistence for pathogens in an epidemic model with variable population size'. *Journal of mathematical biology* **47**(2), 153–168. DOI: 10.1007/s00285-003-0207-9

Ackleh AS & Salceanu PL (2014). 'Robust uniform persistence and competitive exclusion in a nonautonomous multi-strain SIR epidemic model with disease-induced mortality'. *Journal of mathematical biology* (1-2), 453–475. DOI: 10.1007/s00285-012-0636-4

Adams A, Jantzen M, Hamilton R & Fenton M (2012). 'Do you hear what I hear? Implications of detector selection for acoustic monitoring of bats'. *Methods in Ecology and Evolution*. DOI: 10.1126/science.291.5510.1855k

Aldridge H (1987). 'Turning flight of bats'. *Journal of Experimental Biology* **128**(1), 419–425

Allen LJ, Kirupaharan N & Wilson SM (2004). 'SIS epidemic models with multiple pathogen strains'. *Journal of Difference Equations and Applications* **10**(1), 53–75. DOI: 10.1080/10236190310001603680

Allen LJ, Langlais M & Phillips CJ (2003). 'The dynamics of two viral infections in a single host population with applications to hantavirus'. *Mathematical biosciences* **186**(2), 191–217. DOI: 10.1016/j.mbs.2003.08.002

Altizer S, Nunn CL, Thrall PH, Gittleman JL, Antonovics J, Cunnningham AA, Dobson AP, Ezenwa V, Jones KE et al. (2003). 'Social organization and parasite risk in mammals: integrating theory and empirical studies'. *Annual Review of Ecology, Evolution, and Systematics*, 517–547. DOI: 10.1146/annurev.ecolsys.34.030102.151725

Amengual B, Bourhy H, López-Roig M & Serra-Cobo J (2007). 'Temporal dynamics of European bat Lyssavirus type 1 and survival of Myotis myotis bats in natural colonies'. *PLoS One* **2**(6), e566. DOI: 10.1371/journal.pone.0000566

Amman B, Carroll S, Reed Z, Sealy T, Balinandi S, Swanepoel R, Kemp A, Erickson B, Comer J, Campbell S et al. (2012). 'Seasonal pulses of Marburg virus circulation in juvenile *Rousettus aegyptiacus* bats coincide with periods of increased risk of human infection'. *PLoS Pathogens* **8**(10), e1002877

Anderson DR (2001). 'The need to get the basics right in wildlife field studies'. *Wildlife Society Bulletin* **29**(4), 1294–1297. DOI: 10.1211/pj.2013.11124181

Anderson RM & May RM (1979). 'Population biology of infectious diseases: Part I.' *Nature* (280), 361–7. DOI: 10.1038/280361a0

Anthony SJ, Epstein JH, Murray KA, Navarrete-Macias I, Zambrana-Torrelio CM, Solovyov A, Ojeda-Flores R, Arrigo NC, Islam A, Khan SA, Hosseini P, Bogich TL, Olival KJ, Sanchez-Leon MD, Karesh WB, Goldstein T, Luby SP, Morse SS, Mazet JAK, Daszak P & Lipkin WI (2013). 'A strategy to estimate unknown viral diversity in mammals'. *MBio* **4**(5), e00598–13. DOI: 10.1128/mBio.00598-13

Anthony SJ, Islam A, Johnson C, Navarrete-Macias I, Liang E, Jain K, Hitchens PL, Che X, Soloyvov A, Hicks AL, Ojeda-Flores R, Zambrana-Torrelio C, Ulrich W, Rostal MK, Petrosov A, Garcia J, Haider N, Wolfe N, Goldstein T, Morse SS, Rahman M, Epstein JH, Mazet JK, Daszak P & Lipkin WI (2015). 'Non-random patterns in viral diversity'. *Nature communications* **6**. DOI: 10.1038/ncomms9147

Arita HT (1993). 'Rarity in Neotropical bats: correlations with phylogeny, diet, and body mass'. *Ecological Applications*, 506–517. DOI: 10.2307/1941919

Arneberg P (2002). 'Host population density and body mass as determinants of species richness in parasite communities: comparative analyses of directly transmitted nematodes of mammals'. *Ecography* (1), 88–94. DOI: 10.1034/j.1600-0587.2002.250110.x

Barlow J & Taylor B (2005). 'Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey'. *Marine Mammal Science* **21**(3), 429–445. DOI: 10.1111/j.1748-7692.2005.tb01242.x

Barthélemy M, Godreche C & Luck JM (2010). 'Fluctuation effects in metapopulation models: percolation and pandemic threshold'. *Journal of theoretical biology* **267**(4), 554–564

Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL & Purvis A (2007). 'The delayed rise of present-day mammals'. *Nature* **446**(7135), 507–512. DOI: 10.3410/f.1076757.529672

Birt P, McCoy M & Palmer C (2008). 'Little Red Flying-fox, *Pteropus scapulatus*'. In: *The mammals of Australia*. Ed. by S Van Dyck & R Strahan. 3rd ed. Sydney, Australia: Reed New Holland, 446–447

Blackwood JC, Streicker DG, Altizer S & Rohani P (2013). 'Resolving the roles of immunity, pathogenesis, and immigration for rabies persistence in vampire bats'. *Proceedings of the National Academy of Sciences* **110**(51), 20837–20842. DOI: 10.1073/pnas.1308817110

Blumstein DT, Mennill DJ, Clemins P, Girod L, Yao K, Patricelli G, Deppe JL, Krakauer AH, Clark C, Cortopassi KA, Hanser SF, McCowan B, Ali AM & Kirschel ANG (2011). 'Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus'. *Journal of Applied Ecology* **48**(3), 758–767. DOI: 10.1111/j.1365-2664.2011.01993.x

Borchers D, Distiller G, Foster R, Harmsen B & Milazzo L (2014). 'Continuous-time spatially explicit capture–recapture models, with an application to a jaguar camera-trap survey'. *Methods in Ecology and Evolution* **5**(7), 656–665

Bordes F, Blumstein DT & Morand S (2007). 'Rodent sociality and parasite diversity'. *Biology letters* **3**(6), 692–694. DOI: 10.1098/rsbl.2007.0393

Bordes F & Morand S (2011). 'The impact of multiple infections on wild animal hosts: a review'. *Infection ecology & epidemiology* **1**. DOI: 10.3402/iee.v1i0.7346

Bremermann HJ & Thieme H (1989). 'A competitive exclusion principle for pathogen virulence'. *Journal of mathematical biology* **27**, 179–190. DOI: 10.1007/BF00276102

Brierley L, Vonhof M, Olival K, Daszak P & Jones K (2016). 'Quantifying Global Drivers of Zoonotic Bat Viruses: A Process-based Perspective'. *The American Naturalist* **187**

Brinkløv S, Jakobsen L, Ratcliffe J, Kalko E & Surlykke A (2011). 'Echolocation call intensity and directionality in flying short-tailed fruit bats, *Carollia perspicillata* (Phyllostomidae)'. *The Journal of the Acoustical Society of America* **129**(1), 427–435. DOI: 10.3389/fphys.2013.00089

Brusa A & Bunker DE (2014). 'Increasing the precision of canopy closure estimates from hemispherical photography: Blue channel analysis and under-exposure'. *Agricultural and Forest Meteorology* **195**, 102–107. DOI: 10.1016/j.agrformet.2014.05.001

Buckee CO, Koelle K, Mustard MJ & Gupta S (2004). 'The effects of host contact network structure on pathogen diversity and strain structure'. *Proceedings of the National Academy of Sciences of the United States of America* **101**(29), 10839–10844. DOI: 10.1073/pnas.0402000101

Buckland ST & Handel C (2006). 'Point-transect surveys for songbirds: robust methodologies'. *The Auk* **123**(2), 345–357. DOI: 10.1642/0004-8038(2006)123[345:psfsrm]2.0.co;2

Buckland ST, Marsden SJ & Green RE (2008). 'Estimating bird abundance: making methods work'. *Bird Conservation International* **18**(S1), S91–S108. DOI: 10.1017/s0959270908000294

Burland T, Barratt E, Beaumont M & Racey P (1999). 'Population genetic structure and gene flow in a gleaning bat, *Plecotus auritus*'. *Proceedings of the Royal Society of London B: Biological Sciences* **266**(1422), 975–980. DOI: 10.1007/978-3-642-70837-4_20

Burnham KP & Anderson DR (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media. DOI: 10.1007/978-0-387-22456-5_1

Calisher C, Childs J, Field H, Holmes K & Schountz T (2006). 'Bats: important reservoir hosts of emerging viruses'. *Clinical Microbiology Reviews* **19**(3), 531–545. DOI: 10.1128/CMR.00017-06

Campos PR & Gordo I (2006). 'Pathogen genetic variation in small-world host contact structures'. *Journal of Statistical Mechanics: Theory and Experiment* **2006**(12), L12003. DOI: 10.1088/1742-5468/2006/12/L12003

Canals M, Atala C, Grossi B & Iriarte-Díaz J (2005). 'Relative size of hearts and lungs of small bats'. *Acta Chiropterologica* **7**(1), 65–72. DOI: 10.3161/1733-5329(2005)7[65:rsohal]2.0.co;2

Carbone C, Cowlishaw G, Isaac NJ & Rowcliffe JM (2005). 'How far do animals go? Determinants of day range in mammals'. *The American Naturalist* **165**(2), 290–297. DOI: 10.1086/426790

Castillo-Chavez C, Huang W & Li J (1995). 'Dynamics of multiple pathogen strains in heterosexual epidemiological models'. *Differential Equations and*

Applications to Biology and Industry. World Scientific Publishing Co., Singapore, 289–298

Chamberlain SA & Szöcs E (2013). 'taxize: taxonomic search and retrieval in R'. *F1000Research* **2**. DOI: 10.12688/f1000research.2-191.v1

Colizza V & Vespignani A (2007). 'Invasion threshold in heterogeneous metapopulation networks'. *Physical review letters* **99**(14), 148701. DOI: 10.1103/physrevlett.99.148701

Crespo EA, Pedraza SN, Grandi MF, Dans SL & Garaffo GV (2010). 'Abundance and distribution of endangered Franciscana dolphins in Argentine waters and conservation implications'. *Marine Mammal Science* **26**(1), 17–35. DOI: 10.3354/esr00391

Cryan PM, Stricker CA & Wunder MB (2014). 'Continental-scale, seasonal movements of a heterothermic migratory tree bat'. *Ecological Applications* **24**(4), 602–616. DOI: 10.1890/13-0752.1

Damuth J (1981). 'Population density and body size in mammals'. *Nature* **290**(5808), 699–700. DOI: 10.1038/290699a0

Depraetere M, Pavoine S, Jiguet F, Gasc A, Duvail S & Sueur J (2012). 'Monitoring animal diversity using acoustic indices: implementation in a temperate woodland'. *Ecological Indicators* **13**(1), 46–54. DOI: 10.1016/j.ecolind.2011.05.006

Dobson AP (2005). 'What links bats to emerging infectious diseases?' *Science* **310**(5748), 628–629. DOI: 10.1126/science.1120872

Drosten C, Günther S, Preiser W, Van Der Werf S, Brodt HR, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RA et al. (2003). 'Identification of a novel coronavirus in patients with severe acute respiratory syndrome'. *New England Journal of Medicine* **348**(20), 1967–1976. DOI: 10.1056/NEJMoa030747

Everatt KT, Andresen L & Somers MJ (2014). 'Trophic scaling and occupancy analysis reveals a lion population limited by top-down anthropogenic pressure in the Limpopo National Park, Mozambique'. *PloS one* **9**(6), e99389. DOI: 0120099389

Ezenwa VO, Price SA, Altizer S, Vitone ND & Cook KC (2006). 'Host traits and parasite species richness in even and odd-toed hoofed mammals, Artiodactyla and Perissodactyla'. *Oikos* **115**(3), 526–536. DOI: 10.1111/j.2006.0030-1299.15186.x

Fenton A & Perkins SE (2010). 'Applying predator-prey theory to modelling immune-mediated, within-host interspecific parasite interactions'. *Parasitology* **137**(06), 1027–1038. DOI: 10.1017/s0031182009991788

Field H, Young P, Yob J, Mills J, Hall L & Mackenzie J (2001). 'The natural history of Hendra and Nipah viruses'. *Microbes and infection* **3**(4), 307–314. DOI: 10.1016/S1286-4579(01)01384-3

Fleming TH, Eby P, Kunz T & Fenton M (2003). 'Ecology of bat migration'. In: *Bat ecology*. Ed. by T Kunz & M Brock Fenton. Chicago: The University of Chicago Press, 156–208

Foster RJ & Harmsen BJ (2012). 'A critique of density estimation from camera-trap data'. *The Journal of Wildlife Management* **76**(2), 224–236. DOI: 10.1002/jwmg.275

Fritz SA, Bininda-Emonds OR & Purvis A (2009). 'Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics'. *Ecology letters* **12**(6), 538–549. DOI: 10.3410/f.1166120.627034

Funk S & Jansen VA (2010). 'Interacting epidemics on overlay networks'. *Physical Review E* **81**(3), 036118

Gay N, Olival KJ, Bumrungsri S, Siriaroonrat B, Bourgarel M & Morand S (2014). 'Parasite and viral species richness of Southeast Asian bats: Fragmentation of area distribution matters'. *International Journal for Parasitology: Parasites and Wildlife* **3**(2), 161–170. DOI: 10.1016/j.ijppaw.2014.06.003

Goffman W & Newill V (1964). 'Generalization of epidemic theory'. *Nature* **204**(4955), 225–228

Granich R, Gupta S, Hersh B, Williams B, Montaner J, Young B & Zuniga JM (2015). 'Trends in AIDS deaths, new infections and ART coverage in the top 30 countries with the highest AIDS mortality burden; 1990–2013'. *PloS one* **10**(7), e0131353. DOI: 10.1371/journal.pone.0131353

Griffiths EC, Pedersen AB, Fenton A & Petchey OL (2014). 'Analysis of a summary network of co-infection in humans reveals that parasites interact most via shared resources'. *Proceedings of the Royal Society of London B: Biological Sciences* **281**(1782), 20132286. DOI: 10.1098/rspb.2013.2286

Halpin K, Hyatt AD, Fogarty R, Middleton D, Bingham J, Epstein JH, Rahman SA, Hughes T, Smith C, Field HE et al. (2011). 'Pteropid bats are confirmed as the reservoir hosts of henipaviruses: a comprehensive experimental study of virus transmission'. *The American journal of tropical medicine and hygiene* **85**(5), 946–951. DOI: 10.4269/ajtmh.2011.10-0567

Harris D, Matias L, Thomas L, Harwood J & Geissler WH (2013). 'Applying distance sampling to fin whale calls recorded by single seismic instruments in the northeast Atlantic'. *The Journal of the Acoustical Society of America* **134**(5), 3522–3535. DOI: 10.1121/1.4821207

Hassel-Finnegan HM, Borries C, Larney E, Umponjan M & Koenig A (2008). 'How reliable are density estimates for diurnal primates?' *International Journal of Primatology* **29**(5), 1175–1187. DOI: 10.1007/s10764-008-9301-6

Hayman DT (2015). 'Biannual birth pulses allow filoviruses to persist in bat populations'. *Proceedings of the Royal Society of London B: Biological Sciences* **282**(1803), 20142591

Heaney LR, Balete DS, Alviola P, Rickart EA & Ruedi M (2012). '*Nyctalus plancyi* and *Falsistrellus petersi* (Chiroptera: Vespertilionidae) from northern Luzon, Philippines: ecology, phylogeny, and biogeographic implications'. *Acta Chiropterologica* **14**(2), 265–278. DOI: 10.3161/150811012x661602

Heesterbeek J (2002). 'A brief history of R 0 and a recipe for its calculation'. *Acta biotheoretica* **50**(3), 189–204. DOI: 10.1023/A:1016599411804

Henderson LE & Broders HG (2008). 'Movements and resource selection of the Northern Long-Eared Myotis (*Myotis septentrionalis*) in a forest–agriculture landscape'. *Journal of Mammalogy* **89**(4), 952–963. DOI: 10.1644/07-mamm-a-214.1

Holderied M & Von Helversen O (2003). 'Echolocation range and wingbeat period match in aerial-hawking bats'. *Proc. R. Soc. B* **270**(1530), 2293–2299. DOI: 10.1098/rspb.2003.2487

Hughes JM, Wilson ME, Halpin K, Hyatt AD, Plowright RK, Epstein JH, Daszak P, Field HE, Wang L & Daniels PW (2007). 'Emerging viruses: coming in on a wrinkled wing and a prayer'. *Clinical Infectious Diseases* **44**(5), 711–717. DOI: 10.1086/511078

Hulva P, Fornůsková A, Chudárková A, Evin A, Allegrini B, Benda P & Bryja J (2010). 'Mechanisms of radiation in a bat group from the genus Pipistrellus inferred by phylogeography, demography and population genetics'. *Molecular ecology* **19**(24), 5417–5431

Hutchinson JMC & Waser PM (2007). 'Use, misuse and extensions of "ideal gas" models of animal encounter.' *Biological Reviews of the Cambridge Philosophical Society* **82**(3), 335–359. DOI: 10.1111/j.1469-185x.2007.00014.x. URL: http://www.ncbi.nlm.nih.gov/pubmed/17624958

ITIS (2015). *Integrated Taxonomic Information System*. http://www.itis.gov

IUCN (2010). *Red List Of Threatened Species. Version 2010.1*. www.iucnredlist.org

JabRef Development Team (2015). *JabRef*. URL: http://jabref.sf.net

Jirinec V, Campos BR & Johnson MD (2011). 'Roosting behaviour of a migratory songbird on Jamaican coffee farms: landscape composition may affect delivery of an ecosystem service'. *Bird Conservation International* **21**(03), 353–361. DOI: 10.1017/s0959270910000614

Jones KE, Bininda-Emonds OR & Gittleman JL (2005). 'Bats, clocks, and rocks: diversification patterns in Chiroptera'. *Evolution* **59**(10), 2243–2255. DOI: 10.1554/04-635.1

Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL & Daszak P (2008). 'Global trends in emerging infectious diseases'. *Nature* **451**(7181), 990–993. DOI: 10.3201/eid0903.020336

Jones KE, Russ JA, Bashta AT, Bilhari Z, Catto C, Csősz I, Gorbachev A, Győrfi P, Hughes A, Ivashkiv I, Koryagina N, Kurali A, Langton S, Collen A, Margiean G, Pandourski I, Parsons S, Prokofev I, Szodoray-Paradi A, Szodoray-Paradi F, Tilova E, Walters CL, Weatherill A & Zavarzin O (2011). 'Indicator Bats Program: a system for the global acoustic monitoring of bats'. *Biodiversity Monitoring and Conservation: Bridging the Gap between Global Commitment and Local Action*, 211–247. DOI: 10.1002/9781118490747.ch10

Jones KE, Bielby J, Cardillo M, Fritz SA, O'Dell J, Orme CDL, Safi K, Sechrest W, Boakes EH, Carbone C et al. (2009). 'PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals: Ecological Archives E090-184'. *Ecology* **90**(9), 2648–2648. DOI: 10.1890/08-1494.1

Kamiya T, O'Dwyer K, Nakagawa S & Poulin R (2014). 'What determines species richness of parasitic organisms? A meta-analysis across animal, plant and fungal hosts'. *Biological Reviews* **89**(1), 123–134. DOI: 10.1111/brv.12046

Karanth K (1995). 'Estimating tiger (*Panthera tigris*) populations from camera-trap data using capture–recapture models'. *Biological Conservation* **71**(3), 333–338. DOI: 10.1016/0006-3207(94)00057-w

Karrer B & Newman M (2011). 'Competing epidemics on complex networks'. *Physical Review E* **84**(3), 036106. DOI: 10.1103/physreve.84.036106

Kelly MJ, Betsch J, Wultsch C, Mesa B & Mills LS (2012). 'Noninvasive sampling for carnivores.' In: *Carnivore ecology and conservation: a handbook of techniques*. Ed. by L Boitani & R Powell. New York: Oxford University Press, 47–69. DOI: 10.1093/acprof:oso/9780199558520.003.0004

Kerth G, Mayer F & Petit E (2002). 'Extreme sex-biased dispersal in the communally breeding, nonmigratory Bechstein's bat (*Myotis bechsteinii*)'. *Molecular Ecology* **11**(8), 1491–1498. DOI: 10.1046/j.1365-294x.2002.01528.x

Kerth G (2008). 'Causes and consequences of sociality in bats'. *Bioscience* **58**(8), 737–746. DOI: 10.1641/B580810

Kerth G, Perony N & Schweitzer F (2011). 'Bats are able to maintain long-term social relationships despite the high fission–fusion dynamics of their groups'. *Proceedings of the Royal Society of London B: Biological Sciences* **278**(1719), 2761–2767

Kessel S, Cooke S, Heupel M, Hussey N, Simpfendorfer C, Vagle S & Fisk A (2014). 'A review of detection range testing in aquatic passive acoustic telemetry studies'. *Reviews in Fish Biology and Fisheries* **24**(1), 199–218. DOI: 10.1007/s11160-013-9328-4

Kimura S, Akamatsu T, Dong L, Wang K, Wang D, Shibata Y & Arai N (2014). 'Acoustic capture-recapture method for towed acoustic surveys of echolocating porpoises'. *The Journal of the Acoustical Society of America* **135**(6), 3364–3370. DOI: 10.1121/1.4875710

King AM, Adams MJ & Lefkowitz EJ (2011). *Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses*. Vol. 9. San Diego, CA: Elsevier. ISBN: 978-0123846846

Kirupaharan N & Allen LJ (2004). 'Coexistence of multiple pathogen strains in stochastic epidemic models with density-dependent mortality'. *Bulletin of mathematical biology* **66**(4), 841–864

Knobler S, Lemon S, Mack A, Sivitz L & Oberholtzer K (2004). *Learning from SARS: Preparing for the Next Disease Outbreak: Workshop Summary*. National Academies Press Washington, DC, USA: DOI: 10.1016/j.socscimed.2006.08.004

Kuzmin IV, Bozick B, Guagliardo SA, Kunkel R, Shak JR, Tong S & Rupprecht CE (2011). 'Bats, emerging infectious diseases, and the rabies paradigm revisited'. *Emerging health threats journal* **4**. DOI: 10.3402/ehtj.v4i0.7159

Lefebvre A, Fiet C, Belpois-Duchamp C, Tiv M, Astruc K & Glélé LA (2014). 'Case fatality rates of Ebola virus diseases: a meta-analysis of World Health Organization data'. *Médecine et maladies infectieuses* **44**(9), 412–416. DOI: 10.1016/j.medmal.2014.08.005

Leroy EM, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, Yaba P, Délicat A, Paweska JT, Gonzalez JP & Swanepoel R (2005). 'Fruit bats as reservoirs of Ebola virus'. *Nature* **438**(7068), 575–576

Lewis T, Gillespie D, Lacey C, Matthews J, Danbolt M, Leaper R, McLanaghan R & Moscrop A (2007). 'Sperm whale abundance estimates from acoustic surveys of the Ionian Sea and Straits of Sicily in 2003'. *Journal of the Marine Biological Association of the United Kingdom* **87**(01), 353–357. DOI: 10.1111/j.1748-7692.2005.tb01242.x

Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, Wang H, Crameri G, Hu Z, Zhang H, Zhang J, McEachern J, Field H, Daszak P, Eaton BT, Zhang S & Wang LF (2005). 'Bats are natural reservoirs of SARS-like coronaviruses'. *Science* **310**(5748), 676–679. DOI: 10.1126/science.1118391

Li XZ, Liu JX & Martcheva M (2010). 'An age-structured two-strain epidemic model with super-infection'. *Math Biosci Eng* **7**(1), 123–147

Lim BK & Engstrom MD (2001). 'Bat community structure at Iwokrama forest, Guyana'. *Journal of Tropical Ecology* **17**(05), 647–665. DOI: 10.1017/s0266467401001481

Lloyd-Smith JO, Cross PC, Briggs CJ, Daugherty M, Getz WM, Latto J, Sanchez MS, Smith AB & Swei A (2005). 'Should we expect population thresholds for wildlife disease?' *Trends in Ecology & Evolution* **20**(9), 511–519. DOI: 10.1016/j.tree.2005.07.004

López-Baucells A, Rocha R, Fernández-Arellano G, Bobrowiec PED, Palmeirim JM & Meyer CF (2014). 'Echolocation of the big red bat *Lasiurus egregius* (Chiroptera: Vespertilionidae) and first record from the Central Brazilian Amazon'. *Studies on Neotropical Fauna and Environment* **49**(1), 18–25. DOI: 10.1080/01650521.2014.907600

Luby SP, Hossain MJ, Gurley ES, Ahmed BN, Banu S, Khan SU, Homaira N, Rota PA, Rollin PE, Comer JA et al. (2009). 'Recurrent zoonotic transmission of Nipah virus into humans, Bangladesh, 2001–2007'. *Emerging infectious diseases* **15**(8), 1229. DOI: 10.3201/eid1508.081237

Lucas TCD (2015). *palettetown: Use Pokemon Inspired Colour Palettes*. R package version 0.1.0. URL: http://CRAN.R-project.org/package=palettetown

Lucas TCD, Moorcroft EA, Freeman R, Rowcliffe JM & Jones KE (2015). 'A generalised random encounter model for estimating animal density with remote sensor data'. *Methods in Ecology and Evolution* **6**(5), 500–509. DOI: 10.1111/2041-210X.12346

Luis AD, Hayman DTS, O'Shea TJ, Cryan PM, Gilbert AT, Pulliam JRC, Mills JN, Timonin ME, Willis CKR, Cunningham AA, Fooks AR, Rupprecht CE, Wood JLN & Webb CT (2013). 'A comparison of bats and rodents as

reservoirs of zoonotic viruses: are bats special?' *Proceedings of the Royal Society B: Biological Sciences* **280**(1756), 20122753. DOI: 10.1098/rspb.2012.2753

Ma J, Jones G, Zhang S, Shen J, Metzner W, Zhang L & Liang B (2003). 'Dietary analysis confirms that Rickett's big-footed bat (*Myotis ricketti*) is a piscivore'. *Journal of Zoology* **261**(03), 245–248. DOI: 10.1017/s095283690300414x

MacKenzie DI & Royle JA (2005). 'Designing occupancy studies: general advice and allocating survey effort'. *Journal of Applied Ecology* **42**(6), 1105–1114. DOI: 10.1111/j.1365-2664.2005.01098.x

Maganga GD, Bourgarel M, Vallo P, Dallo TD, Ngoagouni C, Drexler JF, Drosten C, Nakouné ER, Leroy EM & Morand S (2014). 'Bat distribution size or shape as determinant of viral richness in african bats'. *PloS one* **9**(6), e100172. DOI: 10.1371/journal.pone.0100172

Marcoux M, Auger-Méthé M, Chmelnitsky EG, Ferguson SH & Humphries MM (2011). 'Local passive acoustic monitoring of narwhal presence in the Canadian Arctic: a pilot project'. *Arctic* **64**(3), 307–316. DOI: 10.14430/arctic4121

Marques TA, Munger L, Thomas L, Wiggins S & Hildebrand JA (2011). 'Estimating North Pacific right whale *Eubalaena japonica* density using passive acoustic cue counting'. *Endangered Species Research* **13**(3), 163–172. DOI: 10.3354/esr00325

Marques TA, Thomas L, Martin SW, Mellinger DK, Ward JA, Moretti DJ, Harris D & Tyack PL (2013). 'Estimating animal population density using passive acoustics'. *Biological Reviews* **88**(2), 287–309. DOI: 10.1111/brv.12001

Martcheva M & Li XZ (2013). 'Competitive exclusion in an infection-age structured model with environmental transmission'. *Journal of Mathematical Analysis and Applications* **408**(1), 225–246. DOI: 10.1016/j.jmaa.2013.05.064

May RM & Anderson RM (1979). 'Population biology of infectious diseases: Part II'. *Nature* **280**(5722), 455–461. DOI: 10.1038/280455a0

May RM & Lloyd AL (2001). 'Infection dynamics on scale-free networks'. *Physical Review E* **64**(6), 066112. DOI: 10.1103/PhysRevE.64.066112

May RM & Nowak MA (1994). 'Superinfection, metapopulation dynamics, and the evolution of diversity'. *Journal of Theoretical Biology* **170**(1), 95–114. DOI: 10.1006/jtbi.1994.1171

McCallum H, Barlow N & Hone J (2001). 'How should pathogen transmission be modelled?' *Trends in ecology & evolution* **16**(6), 295–300. DOI: 10.1016/S0169-5347(01)02144-9

McCracken GF & Bradbury JW (1981). 'Social organization and kinship in the polygynous bat *Phyllostomus hastatus*'. *Behavioral Ecology and Sociobiology* **8**(1), 11–34. DOI: 10.1007/bf00302840

Metcalf C, Birger R, Funk S, Kouyos R, Lloyd-Smith J & Jansen V (2015). 'Five challenges in evolution and infectious diseases'. *Epidemics* **10**, 40–44. DOI: 10.1016/j.epidem.2014.12.003

Morand S (2000). 'Wormy world: comparative tests of theoretical hypotheses on parasite species richness'. In: *Evolutionary Biology of Host-Parasite Relationships*. Ed. by SM R. Poulin & A Skorping. Amsterdam: Elsevier, 63–79. DOI: 10.1017/S0031182001001135

Nowak MA & May RM (1994). 'Superinfection and the evolution of parasite virulence'. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **255**(1342), 81–89. DOI: 10.1098/rspb.1994.0012

Nunes A, da Gama MT & Gomes M (2006). 'Localized contacts between hosts reduce pathogen diversity'. *Journal of theoretical biology* **241**(3), 477–487. DOI: 10.1016/j.jtbi.2005.12.010

Nunn CL, Altizer S, Jones KE & Sechrest W (2003). 'Comparative tests of parasite species richness in primates'. *The American Naturalist* **162**(5), 597–614. DOI: 10.1086/378721

Nunn CL, Jordán F, McCabe CM, Verdolin JL & Fewell JH (2015). 'Infectious disease and group size: more than just a numbers game'. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **370**(1669), 20140111. DOI: 10.1098/rstb.2014.0111

O'Brien TG, Kinnaird MF & Wibisono HT (2003). 'Crouching tigers, hidden prey: Sumatran tiger and prey populations in a tropical forest landscape'. *Animal Conservation* **6**(2), 131–139. DOI: 10.1017/s1367943003003172

O'Donnell CF, Richter S, Dool S, Monks JM & Kerth G (2015). 'Genetic diversity is maintained in the endangered New Zealand long-tailed bat (Chalinolobus tuberculatus) despite a closed social structure and regular population crashes'. *Conservation Genetics*, 1–12. DOI: 10.1007/s10592-015-0746-9

Oleksy R, Racey PA & Jones G (2015). 'High-resolution GPS tracking reveals habitat selection and the potential for long-distance seed dispersal by Madagascan flying foxes *Pteropus rufus*'. *Global Ecology and Conservation* **3**, 678–692. DOI: 10.1016/j.gecco.2015.02.012

Olival KJ, Weekley CC & Daszak P (2015). 'Are bats really "special" as viral reservoirs? What we know and need to know'. In: *Bats and Viruses: A New*

Frontier of Emerging Infectious Diseases. Ed. by LF Wang & C Cowled. New Jersey: Wiley, 281–294. DOI: 10.1002/9781118818824.ch11

Orme D, Freckleton R, Thomas G, Petzoldt T, Fritz S, Isaac N & Pearse W (2012). *caper: Comparative Analyses of Phylogenetics and Evolution in R*. R package version 0.5. URL: http://CRAN.R-project.org/package=caper

Orr TJ & Zuk M (2013). 'Does delayed fertilization facilitate sperm competition in bats?' *Behavioral Ecology and Sociobiology* **67**(12), 1903–1913. DOI: 10.1007/s00265-013-1598-2

O'Shea TJ, Cryan PM, Cunningham AA, Fooks AR, Hayman DT, Luis AD, Peel AJ, Plowright RK & Wood JL (2014). 'Bat flight and zoonotic viruses'. *Emerging infectious diseases* **20**(5), 741. DOI: 10.3201/eid2005.130539

Owen SF, Menzel MA, Ford WM, Chapman BR, Miller KV, Edwards JW & Wood PB (2003). 'Home-range size and habitat used by the northern myotis (*Myotis septentrionalis*)'. *The American midland naturalist* **150**(2), 352–359. DOI: 10.1674/0003-0031(2003)150[0352:hsahub]2.0.co;2

Paradis E, Claude J & Strimmer K (2004). 'APE: analyses of phylogenetics and evolution in R language'. *Bioinformatics* **20**, 289–290. DOI: 10.1093/bioinformatics/btg412

Pastor-Satorras R, Castellano C, Van Mieghem P & Vespignani A (2015). 'Epidemic processes in complex networks'. *Rev. Mod. Phys.* **87** (3), 925–979. DOI: 10.1103/RevModPhys.87.925. URL: http://link.aps.org/doi/10.1103/RevModPhys.87.925

Pastor-Satorras R & Vespignani A (2001). 'Epidemic spreading in scale-free networks'. *Physical review letters* **86**(14), 3200. DOI: 10.1103/PhysRevLett.86.3200

Peel A, Baker K, Crameri G, Barr J, Hayman D, Wright E, Broder C, Fernández-Loras A, Fooks A, Wang L et al. (2012). 'Henipavirus Neutralising Antibodies in an Isolated Island Population of African Fruit Bats'. *PloS one* **7**(1), e30346

Peel AJ, Pulliam J, Luis A, Plowright R, O'Shea T, Hayman D, Wood J, Webb C & Restif O (2014). 'The effect of seasonal birth pulses on pathogen persistence in wild mammal populations'. *Proceedings of the Royal Society B: Biological Sciences* **281**(1786), 20132962. DOI: 10.1098/rspb.2013.2962

Peel AJ, Sargan DR, Baker KS, Hayman DT, Barr JA, Crameri G, Suu-Ire R, Broder CC, Lembo T, Wang LF et al. (2013). 'Continent-wide panmixia of an African fruit bat facilitates transmission of potentially zoonotic viruses'. *Nature communications* **4**. DOI: 10.1038/ncomms3770

Petit E & Mayer F (1999). 'Male dispersal in the noctule bat (Nyctalus noctula): where are the limits?' *Proceedings of the Royal Society of London B: Biological Sciences* **266**(1430), 1717–1722. DOI: 10.1098/rspb.1999.0837

Pinheiro J, Bates D, DebRoy S, Sarkar D & R Core Team (2015). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-122. DOI: 10.1371/journal.pone.0104012. URL: http://CRAN.R-project.org/package=nlme

Plowright RK, Foley P, Field HE, Dobson AP, Foley JE, Eby P & Daszak P (2011). 'Urban habituation, ecological connectivity and epidemic dampening: the emergence of Hendra virus from flying foxes (Pteropus spp.)' *Proceedings of the Royal Society B: Biological Sciences* **278**(1725), 3703–3712. DOI: 10.1098/rspb.2011.0522

Plowright RK, Eby P, Hudson PJ, Smith IL, Westcott D, Bryden WL, Middleton D, Reid PA, McFarlane RA, Martin G et al. (2015). 'Ecological dynamics of emerging bat virus spillover'. *Proceedings of the Royal Society of London B: Biological Sciences* **282**(1798), 20142124. DOI: 10.1098/rspb.2014.2124

Poletto C, Meloni S, Colizza V, Moreno Y & Vespignani A (2013). 'Host mobility drives pathogen competition in spatially structured populations'. *PLoS computational biology* **9**(8), e1003169. DOI: 10.1371/journal.pcbi.1003169

Poletto C, Meloni S, Van Metre A, Colizza V, Moreno Y & Vespignani A (2015). 'Characterising two-pathogen competition in spatially structured environments'. *Scientific reports* **5**. DOI: 10.1038/srep07895

Pons-Salort M, Serra-Cobo J, Jay F, Lopez-Roig M, Lavenir R, Guillemot D, Letort V, Bourhy H & Opatowski L (2014). 'Insights into persistence mechanisms of a zoonotic virus in bat colonies using a multispecies metapopulation model'. *PloS one* **9**(4), e95610. DOI: 10.1371/journal.pone.0095610

Poulin R (1995). 'Phylogeny, ecology, and the richness of parasite communities in vertebrates'. *Ecological Monographs*, 283–302. DOI: 10.2307/2937061

Poulin R (2010). 'Latitudinal gradients in parasite diversity: bridging the gap between temperate and tropical areas'. *Neotropical Helminthology* **4**(2), 169–177

Poulin R (2014). 'Parasite biodiversity revisited: frontiers and constraints'. *International journal for parasitology* **44**(9), 581–589. DOI: 10.1016/j.ijpara.2014.02.003

Poulin R & Morand S (2000). 'The diversity of parasites'. *Quarterly Review of Biology*, 277–293

Purvis A, Gittleman JL, Cowlishaw G & Mace GM (2000). 'Predicting extinction risk in declining species'. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **267**(1456), 1947–1952. DOI: 10.1098/rspb.2000.1234

Qiu Z, Kong Q, Li X & Martcheva M (2013). 'The vector–host epidemic model with multiple strains in a patchy environment'. *Journal of Mathematical Analysis and Applications* **405**(1), 12–36. DOI: 10.1016/j.jmaa.2013.03.042

R Development Core Team (2010). *R: A Language And Environment For Statistical Computing*. R Foundation For Statistical Computing. Vienna, Austria. URL: http://www.R-Project.org

Racey PA (2015). 'The uniqueness of bats'. In: *Bats and Viruses: A New Frontier of Emerging Infectious Diseases. Wiley Blackwell*. Ed. by LF Wang & C Cowled, 1–22. DOI: 10.1002/9781118818824.ch1

Rand D, Keeling M & Wilson H (1995). 'Invasion, stability and evolution to criticality in spatially extended, artificial host-pathogen ecologies'. *Proceedings of the Royal Society of London B: Biological Sciences* **259**(1354), 55–63. DOI: 10.1098/rspb.1995.0009

Revell LJ (2010). 'Phylogenetic signal and linear regression on species data'. *Methods in Ecology and Evolution* **1**(4), 319–329. DOI: 10.1111/j.2041-210x.2010.00044.x

Richter H & Cumming G (2008). 'First application of satellite telemetry to track African straw-coloured fruit bat migration'. *Journal of Zoology* **275**(2), 172–176. DOI: 10.1111/j.1469-7998.2008.00425.x

Rifkin JL, Nunn CL & Garamszegi LZ (2012). 'Do animals living in larger groups experience greater parasitism? A meta-analysis'. *The American naturalist* **180**(1), 70–82. DOI: 10.1086/666081

Rogers TL, Ciaglia MB, Klinck H & Southwell C (2013). 'Density can be misleading for low-density species: benefits of passive acoustic monitoring'. *Public Library of Science One* **8**(1), e52542. DOI: 10.1371/journal.pone.0052542

Rohani P, Green C, Mantilla-Beniers N & Grenfell B (2003). 'Ecological interference between fatal diseases'. *Nature* **422**(6934), 885–888. DOI: 10.1038/nature01542

Rovero F, Zimmermann F, Berzi D & Meek P (2013). '"Which camera trap type and how many do I need?" A review of camera features and study designs for a range of wildlife research applications.' *Hystrix* **24**(2), 148–156. DOI: 10.4404/hystrix-24.2-8789

Rowcliffe JM & Carbone C (2008). 'Surveys using camera traps: are we looking to a brighter future?' *Animal Conservation* **11**(3), 185–186

Rowcliffe J, Field J, Turvey S & Carbone C (2008). 'Estimating animal density using camera traps without the need for individual recognition'. *Journal of Applied Ecology* **45**(4), 1228–1236. DOI: 10.1111/j.1365-2664.2008.01473.x

Royle JA & Nichols JD (2003). 'Estimating abundance from repeated presence-absence data or point counts'. *Ecology* **84**(3), 777–790. DOI: 10.1890/0012-9658(2003)084[0777:eafrpa]2.0.co;2

Schielzeth H (2010). 'Simple means to improve the interpretability of regression coefficients'. *Methods in Ecology and Evolution* **1**(2), 103–113. DOI: 10.1111/j.2041-210x.2010.00012.x

Schmidt BR (2003). 'Count data, detection probabilities, and the demography, dynamics, distribution, and decline of amphibians'. *Comptes Rendus Biologies* **326**, 119–124. DOI: 10.1016/s1631-0691(03)00048-9

Shump KA & Shump AU (1982). 'Lasiurus borealis'. *Mammalian species*, 1–6. DOI: 10.2307/3503843

Slatkin M (1995). 'A measure of population subdivision based on microsatellite allele frequencies.' *Genetics* **139**(1), 457–462

Smith KF, Goldberg M, Rosenthal S, Carlson L, Chen J, Chen C & Ramachandran S (2014). 'Global rise in human infectious disease outbreaks'. *Journal of The Royal Society Interface* **11**(101), 20140950. DOI: 10.1098/rsif.2014.0950

Smouse PE, Focardi S, Moorcroft PR, Kie JG, Forester JD & Morales JM (2010). 'Stochastic modelling of animal movement'. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**(1550), 2201–2211. DOI: 10.1098/rstb.2010.0078

SymPy Development Team (2014). *SymPy: Python library for symbolic mathematics*. URL: http://www.sympy.org

Taylor LH, Latham SM & Mark E (2001). 'Risk factors for human disease emergence'. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **356**(1411), 983–989. DOI: 10.1093/acprof:oso/9780198525738.003.0035

Thomas L & Marques TA (2012). 'Passive acoustic monitoring for estimating animal density'. *Acoustics Today* **8**(3), 35–44. DOI: 10.1121/1.4753915

Turmelle AS & Olival KJ (2009). 'Correlates of viral richness in bats (order Chiroptera)'. *EcoHealth* **6**(4), 522–539. DOI: 10.1007/s10393-009-0263-8

Turner K & Garnett G (2002). 'The impact of the phase of an epidemic of sexually transmitted infection on the evolution of the organism'. *Sexually transmitted infections* **78**, i20–i30. DOI: 10.1136/sti.78.suppl_1.i20

Van de Bovenkamp R, Kuipers F & Van Mieghem P (2014). 'Domination-time dynamics in susceptible-infected-susceptible virus competition on networks'. *PHYSICAL REVIEW E Phys Rev E* **89**(042818), 042818

Vespignani A (2008). 'Reaction-diffusion processes and epidemic metapopulation models in complex networks'. *The European Physical Journal B* **64**(3), 349–353. DOI: 10.1140/epjb/e2008-00302-y

Vitone ND, Altizer S & Nunn CL (2004). 'Body size, diet and sociality influence the species richness of parasitic worms in anthropoid primates'. *Evolutionary Ecology Research* **6**(2), 183–199

Vonhof MJ, Russell AL & Miller-Butterworth CM (2015). 'Range-wide genetic analysis of little brown bat (*Myotis lucifugus*) populations: Estimating the risk of spread of white-nose syndrome'. *PloS one* **10**(7). DOI: 10.1371/journal.pone.0128713

Walters CL, Collen A, Lucas T, Mroz K, Sayer CA & Jones KE (2013). 'Challenges of using bioacoustics to globally monitor bats'. In: *Bat Evolution, Ecology, and Conservation*. Ed. by RA Adams & SC Pedersen. Springer, 479–499. DOI: 10.1007/978-1-4614-7397-8_23

Walters CL, Freeman R, Collen A, Dietz C, Brock Fenton M, Jones G, Obrist MK, Puechmaille SJ, Sattler T, Siemers BM, Parsons S & Jones KE (2012). 'A continental-scale tool for acoustic identification of European bats'. *Journal of Applied Ecology* **49**(5), 1064–1074. DOI: 10.1111/j.1365-2664.2012.02182.x

Wang L, Walker P & Poon L (2011). 'Mass extinctions, biodiversity and mitochondrial function: are bats 'special' as reservoirs for emerging viruses?' *Current Opinion in Virology*. DOI: 10.1016/j.coviro.2011.10.013

Wang W, Chen M, Min Y & Jin X (2016). 'Structural diversity effects of multilayer networks on the threshold of interacting epidemics'. *Physica A: Statistical Mechanics and its Applications* **443**, 254–262. DOI: 10.1016/j.physa.2015.09.064

Wardeh M, Risley C, McIntyre MK, Setzkorn C & Baylis M (2015). 'Database of host-pathogen and related species interactions, and their global distribution'. *Scientific data* **2**. DOI: 10.1038/sdata.2015.49

Whittingham MJ, Stephens PA, Bradbury RB & Freckleton RP (2006). 'Why do we still use stepwise modelling in ecology and behaviour?' *Journal of animal ecology* **75**(5), 1182–1189. DOI: 10.1111/j.1365-2656.2006.01141.x

Whittingham MJ, Swetnam RD, Wilson JD, Chamberlain DE & Freckleton RP (2005). 'Habitat selection by yellowhammers *Emberiza citrinella* on lowland farmland at two spatial scales: implications for conservation management'. *Journal of applied ecology* **42**(2), 270–280. DOI: 10.1111/j.1365-2664.2005.01007.x

Wickham H (2009). *ggplot2: elegant graphics for data analysis*. Springer New York. ISBN: 978-0-387-98140-6. URL: http://had.co.nz/ggplot2/book

Wickham H (2015). *rvest: Easily Harvest (Scrape) Web Pages*. R package version 0.2.0. URL: http://CRAN.R-project.org/package=rvest

Wilmer JW, Moritz C, Hall L & Toop J (1994). 'Extreme population structuring in the threatened ghost bat, Macroderma gigas: evidence from mitochondrial DNA'. *Proceedings of the Royal Society of London B: Biological Sciences* **257**(1349), 193–198. DOI: 10.1098/rspb.1994.0115

Wilson DE & Reeder DM (2005). *Mammal species of the world: a taxonomic and geographic reference*. Vol. 12. JHU Press. ISBN: 9780801882210

Woolhouse ME & Gowtage-Sequeria S (2006). 'Host range and emerging and reemerging pathogens'. In: *Ending the War Metaphor:: The Changing Agenda for Unraveling the Host-Microbe Relationship-Workshop Summary*. Vol. 192. National Academies Press

World Bank (2014). *Update on the economic impact of the 2014 Ebola epidemic on Liberia, Sierra Leone, and Guinea*. Washington, DC. DOI: 10.1211/pj.2014.20066832. URL: http://documents.worldbank.org/curated/en/2014/12/20454884/update-economic-impact-2014-ebola-epidemic-liberia-sierra-leone-guinea

World Bank (2015). 'World Development Indicators: Size of the economy'. URL: http://wdi.worldbank.org/table/1.1

Wright SJ & Hubbell SP (1983). 'Stochastic extinction and reserve size: a focal species approach'. *Oikos*, 466–476. DOI: 10.2307/3544106

Wu Q, Zhang H, Small M & Fu X (2013). 'Threshold analysis of the susceptible-infected-susceptible model on overlay networks'. *Communications in Nonlinear Science and Numerical Simulation*

Xie Y (2015). *knitr: A General-Purpose Package for Dynamic Report Generation in R*. R package version 1.11. URL: http://CRAN.R-project.org/package=knitr

Yapp W (1956). 'The theory of line transects'. *Bird study* **3**(2), 93–104. DOI: 10.1080/00063655609475840

Yu G (2015). 'ggtree: an R package for versatile annotation and visualization of phylogenetic tree'

Zero VH, Sundaresan SR, O'Brien TG & Kinnaird MF (2013). 'Monitoring an endangered savannah ungulate, Grevy's zebra (*Equus grevyi*): choosing a method for estimating population densities'. *Oryx* **47**(03), 410–419. DOI: 10.1017/s0030605312000324

Zhang L, Jones G, Zhang J, Zhu G, Parsons S, Rossiter SJ & Zhang S (2009). 'Recent surveys of bats (Mammalia: Chiroptera) from China. I. Rhinolophidae and Hipposideridae'. *Acta Chiropterologica* **11**(1), 71–88. DOI: 10.3161/150811010x504626