

Dissertação de Mestrado

**Uso de drones para estimativas de abundância:
como lidar com os erros de detecção**

ISMAEL VERRASTRO BRACK

Porto Alegre, Agosto de 2016

Uso de drones para estimativas de abundância: como lidar com os erros de detecção

ISMAEL VERRASTRO BRACK

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ecologia, Instituto de Biociências da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do título de Mestre em Ecologia.

Orientador: Prof. Dr. Luiz Flamarion Barbosa de Oliveira

Comissão Examinadora

Prof. Dr. Gonçalo Nuno Côrte-Real Ferraz de Oliveira

Dr. Murilo Guimarães Rodrigues

Prof. Dr. Luiz Gustavo Rodrigues Oliveira Santos

Porto Alegre, Agosto de 2016

*Cuando creíamos que teníamos todas las respuestas,
de pronto,
cambiaron todas las preguntas*

Mario Benedetti

poeta e escritor uruguayo

(1920 – 2009)

AGRADECIMENTOS

À UFRGS e ao PPG Ecologia, pelo suporte e estrutura. À Fundação Grupo Boticário pelo apoio financeiro.

À CAPES, pela bolsa. Mas não pelos cortes. Aliás, pra não esquecer: **Fora Temer!**

Ao Flamarion, pela orientação, por acreditar e topar fazer parte desse plano maluco.

Ao Andreas, pela parceria nessa coorientação informal. Por ter me colocado nessa enrascada. Por ter me confiado a missão de operar esse “brinquedinho”, sempre demonstrar confiança e ter me incentivado até o final.

Antecipadamente, aos membros da banca, Gustavo, Murilo e Gonçalo, pelas contribuições. E ao Gonçalo, por ter oportunizado e incentivado minha introdução ao fascinante e complexo mundo dos modelos hierárquicos.

Ao Alexandre Krob, e ao Curicaca, por ter me possibilitado a oportunidade com o “aviãozinho”.

Aos que me ajudaram em campo, Mi, Carol, Juan (e por ter resgatado o avião do fundo de um açude...) e especialmente o Fábio por ter compartilhado os perrengues e momentos de frustração.

Aos voluntários, Taís, Lari e Ezequiel, que dedicaram tempo para contar as vaquinhas (e ovelhas, cavalos... e... leões) e possibilitaram a concretização do primeiro capítulo.

Aos colegas do PPG por tornar as disciplinas e os dias na UFRGS mais leves, especialmente aos parceiros de mestrado Claudinho e Paulo. Ao pessoal do NERF, pela acolhida e aos flamarionetes, parceiros de modelagem, Igor e Juan, pelas discussões. Ao Vinícius Bastazini, pela revisão de texto e discussões sobre o modelo.

À gurizada nada menos do que incrível da bio: Claudinho, Paulo, Marcelão, Ronaldo, Alexis, Rafa, Lucas, Thay, Mila, Mari, Bruna, entre outros, pela amizade, diversão e pelos melhores carnavais. Ao Paulo, Marcelão e Ronaldo pela melhor autocervejaria que poderíamos ter! Sem dúvida, a dALE! foi o combustível desse mestrado.

Aos meus pais, pelo suporte e apoio incondicionais. Pelo exemplo de vida que são. À minha irmã Gabi, e minhas avós Consuelo e Marta. À toda a minha família uruguaia. E a minha nova família, Abadie-Vasconcellos.

À Mi, minha companheira de tudo, por sempre me fazer ver o mundo de um jeito melhor.

RESUMO

Abundância é uma das medidas centrais da ecologia, sendo empregada tanto para investigar questões ecológicas básicas quanto para avaliação questões de conservação e manejo. Porém, a abundância dificilmente pode ser acessada de maneira direta; diversas fontes de erro de detecção (ou medição) podem produzir vieses nas estimativas de abundância. Ao longo da história, diferentes técnicas de coleta e métodos de estimativa vêm sendo aplicados com objetivo de obter estimativas mais acuradas e precisas. Recentemente, os *drones* – ou veículos aéreos não tripulados – têm surgido como uma alternativa nos levantamentos aéreos de populações animais e vêm sendo aplicados em estudos de várias espécies. Entretanto, para a consolidação dessa nova ferramenta em estudos ecológicos, ainda são necessárias avaliações sobre a acurácia e precisão de suas estimativas. Nesta dissertação, verificamos as fontes de erro que podem produzir vieses nas estimativas de abundância em amostragens com *drones* e propomos uma abordagem que visa acomodar os principais erros de detecção. No primeiro capítulo, comparando estimativas com o tamanho populacional real conhecido, demonstramos que caso as fontes de erro sejam ignoradas, ou se não corrigidas adequadamente, as estimativas de abundância resultam fortemente enviesadas. No segundo capítulo, em um estudo simulado, propomos uma abordagem de modelagem hierárquica para acomodar os erros falsos negativos em levantamentos aéreos espacial e temporalmente replicados com *drones*, com revisão das imagens por múltiplos observadores. Assim, este estudo traz contribuições para o desenvolvimento de abordagens robustas que permitirão o uso dos *drones* para investigar questões ecológicas, monitorar populações em longo prazo, entre outras das aplicações das informações de abundância.

Palavras-chave: veículo aéreo não-tripulado, levantamento aéreo, tamanho populacional, erros de medição, detecção imperfeita, falsos negativos, falsos positivos, *N-mixture model*, modelo hierárquico

ABSTRACT

Abundance is one of the central measures of ecology and it has been used to investigate from basic ecological questions to conservation and management issues. However, abundance can rarely be directly accessed; there are sources of measurement errors that could produce biases in abundance estimates. Historically, different sampling techniques and estimation methods have been applied in order to get more accurate and precise abundance estimates. Recently, drones – or unmanned aerial systems (UAS) – emerged as an alternative in aerial surveys of animal populations and they have been used to study several species. However, accuracy and precision evaluations are still needed for the consolidation of this new ecological tool. In this dissertation, we assessed the sources of error that could produce biases in abundance estimates in drone surveys and we propose an approach to accommodate the main measurement errors. In the first chapter, we compared estimates with a population of known size and we demonstrated that if sources of error are ignored, or they are not adequately corrected, abundance estimates results can be strongly biased. In the second chapter, in a simulation study, we propose a hierarchical modeling approach to accommodate false negative errors in spatially replicated drone surveys, and image reviews by multiple observers. Then, this study brings contributions to the development of robust approaches that will allow the utilization of drones to investigate ecological questions, long-term population monitoring, among many other applications of abundance data.

Key-words: unmanned aerial system, aerial survey, population size, measurement errors, imperfect detection, false negative, false positive, N-mixture model, hierarchical model

SUMÁRIO

Introdução geral.....	5
Capítulo 1: Ignorar fontes de erro em levantamentos com drones resulta em estimativas de abundância enviesadas.....	10
Resumo.....	11
Introdução.....	11
Métodos.....	14
Resultados.....	16
Discussão.....	17
Referências.....	19
Material Suplementar.....	22
Capítulo 2: Uma abordagem de modelagem hierárquica para estimar abundância com drones.....	24
Resumo.....	25
Introdução.....	26
Métodos.....	28
Resultados.....	31
Discussão.....	33
Referências.....	36
Material Suplementar.....	40
Conclusões.....	46
Referências.....	48

LISTA DE FIGURAS

Capítulo 1

Figura 1. Fontes de erro de medição em levantamentos com *drones*. Números pretos são falsos positivos e números brancos, falsos negativos.....13

Figura 2. Imagem da área amostrada, gerada a partir das imagens obtidas com o sistema aéreo não tripulado Echar 20B. O ortomosaico foi construído a partir de aprox. 800 fotos tiradas a 300 m acima do solo, com resolução de 4,6 cm/px. Em destaque, zoom da área com algumas cabeças de gado.....15

WebFigura 1. Procedimento de captura-recaptura com três observadores independentes em imagem gerada por um sobrevoo com drone. A procura foi realizada no ArcGIS (ESRI). Em uma camada específica para cada observador (representados por cores diferentes), um ponto foi posicionado no centro dos indivíduos visualizados. A partir da sobreposição de buffers de 30cm ao redor dos pontos, um histórico de captura (1 = detectado; 0 = não detectado) foi gerado para cada indivíduo. Imagem coletada pelo sistema aéreo não tripulado Echar 20B a 300 m acima do solo, com resolução de 4,6 cm/px.....22

Capítulo 2

Figura 1. Viés relativo das estimativas de abundância local ($\lambda=4$) de um *multinomial N-mixture model* com emigração temporária, em cenários simulados de número de observadores e de réplicas temporais. Os boxplots indicam a distribuição de 999 vieses relativos, calculados a partir da diferença entre as médias posteriores e o valor real utilizado para gerar o conjunto de dados. Os limites inferior e superior da caixa indicam, respectivamente, os quartis de 25% e 75%, a linha preta no centro da caixa representa a mediana, os limites das hastes estendem as estimativas a 1,5 vezes os valores dos interquartis e os círculos são os valores extremos.....31

Figura 2. Coeficiente de variação das estimativas de abundância local ($\lambda=4$) de um *multinomial N-mixture model* com emigração temporária, em cenários simulados de número de observadores e de réplicas temporais. Os boxplots indicam a distribuição de 999 coeficientes de variação, calculados a partir do quociente entre o desvio padrão e a média posterior. A simbologia é a mesma da figura 1.....32

Figura 1 S2. Coeficiente de variação das estimativas de abundância local de um *multinomial N-mixture model* com emigração temporária em diferentes cenários de disponibilidade e abundância local, sob três estratégias de amostragem. Para uma amostragem de 200 voos, s1: R = 67, T = 3; s2: R = 50, T = 4; s3: R = 40, T = 5. Onde R = unidades espaciais e T = replicações temporais. Os boxplots indicam a distribuição de 300 coeficientes de variação, calculados a partir do quociente entre o desvio padrão e a média posterior.....45

LISTA DE TABELAS

Capítulo 1

Tabela 1. Comparação de diferentes métodos de estimativa de abundância aplicados à três levantamentos aéreos com drone, sobre uma população de 414 indivíduos de gado.....17

WebTabela 1. Estimativa do número de indivíduos disponíveis para detecção para os três sobrevoos de *drone*. Resultados dos modelos de captura-recaptura de população fechada para observadores múltiplos.....23

Capítulo 2

Tabela 1. Número de replicações espaciais (sítios amostrais) necessárias para atingir um coeficiente de variação menor que 20% em diferentes condições de abundância local (λ) e disponibilidade (ϕ)33

Tabela 1 S2. Cobertura do intervalo de credibilidade de 95% (IC95%) das estimativas de abundância local ($\lambda=4$) de um *multinomial N-mixture model* com emigração temporária, em cenários simulados de número de observadores e de réplicas temporais. A cobertura foi calculada a partir da proporção dos 999 IC95% que abrangeu o valor real utilizado para simular os dados.....44

INTRODUÇÃO GERAL

Abundância é uma das medidas fundamentais em ecologia. Informações de abundância são usadas para desvendar a relação dos indivíduos com seu ambiente biótico e abiótico, assim como a dinâmica dessas relações no tempo e no espaço, através dos diferentes níveis biológicos. Entre as diversas aplicações dos dados de abundância estão a definição do status de conservação de espécies, o monitoramento em longo prazo de populações, o planejamento do uso de estoques pesqueiros e o manejo de espécies exóticas invasoras. Acessar a abundância ajuda a elucidar padrões ecológicos e a embasar ações de conservação ou de manejo de espécies com interesse de uso pelo ser humano.

Porém, obter informações de abundância em populações silvestres não é uma tarefa fácil, principalmente em populações animais. Censos populacionais – ou seja, levantar todos os indivíduos de uma população – são, salvo raras exceções, impraticáveis nos diversos sistemas ecológicos. Como consequência, para estimar tamanhos populacionais, precisamos lançar mão de diferentes técnicas de amostragem e estatísticas, as quais muitas vezes estão associadas. A maneira talvez mais simples de medir a abundância, e historicamente mais utilizada, é simplesmente contar os indivíduos vistos em determinada área e em determinado intervalo de tempo. Entretanto, apesar de facilmente implementável em campo, e portanto ainda muito usada, essa abordagem pode apresentar graves problemas (Anderson 2001; Williams *et al.* 2002). Durante a amostragem de populações animais, o número de indivíduos observados dificilmente é o número real. **Erros de medição** (i.e. detecção imperfeita) produzem vieses sistemáticos nas estimativas como consequência de diversos fatores como o tipo de técnica de amostragem (p. ex., erro do observador ou falha de detecção por um dispositivo de coleta de dados), o momento da amostragem (p. ex., horário do dia ou condição do tempo), características da área amostrada (p. ex., dificuldade de locomoção ou obstrução da vegetação) e/ou características da espécie estudada (p. ex., comportamento, cor ou tamanho). Portanto, a utilização de contagens brutas como tamanho populacional absoluto ou relativo (no último caso, denominadas índices de abundância) exige assumir pressupostos muito improváveis de detecção completa ou constante no espaço e no tempo (Williams *et al.* 2002). Como decorrência da violação desses pressupostos, as estimativas resultam ser fortemente enviesadas (MacKenzie *et al.* 2002; Martin *et al.* 2005).

Ao longo da história, diversas abordagens foram propostas para tentar lidar de alguma forma com os erros de medição nas estimativas de abundância, inerentes à grande maioria dos estudos em ecologia animal. Todas partem de uma fórmula básica, denominada **forma canônica**:

$$C = N * \beta^1.$$

Em palavras, uma contagem (C) é resultado do produto do número de indivíduos em uma população (N) e da probabilidade de detectá-los (β). Uma evidente e poderosa interpretação dessa equação é que a informação de abundância coletada é resultado de dois processos: um processo biológico latente e um processo de observação que mascara o processo biológico. Assim, os dados coletados contêm informação sobre a variável latente (neste caso, a abundância), mas não a representam diretamente.

As inúmeras abordagens que lidam de alguma forma com a detecção imperfeita vão desde correções e calibrações com informações adicionais até a modelagem explícita dos processos que geram os dados observados (Williams *et al.* 2002; Royle & Dorazio 2006). Tradicionalmente, as fontes de erro no processo de observação são atribuídas às falsas ausências, também chamadas de falsos negativos ou falsos zeros (MacKenzie *et al.* 2002; Martin *et al.* 2005). As falsas ausências são causadas basicamente por dois processos: 1) ausência temporária de um indivíduo no sítio amostrado, durante a amostragem e 2) a detecção imperfeita propriamente dita. Contudo, há outro tipo de fonte de erro menos considerado nos estudos de ecologia animal, os falsos positivos. Os falsos positivos podem ser causados por erro de identificação de espécie ou por uma contagem dupla de um mesmo indivíduo (Royle & Link 2006; Guillera-Arroita *et al.* 2012). Enquanto os falsos negativos, amplamente consolidados na literatura, produzem vieses negativos nas estimativas de abundância, os falsos positivos, muitas vezes ignorados, podem resultar em vieses positivos significativos.

Diversas técnicas de amostragem estão disponíveis para estudos com populações animais, cada qual com diferentes influências das fontes de erro de medição. Essas técnicas podem ser divididas em duas grandes “famílias”: populações com indivíduos **marcados** ou **não**

¹ Uma versão mais completa da forma canônica pode também levar em conta, além da probabilidade de detecção, o fato de que geralmente amostramos somente uma fração (α) da área onde uma população está presente, resultando em $C = N * \beta * \alpha$ (Williams *et al.* 2002).

marcados. Os métodos de modelagem de população com indivíduos marcados, ou métodos de captura-recaptura, já são bem estabelecidos nos estudos ecológicos e é possível encontrar uma vasta literatura a respeito (Royle *et al.* 2014; Cooch & White 2016). Porém, nem sempre é possível aplicar tais estudos em campo. Geralmente, estudos de captura-recaptura exigem um intenso esforço amostral que muitas vezes não é viável e, quando viável, limitam os estudos a escalas relativamente pequenas. Além disso, as técnicas de marcação comumente são invasivas, podendo causar impacto sobre as populações (Murray & Fuller 2000). Por outro lado, os estudos com animais não marcados resultam em uma alternativa sem a necessidade de identificar individualmente os animais. São exemplos de métodos para animais não marcados a amostragem com distâncias (*distance sampling*; Buckland *et al.* 2001), modelagem de ocupação de sítios (Mackenzie *et al.* 2006) e contagens replicadas espacial e temporalmente (Royle 2004; Kéry *et al.* 2005).

O que os métodos exemplificados acima têm em comum é que modelam explicitamente os processos subjacentes aos dados, com o objetivo de realizar inferências robustas sobre variáveis biológicas latentes (como abundância e distribuição de populações) e os parâmetros demográficos associados (p. ex., sobrevivência, recrutamento, extinção e colonização). Esses modelos são compostos, na mesma lógica da forma canônica, por pelo menos dois níveis: um para o processo biológico e outro para o processo de observação.

Recentemente, a modelagem hierárquica tem agrupado esses diversos métodos sob uma mesma abordagem para modelar a dinâmica espaço-temporal das variáveis biológicas de interesse, ao mesmo tempo em que leva em conta as possíveis fontes de erro na observação (Royle & Dorazio 2008; Kéry & Schaub 2012). A estrutura dos modelos hierárquicos é extremamente flexível podendo integrar mais de um método de coleta de dados e modelar diversas fontes de erro em um mesmo modelo (Kéry & Royle 2016). Além disso, os modelos podem ser implementados sob diferentes perspectivas estatísticas, ou seja, a abordagem por máxima verossimilhança ou bayesiana (Kéry & Royle 2016). Por outro lado, para poder separar os processos subjacentes aos dados, esses modelos exigem, na maioria dos casos, replicações temporais e/ou espaciais (Kéry & Royle 2016). Adicionalmente, para atingir bons níveis de precisão, os modelos hierárquicos demandam uma grande quantidade de informação, que aumenta conforme aumenta a complexidade de sua estrutura.

Entre as técnicas utilizadas para amostragem de indivíduos não marcados, os **levantamentos aéreos** têm sido utilizados em situações de amostragem de áreas extensas remotas e/ou inacessíveis. Tradicionalmente, essa técnica é aplicada para acessar a abundância de animais de porte grande que ocupam áreas abertas como, por exemplo, grandes mamíferos terrestres (e.g. Ferreira & Aarde 2009; Griffin *et al.* 2013), mamíferos marinhos (e.g. Pollock *et al.* 2006; Edwards *et al.* 2007), aves aquáticas (e.g. Rodgers *et al.* 2005) e crocodilianos (e.g. Mourão *et al.* 2000).

As principais fontes de erro nas medidas de abundância com levantamentos aéreos são **i)** indivíduos escondidos, que estão indisponíveis para serem visualizados (p. ex., debaixo de árvores ou mergulhados) e **ii)** dado que o indivíduo está disponível, não ser visualizado pelo observador (Pollock *et al.* 2006). Caso não consideradas, essas fontes de erro podem gerar grandes vieses – denominados, respectivamente, viés da disponibilidade e viés da visualização – nas estimativas de abundância. O viés da visualização é facilmente corrigido através de métodos de amostragem com distâncias (Buckland *et al.* 2001) ou observadores múltiplos (Williams *et al.* 2002). Por outro lado, a disponibilidade é pouco acessada em levantamentos aéreos e, quando acessada, utiliza informações adicionais, as quais raramente são factíveis de obtenção (p. ex., biotelemetria, indivíduos marcados, Edwards *et al.* 2007).

O recente desenvolvimento de sistemas aéreos remotamente pilotados (RPAS), conhecidos popularmente como **drones**, tem tornado o uso dessa ferramenta uma alternativa para os levantamentos aéreos convencionais com aviões tripulados. Os **drones** permitem amostragens seguras e com logística simplificada, ao mesmo tempo em que podem reduzir os custos por voo (Anderson & Gaston 2013). A praticidade e versatilidade dos **drones** prometem revolucionar a maneira pela qual coletamos informações ecológicas (Anderson & Gaston 2013; Christie *et al.* 2016). Por outro lado, ainda existem algumas limitações para o uso de RPAS como restrições legais e burocráticas nas autorizações e uma baixa autonomia em comparação a aviões convencionais.

Diversos estudos já foram publicados provando a utilidade dos **drones** para amostrar diversas espécies (Linchant *et al.* 2015; Hodgson *et al.* 2016). Entretanto, a maioria das estimativas de abundância geradas utilizou contagens brutas e poucos estudos aplicaram correções para levar em conta alguma fonte de erro (Linchant *et al.* 2015). O surgimento dessa

nova tecnologia para estudos de ecologia populacional permite o desenvolvimento e a aplicação de métodos mais robustos para estimar populações, como é o caso da modelagem hierárquica.

O objetivo desta dissertação de mestrado é i) avaliar a influência das possíveis fontes de erro em levantamentos aéreos de populações animais com *drones* e ii) propor uma abordagem geral que leve em conta esses erros de detecção para estimar com robustez o tamanho populacional. Cada objetivo corresponde a um capítulo da dissertação. O primeiro capítulo, intitulado “*IGNORAR FONTES DE ERRO EM LEVANTAMENTOS COM DRONES RESULTA EM ESTIMATIVAS DE ABUNDÂNCIA ENVIESADAS*”, pretende demonstrar que se não considerarmos as fontes de erro nas amostragens com *drones*, ou se não as corrigimos adequadamente, as estimativas de abundância resultam fortemente enviesadas. Para isso, conduzimos três sobrevoos sobre uma população de gado de tamanho real conhecido e aplicamos os principais métodos utilizados para estimar tamanhos populacionais com *drones*. No segundo capítulo, “*UMA ABORDAGEM DE MODELAGEM HIERÁRQUICA PARA ESTIMAR ABUNDÂNCIA COM DRONES*”, propomos uma abordagem de modelagem hierárquica para estimativas populacionais com *drones*, que modela explicitamente as fontes de erro de detecção. Nesse estudo, avaliamos o desempenho dessa abordagem em diversos cenários de amostragem simulados e provemos algumas recomendações gerais para planejar os levantamentos com *drones*.

CAPÍTULO 1

Ignorar fontes de erro em levantamentos com drones resulta em estimativas de abundância enviesadas

*Artigo a ser submetido como Research
Communication no periódico Frontiers
in Ecology and the Environment.*

IGNORAR FONTES DE ERRO EM LEVANTAMENTOS COM DRONES RESULTA EM ESTIMATIVAS DE ABUNDÂNCIA ENVIESADAS

Ismael V Brack¹, Andreas Kindel², Luiz FB Oliveira³

¹ PPG Ecologia UFRGS; ² Dep. Ecologia UFRGS; ³ Dep. Vertebrados MN/UFRJ

RESUMO

Drones tem emergido como uma ferramenta versátil que promete facilitar a coleta de dados em estudos ecológicos. A consolidação dos *drones* para levantamentos aéreos exige uma avaliação da influência dos erros de medição, falsos negativos e falsos positivos, para estimativas de abundância. Aqui, nós verificamos a influência das possíveis fontes de erro em amostragem com *drones* e comparamos métodos de estimativa de abundância em uma população de gado com tamanho real conhecido. Surpreendentemente, todos os métodos de estimativa apresentaram vieses negativos, causados principalmente devido aos indivíduos escondidos. O desenvolvimento de novas abordagens estatísticas, que levem em conta as diversas fontes de erro em amostragens com *drones*, é imprescindível para a obtenção de estimativas robustas de abundância.

Palavras-chave: sistema aéreo não tripulado, tamanho populacional, falso negativo, falso positivo, detecção imperfeita

INTRODUÇÃO

A ampla disseminação dos sistemas de aeronaves remotamente pilotadas (RPAS), popularmente conhecidos como *drones*, tem possibilitado o uso dessa ferramenta nas mais variadas áreas da ecologia (Anderson and Gaston 2013; Linchant *et al.* 2015). A versatilidade dos *drones* facilita a obtenção de informações com maior resolução espacial e temporal, ao mesmo tempo em que permite o uso de múltiplos sensores (eg RGB, termal, infravermelho próximo; Christie *et al.* 2016). Essas vantagens tornam os *drones* uma alternativa eficiente na coleta de dados de abundância para suas diversas aplicações, que vão desde investigar questões ecológicas básicas até monitorar populações e planejar ações de manejo e conservação.

Porém não podemos cair na tentação de pensar que com os levantamentos aéreos com *drones* todos nossos problemas estarão resolvidos. Durante uma amostragem, os dados

coletados são influenciados por processos tanto biológicos quanto de medição e devemos conseguir diferenciá-los para alcançar conclusões biológicas confiáveis (Williams *et al.* 2002). Os erros de medição (ie detecção imperfeita) são inerentes à maioria dos estudos em ecologia animal e, caso não sejam considerados, podem resultar em estimativas de abundância fortemente enviesadas (Mackenzie *et al.* 2006; Royle and Dorazio 2008). Ignorar esses vieses acarreta em inferências biológicas fracas ou até mesmo equivocadas.

Os erros de medição em levantamentos aéreos são atribuídos comumente aos falsos negativos e são divididos em dois componentes: a disponibilidade de um indivíduo para ser detectado (eg não escondido, sob dossel ou mergulhado) e a sua visualização, dado que ele esteja disponível (Figura 1) (Pollock *et al.* 2006). Por outro lado, os falsos positivos (eg erro de identificação de espécie, por outra espécie ou mancha no solo) costumam ser ignorados, assumindo que esses erros não ocorrem. Entretanto, a maneira como as imagens são coletadas pelos *drones* pode gerar falsos positivos significativos. Para realizar o mapeamento e construir o mosaico de imagens, as fotos consecutivas e as linhas de voo geralmente são sobrepostas, podendo gerar contagens duplas de um mesmo indivíduo (Figura 1) (Hodgson *et al.* 2013).

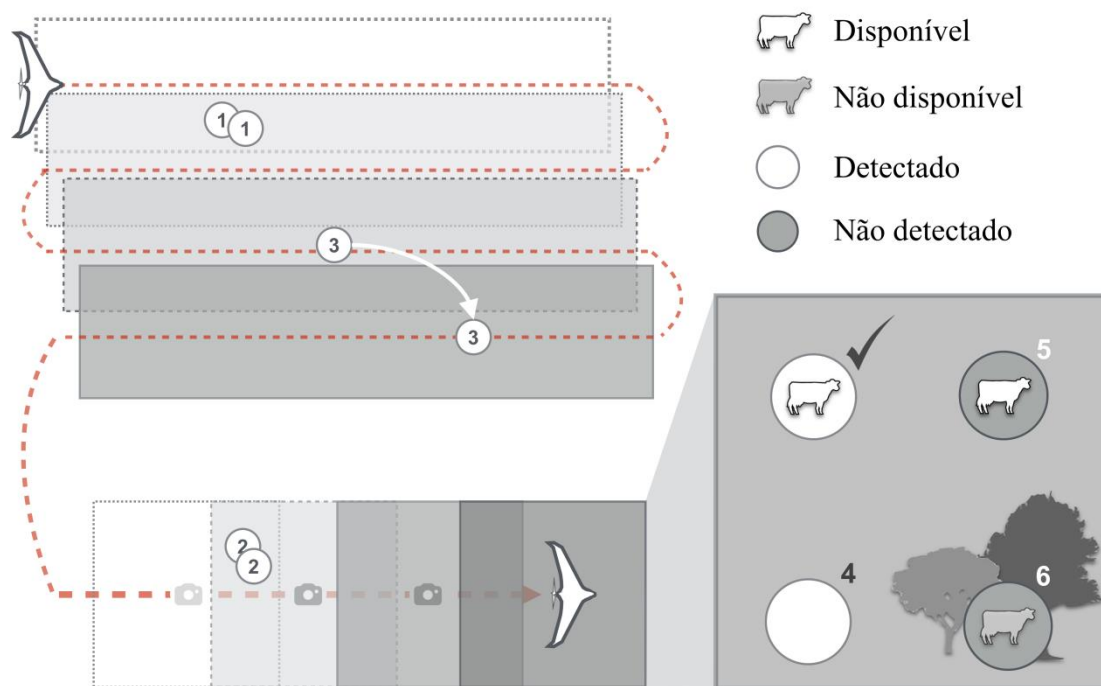


Figura 1. Fontes de erro de medição em levantamentos com *drones*. Números pretos são falsos positivos e números brancos, falsos negativos. **1)** Contagem dupla de um mesmo indivíduo devido à sobreposição entre as linhas de voo; **2)** Contagem dupla de um mesmo indivíduo devido à sobreposição entre fotos sequenciais; **3)** Contagem dupla de um mesmo indivíduo devido ao movimento entre as linhas de voo; **4)** Erro de identificação da espécie; **5)** Indivíduo disponível não detectado – viés de visualização; **6)** Indivíduo não disponível para ser detectado – viés de disponibilidade. O símbolo *check* indica uma detecção correta. Arte por Juan Rodriguez.

Na contramão da atual tendência dos estudos ecológicos (Gimenez *et al.* 2014), a maioria dos tamanhos populacionais acessados utilizando *drones* é baseada em simples contagens brutas (Linchant *et al.* 2015). Contagens brutas feitas com *drones* podem produzir estimativas mais precisas quando comparadas com outros métodos (Hodgson *et al.* 2016). No entanto, os resultados ainda são inconsistentes com relação a sua acurácia (Christie *et al.* 2016). Poucos estudos consideraram que a revisão das imagens pode gerar contagens imperfeitas devido ao erro do observador e corrigiram esse viés de visualização (erro 5) utilizando observadores múltiplos (Vermeulen *et al.* 2013). A indisponibilidade dos indivíduos (erro 6) é frequentemente citada como uma possível fonte de erro nas amostragens com *drones* (Martín *et al.* 2012; Linchant *et al.* 2015; Chrétien *et al.* 2016), porém raras vezes é levada em conta nas estimativas. Barasona *et al.* (2014) e Mulero-Pázmány *et al.* (2015) aplicaram uma correção

às contagens para levar em conta os indivíduos escondidos (ie indisponíveis) sob vegetação. Para isso, utilizaram um “coeficiente de detecção” (daqui em diante, coeficiente de disponibilidade) calculado a partir da classificação de pontos aleatorizados nas imagens. Quanto aos falsos positivos, alguns estudos mencionam medidas de precaução para evitar erros de identificação (erro 4), enquanto Hodgson *et al.* (2013) aplicaram uma correção para contagens duplas em áreas de sobreposição de imagens (erros 1 e 2).

A consolidação dos *drones* para a coleta de dados em estudos ecológicos exige entendimento de como essas fontes de erro podem afetar as estimativas de parâmetros populacionais. Neste estudo, nós avaliamos a influência das possíveis fontes de erro de medição em estimativas de abundância com *drones*. Para isso, aplicamos, em uma população de gado de tamanho real conhecido, métodos de estimativa comumente usados em amostragens com *drones*. Nós demonstramos que, caso as fontes de erro não sejam consideradas, as estimativas de abundância podem apresentar grandes vieses. Além disso, se o objetivo é atingir estimativas acuradas, nenhum dos métodos empregados até o momento se mostrou confiável.

MÉTODOS

Amostramos uma área de 470 ha, localizada na Estação Experimental Agronômica da Universidade Federal do Rio Grande do Sul (30°6’S 51°41’W; Figura 2), no sul do Brasil. A área é composta por pastagens nativas com a presença de mata de galeria e algumas árvores esparsas. Durante os sobrevoos, havia na área 414 indivíduos bovinos e equinos de diversas pelagens e tamanhos.

Conduzimos três sobrevoos de aproximadamente 40 minutos sobre a área; às 11h30min, 13h e 15h do dia 01 de março de 2016. Os voos foram realizados com uma aeronave remotamente pilotada de asa-fixa, o Echar 20B (por Xrobots, São Carlos, SP, Brasil). Os sobrevoos foram executados a partir de missões pré-programadas a uma velocidade média de 70 km/h e a 300 m de altura acima do solo. A missão foi programada para que as imagens apresentassem 60 % de sobreposição frontal (entre fotos sequenciais) e 50 % de sobreposição lateral (entre linhas de voo). No momento da foto, o sistema da aeronave registrou as coordenadas geográficas e informações dos ângulos (*yaw*, *pitch* e *roll*) para subsidiar a montagem do ortomosaico de fotos. As fotos foram capturadas utilizando uma câmera digital

embarcada SRL (Canon EOS Rebel T3i) de 18 Mpx (5184 x 3456 px) com uma distância focal de 28 mm (Lente Canon EF-S 18-55 mm). A resolução no nível do solo (*ground sampled distance* - GSD) foi de 4,6 cm/px. As aproximadamente 800 fotos obtidas em cada voo foram fundidas em ortomosaicos georreferenciados (Figura 2) no *software* Agisoft PhotoScan 1.2.4. O procedimento de montagem do ortomosaico exclui áreas sobrepostas das imagens e, portanto, evita contagem dupla de um mesmo indivíduo (erros 1 e 2).

Três observadores revisaram independentemente os ortomosaicos de cada sobrevoo no ArcGIS 10.3 (ESRI). A procura foi guiada por uma grade de quadrículas de 100 x 100 m e os observadores foram estimulados a procurar livremente os indivíduos em cada quadrícula utilizando zoom (Figura Suplementar 1). Os revisores foram instruídos a evitar falsos positivos, ou seja, somente marcar um indivíduo quando houvesse certeza. Bovinos e quinos foram considerados sem diferenciação, devido à impossibilidade de diferenciar alguns indivíduos nessa resolução no nível de solo.



Figura 2. Imagem da área amostrada, gerada a partir das imagens obtidas com o sistema aéreo não tripulado Echar 20B. O ortomosaico foi construído a partir de aprox. 800 fotos tiradas a 300 m acima do solo, com resolução de 4,6 cm/px. Em destaque, zoom da área com algumas cabeças de gado.

Nós comparamos os métodos de estimativa de abundância que já foram utilizados em estudos com *drones* com o tamanho populacional real conhecido para a área. Para as contagens dos observadores em cada sobrevoo, aplicamos três diferentes métodos: i) contagem máxima, número total de indivíduos detectados pelos três observadores; ii) observadores múltiplos, estimativa gerada por modelos de captura-recaptura de população fechada (Williams *et al.* 2002); iii) contagem corrigida, contagem máxima corrigida pelo coeficiente de disponibilidade. O coeficiente de disponibilidade foi calculado a partir da proporção de pontos detectáveis gerados sobre a área amostrada. Utilizamos 500 pontos aleatórios sobre a área e classificamos cada um como “não detectável” se localizado sobre árvores ou arbustos e “detectável” caso situado em local sem cobertura de dossel (Barasona *et al.* 2014). Excluímos os pontos localizados sobre áreas não acessíveis (eg corpos d’água ou edificações). O coeficiente de disponibilidade calculado foi de 0,85. Os modelos de captura-recaptura para os observadores múltiplos foram construídos no programa MARK (White and Burnham 1999), através do pacote RMark (Laake 2013) no ambiente R. Montamos os modelos de população fechada permitindo uma probabilidade de detecção para cada observador.

RESULTADOS

Os resultados encontrados estão resumidos na Tabela 1. Todas as estimativas apresentaram um viés negativo com relação ao tamanho populacional real (414 indivíduos). A probabilidade de um indivíduo visível ser detectado por pelo menos um observador foi de 0,999 para todos os voos (Tabela Suplementar 1). Como consequência, as estimativas de abundância dos modelos de captura-recaptura para observadores múltiplos indicaram que todos os indivíduos visíveis foram detectados pelos observadores (contagem máxima = obs. múltiplos). Os valores inferiores e superiores do intervalo de confiança de 95% dessas estimativas foram iguais à média (SE < 0,001). Das 1856 detecções dos observadores nos três sobrevoos, apenas duas foram identificadas como falsos positivos por erro de identificação (erro 4).

Tabela 1. Comparação de diferentes métodos de estimativa de abundância aplicados à três levantamentos aéreos com *drone*, sobre uma população de 414 indivíduos de gado.

Voo (hora)	Contagem bruta			Contagem máxima	Obs. múltiplos		Contagem corrigida
	Obs. 1	Obs. 2	Obs. 3		\hat{N}	Detecção (IC95%)	
1 (11h)	173	171	175	181	181	0,96 (0,93-0,97)	208
2 (13h)	265	279	276	283	283	0,96 (0,95-0,97)	325
3 (15h)	256	265	261	276	276	0,94 (0,93-0,96)	317

Notas: Os valores dos intervalos de confiança de 95% de \hat{N} são iguais à média. A detecção é a probabilidade média para os três observadores.

DISCUSSÃO

Nós aplicamos métodos presentes na literatura para acessar a abundância em levantamentos aéreos com *drones* e comparamos as estimativas com o tamanho populacional real conhecido. Todos os métodos apresentaram grandes vieses nas estimativas e subestimaram o tamanho populacional. O viés negativo encontrado indica que os falsos negativos são os erros que mais influenciaram as estimativas.

Uma vez que a probabilidade de detecção pelos observadores foi muito alta, o viés negativo observado em todas as estimativas avaliadas teve como principal fonte de erro os indivíduos escondidos sob árvores (ie indisponíveis). O viés da disponibilidade é bem conhecido nos levantamentos aéreos convencionais, tanto para animais terrestres (escondidos sob vegetação) quanto para aquáticos (mergulhados). Para lidar com esse viés, as contagens feitas com aviões tripulados são corrigidas com base em informações adicionais providas por outros métodos de coleta (eg contagens no solo, subpopulação de indivíduos marcados ou dados de telemetria; Pollock *et al.* 2006; Edwards *et al.* 2007). Contudo, são raros os casos onde há disponibilidade desse tipo de dado e isso limita muito a usabilidade de *drones*.

O coeficiente de disponibilidade aplicado para ajustar o viés da disponibilidade foi insuficiente para corrigir acuradamente as contagens. Uma das grandes limitações desse método é assumir que os indivíduos estão aleatoriamente distribuídos no espaço e que a disponibilidade é constante no tempo. Neste estudo, se desconsiderarmos os falsos positivos, a disponibilidade variou de 44% (voo1) a 68% em um intervalo de poucas horas. A disponibilidade dos indivíduos para serem detectados pode variar em consequência de diversos fatores como, por exemplo, a

espécie estudada, seu horário de atividade, idade dos indivíduos e a cobertura vegetal da área amostrada. Além disso, se essa correção for aplicada a espécies que evitam áreas sombreadas, as estimativas podem resultar enviesadas positivamente. Possivelmente, em alguns casos específicos onde *drones* vem sendo usados, a indisponibilidade apresenta um efeito insignificante nas contagens, como na amostragem de colônias de aves aquáticas (Sardà-Palomera *et al.* 2012; Chabot *et al.* 2015).

O uso de observadores múltiplos é útil e prático para corrigir a falha de visualização na revisão das imagens e assim estimar o número de indivíduos disponíveis para detecção. Embora amplamente utilizado em levantamentos aéreos convencionais (Williams *et al.* 2002), poucos estudos aplicaram esse método em amostragem com *drones* (Vermeulen *et al.* 2013). Apesar de neste estudo havermos encontrado uma alta probabilidade de detecção pelos observadores, a visualização deve ser mais baixa em situações de menor resolução ou de espécies mais crípticas. O longo tempo de revisão das imagens (neste estudo, média de 3 h para uma área de 470 ha) pode gerar falhas na detecção de indivíduos. A medida que os bancos de dados forem aumentando e o tempo de revisão se tornar impraticável, métodos semiautomatizados de busca serão imprescindíveis (Conn *et al.* 2013). Além disso, o uso de sensores térmicos nos levantamentos com *drones* pode trazer vantagens para aperfeiçoar a detecção dos indivíduos (Chrétien *et al.* 2016).

Os falsos positivos não apresentaram grande influência nas contagens. Utilizar a imagem fundida da área permitiu que excluíssemos as contagens duplas nas áreas de sobreposição entre fotos (veja Hodgson *et al.* 2013 para uma correção desses falsos positivos). A instrução aos observadores de evitar marcar indivíduos em caso de dúvida, praticamente anulou a ocorrência de erros de identificação. Por outro lado, nós não acessamos diretamente as contagens duplas causadas pelo movimento entre as linhas de voo. Entretanto, supomos que, caso esse fato ocorresse frequentemente, as contagens não apresentariam um viés negativo tão grande. A influência dessa fonte de erro deve variar conforme a mobilidade da espécie e características do plano de voo (eg altura, sobreposição ou distância entre linhas) e haverá casos onde ela terá maior efeito.

Os *drones* oferecem uma alternativa promissora para amostragens de populações animais. Porém, antes do uso indiscriminado, os dados coletados pelos *drones* devem ser avaliados com

cautela. Avaliações e medidas de conservação exigem informações acuradas de abundância para acessar o status e detectar tendências populacionais. Além disso, a fim de entender como fatores bióticos e abióticos afetam a abundância no espaço e no tempo, devemos ter a habilidade de levar em conta o efeito que os erros de amostragem geram nas estimativas. O desenvolvimento de novas abordagens estatísticas vem sendo requisitada para estimativas confiáveis em monitoramentos de populações com *drones* (Linchant *et al.* 2015). A facilidade que os *drones* possuem para replicar voos, devido a sua versatilidade, permite o uso de métodos estatísticos mais robustos para acomodar as fontes de erro de medição, como por exemplo a modelagem hierárquica (Royle and Dorazio 2008; Kéry and Royle 2016).

REFERÊNCIAS

- Anderson K and Gaston KJ. 2013. Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Front Ecol Environ* **11**: 138–46.
- Barasona JÁ, Mulero-Pázmány M, Acevedo P, *et al.* 2014. Unmanned Aircraft Systems for Studying Spatial Abundance of Ungulates: Relevance to Spatial Epidemiology. *PLoS One* **9**: e115608.
- Chabot D, Craik SR, and Bird DM. 2015. Population Census of a Large Common Tern Colony with a Small Unmanned Aircraft. *PLoS One* **10**: e0122588.
- Chrétien L-P, Théau J, and Ménard P. 2016. Visible and thermal infrared remote sensing for the detection of white-tailed deer using an unmanned aerial system. *Wildl Soc Bull* **40**: 181–91.
- Christie KS, Gilbert SL, Brown CL, *et al.* 2016. Unmanned aircraft systems in wildlife research: current and future applications of a transformative technology. *Front Ecol Environ* **14**: 241–51.
- Conn PB, Hoef JM Ver, McClintock BT, *et al.* 2014. Estimating multispecies abundance using automated detection systems: ice-associated seals in the Bering Sea. *Methods Ecol Evol* **5**: 1280–93.
- Edwards HH, Pollock KH, Ackerman BB, *et al.* 2007. Estimation of Detection Probability in Manatee Aerial Surveys at a Winter Aggregation Site. *J Wildl Manage* **71**: 2052–60.
- Gimenez O, Buckland ST, Morgan BJT, *et al.* 2014. Statistical ecology comes of age. *Biol*

- Lett* **10**: 20140698.
- Hodgson JC, Baylis SM, Mott R, *et al.* 2016. Precision wildlife monitoring using unmanned aerial vehicles. *Sci Rep* **6**: 22574.
- Hodgson A, Kelly N, and Peel D. 2013. Unmanned Aerial Vehicles (UAVs) for Surveying Marine Fauna: A Dugong Case Study. *PLoS One* **8**: e79556.
- Kéry M and Royle JA. 2016. Applied hierarchical modeling in ecology, volume 1 (Prelude and Static model). New York: Academic Press.
- Laake JL. 2013. RMark: An R Interface for Analysis of Capture-Recapture Data with MARK. AFSC Processed Rep. 2013-01. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv. Seattle, WA.
- Linchant J, Lisein J, Semeki J, *et al.* 2015. Are unmanned aircraft systems (UASs) the future of wildlife monitoring? A review of accomplishments and challenges. *Mamm Rev* **45**: 239–52.
- Mackenzie DI, Nichols JD, Royle JA, *et al.* 2006. Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence. San Diego: Elsevier.
- Martin J, Edwards HH, Burgess MA, *et al.* 2012. Estimating Distribution of Hidden Objects with Drones: From Tennis Balls to Manatees. *PLoS One* **7**: e38882.
- Mulero-Pázmány M, Barasona JÁ, Acevedo P, *et al.* 2015. Unmanned Aircraft Systems complement biologging in spatial ecology studies. *Ecol Evol* **5**: 4808–18.
- Pollock KH, Marsh HD, Lawler IR, and W AM. 2006. Estimating Animal Abundance in Heterogeneous Environments: An Application to Aerial Surveys for Dugongs. *J Wildl Manage* **70**: 255–62.
- Royle JA and Dorazio RM. 2008. Hierarchical Modeling and Inference in Ecology. New York: Academic Press.
- Sardà-Palomera F, Bota G, Viñolo C, *et al.* 2012. Fine-scale bird monitoring from light unmanned aircraft systems. *IBIS Int J Avian Sci* **154**: 177–83.
- Vermeulen C, Lejeune P, Lisein J, *et al.* 2013. Unmanned Aerial Survey of Elephants. *PLoS One* **8**: e54700.
- White GC and Burnham KP. 1999. Program MARK: survival estimation from populations of marked animals. *Bird Study* **46**: S120–39.

Williams BK, Nichols JD, and Conroy MJ. 2002. *Analysis and Management of Animals Populations*. San Diego: Academic Press.

MATERIAL SUPLEMENTAR



Figura 1. Procedimento de captura-recaptura com três observadores independentes em imagem gerada por um sobrevoo com *drone*. A procura foi realizada no ArcGIS (ESRI). Em uma camada específica para cada observador (representados por cores diferentes), um ponto foi posicionado no centro dos indivíduos visualizados. A partir da sobreposição de buffers de 30cm ao redor dos pontos, um histórico de captura (1 = detectado; 0 = não detectado) foi gerado para cada indivíduo. Imagem coletada pelo sistema aéreo não tripulado Echar 20B a 300 m acima do solo, com resolução de 4,6 cm/px.

Tabela 1. Estimativa do número de indivíduos disponíveis para detecção para os três sobrevoos de *drone*. Resultados dos modelos de captura-recaptura de população fechada para observadores múltiplos.

Voo (hora)	Frequências de históricos de captura							\hat{N} (IC95%)	Prob. de detecção (IC95%)		
	<i>111</i>	<i>110</i>	<i>101</i>	<i>100</i>	<i>011</i>	<i>010</i>	<i>001</i>		<i>Obs1</i>	<i>Obs2</i>	<i>Obs3</i>
1 (11h)	165	2	3	3	3	1	4	181 (181-181)	0,96 (0,91-0,98)	0,94 (0,90-0,97)	0,97 (0,93-0,98)
2 (13h)	259	4	1	1	10	4	4	283 (283-283)	0,94 (0,90-0,96)	0,98 (0,95-0,99)	0,97 (0,94-0,98)
3 (15h)	240	9	5	2	12	4	4	276 (276-276)	0,93 (0,89-0,95)	0,96 (0,93-0,98)	0,95 (0,91-0,97)

CAPÍTULO 2

Uma abordagem de modelagem hierárquica para estimar abundância com drones

*Artigo a ser submetido no periódico
Methods in Ecology and Evolution.*

UMA ABORDAGEM DE MODELAGEM HIERÁRQUICA PARA ESTIMAR ABUNDÂNCIA COM DRONES

Ismael V Brack¹, Andreas Kindel², Luiz FB Oliveira³

¹ PPG Ecologia UFRGS; ² Dep. Ecologia UFRGS; ³ Dep. Vertebrados MN/UFRJ

RESUMO

1. Os *drones* vêm prometendo revolucionar a obtenção de informações sobre distribuição e abundância de populações animais. A maioria das estimativas de abundância feitas com *drones* utilizaram contagens brutas, desconsiderando as principais fontes de erro em levantamentos aéreos: vieses de disponibilidade e visualização. Esperançosamente, a facilidade de replicação de amostragem com *drones* permite o desenvolvimento de novas abordagens estatísticas para acomodar os erros de medição nas estimativas de abundância.
2. Nós propomos o uso de um *multinomial N-mixture model* com emigração temporária para estimar abundância em amostragens temporal e espacialmente replicadas com *drones* e observadores múltiplos na revisão das imagens. Nós avaliamos o desempenho do modelo em diversos cenários biológicos e amostrais simulados e provemos algumas recomendações na amostragem para aperfeiçoar as estimativas.
3. O modelo proposto apresentou estimativas acuradas e bons níveis de precisão na maioria das situações simuladas, exceto em cenários de baixa disponibilidade de indivíduos. Além disso, o uso de mais de um observador na revisão das imagens permitiu particionar o processo de observação e aperfeiçoar as estimativas de abundância.
4. O conhecimento acerca do comportamento da espécie estudada pode ser utilizado para planejar a amostragem e maximizar a probabilidade de disponibilidade dos indivíduos. Do mesmo modo, estudos simulados como este podem ser úteis no delineamento de amostragem mais eficientes.
5. A modelagem hierárquica se mostrou uma abordagem promissora para levar em conta os diferentes erros de medição em estudos populacionais com *drones*, possibilitando o estabelecimento dessa ferramenta para as diversas aplicações em ecologia e conservação.

Palavras-chave: veículo aéreo não-tripulado, *N-mixture model*, levantamento aéreo, viés de disponibilidade, desenho amostral, estudo de simulação

INTRODUÇÃO

Drones, ou veículos aéreos não-tripulados (VANT), têm emergido como uma ferramenta prática e acessível na ecologia, prometendo revolucionar a obtenção de informações sobre distribuição e abundância dos organismos (Anderson & Gaston 2013). As potencialidades do uso dessa ferramenta para investigar e monitorar populações já foram sugeridas para diversas espécies (Linchant *et al.* 2015; Christie *et al.* 2016; Hodgson *et al.* 2016). Apesar de alguns estudos terem abordado diretamente a capacidade dos *drones* para acessar a abundância de animais, o desenvolvimento de técnicas de amostragem e estatísticas específicas permanece praticamente inexplorado (Linchant *et al.* 2015).

O uso dos *drones* em estudos populacionais tem surgido principalmente como uma alternativa aos levantamentos aéreos de grandes mamíferos de áreas abertas (e.g., elefantes, Vermeulen *et al.* 2013; cervídeos, Chrétien *et al.* 2016), mamíferos marinhos (e.g., dugongos, Hodgson *et al.* 2013) e aves aquáticas (e.g., andorinha-do-mar, Chabot *et al.* 2015; flamingos, patos e maçaricos, Vas *et al.* 2015). Os *drones* apresentam algumas vantagens com relação aos levantamentos convencionais com aviões tripulados, entre elas, sobrevoos sem risco para os pesquisadores e piloto, uma logística mais simples de decolagem e pouso e um custo mais baixo por voo (Anderson & Gaston 2013). As principais desvantagens são uma baixa autonomia quando comparado aos aviões tripulados (Anderson & Gaston 2013) e as duras restrições legais e burocráticas que limitam sua utilização (Vincent *et al.* 2015).

As amostragens aéreas, tanto com aviões tripulados como não-tripulados (*drones*), apresentam potenciais fontes de erro que, caso não sejam consideradas, podem acarretar em grandes vieses nas estimativas de abundância (Capítulo 1). Abordagens que levem em conta essas fontes de erro no processo de observação (i.e. detecção imperfeita) são necessárias para a obtenção de estimativas de abundância acuradas e confiáveis (Mackenzie *et al.* 2006; Royle & Dorazio 2008). O processo de detecção em levantamentos aéreos é comumente dividido em dois componentes: i) a disponibilidade dos indivíduos para serem detectados (i.e. indivíduos não escondidos) e ii) sua visualização, dado que estão disponíveis (Pollock *et al.* 2006). O viés

da visualização é facilmente corrigido em levantamentos aéreos através do uso de observadores múltiplos (Williams *et al.* 2002) e/ou amostragem com distâncias (Buckland *et al.* 2001). Por outro lado, a disponibilidade é pouco acessada em levantamentos aéreos convencionais e, quando acessada, utiliza informações adicionais, as quais raramente são factíveis de obtenção (e.g. biotelemetria, contagens no solo, Edwards *et al.* 2007).

Estimativas de abundância com *drones* têm sido acessadas principalmente através de contagens brutas (Linchant *et al.* 2015). Em alguns casos – como em amostragens de colônias ou bandos de aves aquáticas (Sardà-Palomera *et al.* 2012; Chabot *et al.* 2015) – pode haver pouco ou nenhum viés de visualização e/ou disponibilidade nessas estimativas. Porém, devido à diversidade de potenciais aplicações dos *drones*, desconsiderar os indivíduos escondidos (e.g. debaixo de árvores ou mergulhados) e o erro do observador, na maioria dos casos, pode induzir a vieses significativos (Joseph *et al.* 2009; Dénes *et al.* 2015). Alguns estudos demonstraram certa preocupação em levar em conta algumas fontes de erro. Vermeulen *et al.* 2013 utilizaram mais de um revisor nas imagens para acessar o erro de visualização, enquanto Barasona *et al.* 2014 realizaram correções arbitrárias nas contagens para lidar com o viés de disponibilidade.

Uma grande vantagem que o uso de *drones* oferece para análises estatísticas mais robustas é a capacidade de realizar sobrevoos sequenciais em um curto espaço de tempo (Anderson & Gaston 2013). Essa facilidade de replicação permite a utilização de abordagens de modelagem hierárquica, capazes de lidar com diversas fontes de erro de observação (Royle & Dorazio 2008; Kéry & Royle 2016). Assim, é possível realizar inferências robustas sobre as variáveis de interesse que não são diretamente observadas (i.e. latentes; e.g. ocupação, abundância, sobrevivência). Martin *et al.* (2012) foram, até onde sabemos, os únicos que propuseram uma abordagem que modela explicitamente os processos de erro na detecção em amostragem com *drones*. Porém, sua abordagem é bastante limitada por assumir pressupostos improváveis, como por exemplo, de não-movimentação dos animais entre unidades amostrais pequenas e adjacentes.

Neste estudo, propomos a utilização de um modelo hierárquico para estimar abundância de animais através de amostragens aéreas com *drones*. Para isso, avaliamos em cenários simulados o desempenho de um *multinomial N-mixture model* que leva em conta os vieses de disponibilidade e visualização. O modelo proposto requer uma amostragem espacial e

temporalmente replicada com múltiplos observadores. O estudo simulado está dividido em três partes. Primeiro, avaliamos a acurácia e a precisão do modelo variando o número de observadores e o número de replicações temporais (i.e. sobrevoos repetidos). A seguir, verificamos o número de unidades espaciais necessárias para atingir uma precisão desejada, em diferentes cenários de abundância e disponibilidade de indivíduos. Por último, consideramos uma situação de limitação de número de voos (200 voos) e verificamos qual a melhor estratégia de amostragem para alocar os voos em replicações espaciais ou temporais. A partir dos resultados encontrados aqui, provemos algumas recomendações gerais para os desenhos amostrais a fim de aperfeiçoar as estimativas de abundância geradas com *drones*.

MÉTODOS

Modelo

O modelo proposto é um *multinomial N-mixture model* para população aberta, onde a dinâmica temporal é determinada por um modelo de emigração temporária (Chandler *et al.* 2011). Esse modelo assume uma amostragem com “desenho meta-populacional” (ver Kéry & Royle 2016), ou seja, replicada espacialmente em $i = 1, 2, \dots, R$ unidades amostrais. Considera uma amostragem replicada temporalmente em $t = 1, 2, \dots, T$ ocasiões para cada unidade amostral. Neste caso, cada ocasião t corresponde a um sobrevoo nas unidades espaciais. Para cada sobrevoo t em cada unidade i , observadores independentes revisam as imagens geradas pelo drone e marcam cada indivíduo visualizado (assim como em um procedimento de captura-recaptura), gerando um histórico de capturas.

Assumimos que a abundância nas unidades amostrais (processo biológico) é determinada por uma distribuição de probabilidade discreta (Poisson ou Binomial negativa) com média λ :

$$M_i \sim \text{Poisson}(\lambda)$$

onde M_i é a abundância local em cada sítio, denominada superpopulação. Em cada sobrevoo t , dos M_i indivíduos da superpopulação, N_{it} indivíduos estão disponíveis para serem detectados com uma probabilidade ϕ (emigração temporária = $1 - \phi$):

$$N_{it} \sim \text{Binomial}(M_{it}; \phi).$$

Os N_{it} indivíduos disponíveis têm uma probabilidade p de serem visualizados por cada observador. Consideramos um modelo de observação multinomial onde os históricos de captura

são agrupados em frequências de históricos de captura (e.g., para três observadores: “111”, “110”, “101”, “100”, “011”, “010” e “001” históricos observáveis). Assim, cada possibilidade de histórico possui uma probabilidade celular π , condicionada à probabilidade de visualização p :

$$\mathbf{y}_{it} \sim \text{Multinomial}(N_{it}; \boldsymbol{\pi}(p))$$

onde \mathbf{y}_{it} é um vetor contendo o número de indivíduos visualizados para cada histórico observável. Através da inclusão de covariáveis nos parâmetros (λ , ϕ e p), essa abordagem nos permite aperfeiçoar as inferências sobre os processos envolvidos na variação espacial da abundância, ao mesmo tempo em que modelamos a heterogeneidade no processo de observação. Para mais detalhes sobre o modelo ver Chandler *et al.* (2011) e a função *gmultmix* do pacote *unmarked* (Fiske & Chandler 2011).

Simulações

Conduzimos o estudo de simulação em três partes. Primeiro, avaliamos a acurácia e a precisão do modelo para estimar a abundância média (λ), variando o número de observadores (K) e o número de replicações temporais (nº de voos; T). Acessamos a acurácia através do viés relativo (i.e. diferença da estimativa para o valor real, dividido pelo valor real) e da taxa de cobertura (i.e. proporção das simulações nas quais o intervalo de credibilidade de 95% contém o valor real usado para simular os dados). A precisão foi acessada através do coeficiente de variação (i.e. desvio padrão estimado dividido pela média estimada). Consideramos $T = 3, 5, 10$ e $K = 1, 2, 3$ – totalizando nove cenários – e fixamos $R = 40$ unidades espaciais. Note que quando $K = 1$ não é possível separar o processo de observação em dois níveis, resultando em um simples *Binomial N-mixture model* onde a probabilidade de detecção é o produto de ϕ e p . Cada cenário foi simulado 999 vezes e para cada simulação os dados foram gerados a partir de uma abundância média por sítio $\lambda = 4$ (distribuição Poisson), uma disponibilidade $\phi = 0,6$ e uma probabilidade de visualização $p = 0,7$. Utilizamos uma abordagem bayesiana para estimar os parâmetros. Rodamos três cadeias Monte Carlo Markov Chain (MCMC) com 30.000 iterações, descartando as 10.000 primeiras (*burn-in*) e mantendo todas as amostras (*thinned*), restando 60.000 amostras da distribuição *a posteriori*. Nós excluimos das análises aquelas

simulações que não atingiram convergência segundo o diagnóstico de Brooks-Gelman-Rubin (Brooks & Gelman 1998), ou seja, $R\text{-hat} < 1,1$.

Na segunda parte, simulamos diferentes cenários biológicos de abundância local e disponibilidade dos indivíduos para avaliar as replicações espaciais necessárias para atingir determinada precisão. Para isso, verificamos o número de unidades espaciais que precisaríamos para atingir um coeficiente de variação de 20% no parâmetro da abundância local (λ). Consideramos todos os cenários combinados entre $\lambda = (0,5; 2; 5; 10)$ e $\phi = (0,2; 0,4; 0,6; 0,8)$, totalizando 16 cenários, e fixamos $T = 3$, $K = 3$ e $p = 0,7$. Cada cenário começou com $R = 20$ e se, ao final de 100 simulações, 90% delas apresentassem um coeficiente de variação maior que 20% na estimativa de λ , era adicionada mais uma unidade espacial. Esse procedimento foi repetido até obtermos 90% das simulações com $CV \leq 20\%$. Em cada simulação rodamos três cadeias MCMC com 50.000 iterações, descartando as 10.000 primeiras (*burn-in*) e mantendo todas as amostras (*thinned*), restando 120.000 amostras da distribuição *a posteriori*.

Finalmente, avaliamos se é melhor priorizar replicações espaciais ou temporais com o objetivo de atingir estimativas mais acuradas de abundância, em diferentes cenários biológicos de abundância local e disponibilidade. Para isso, limitamos a amostragem a 200 voos e consideramos três estratégias: s1) $R = 67$, $T = 3$; s2) $R = 50$, $T = 4$; e s3) $R = 40$, $T = 5$. Analisamos os mesmos 16 cenários combinados entre $\lambda = (0,5; 2; 5; 10)$ e $\phi = (0,2; 0,4; 0,6; 0,8)$ e fixamos $K = 3$ e $p = 0,7$. Para cada estratégia em cada cenário, realizamos 300 simulações e verificamos qual das estratégias atingiu o menor coeficiente de variação. Em cada simulação, utilizamos três cadeias MCMC com 50.000 iterações, descartando as 10.000 primeiras (*burn-in*) e mantendo todas as amostras (*thinned*), restando 120.000 amostras da distribuição *a posteriori*. Excluímos das análises aquelas simulações que não atingiram convergência $R\text{-hat} < 1,1$.

Nós simulamos os dados no ambiente R (versão 3.1.1) e ajustamos os modelos utilizando JAGS (versão 3.4.0, Plummer 2003) via pacote jagsUI (Kellner 2015). Para evitar o problema com o tamanho amostral aleatório na distribuição multinomial no JAGS, construímos os modelos condicionando o modelo de observação multinomial aos indivíduos observados. Para isso, inserimos um nível para os indivíduos observados $n_{it} \sim \text{Binomial}(N_{it}; 1 - \pi_0)$, onde $1 - \pi_0$ é a probabilidade do indivíduo ser visualizado pelo menos uma vez (ver Apêndice S1). Em

todos os cenários, consideramos os parâmetros constantes (i.e. sem efeitos de covariáveis) e utilizamos distribuições *a priori* “vagas” para todos os parâmetros (ver detalhes no Apêndice S1).

RESULTADOS

Quando o modelo foi ajustado em diferentes situações de número de observadores e de replicações temporais, a abundância local foi acuradamente estimada em todos os cenários e para os dois critérios empregados (viés relativo, Fig. 1; cobertura do IC95%, Tabela 1 do Apêndice S2). A precisão das estimativas foi mais baixa quando utilizado apenas um observador e com apenas três replicações temporais (Figura 2). Não obtivemos diferenças na precisão entre os cenários com dois e três observadores.

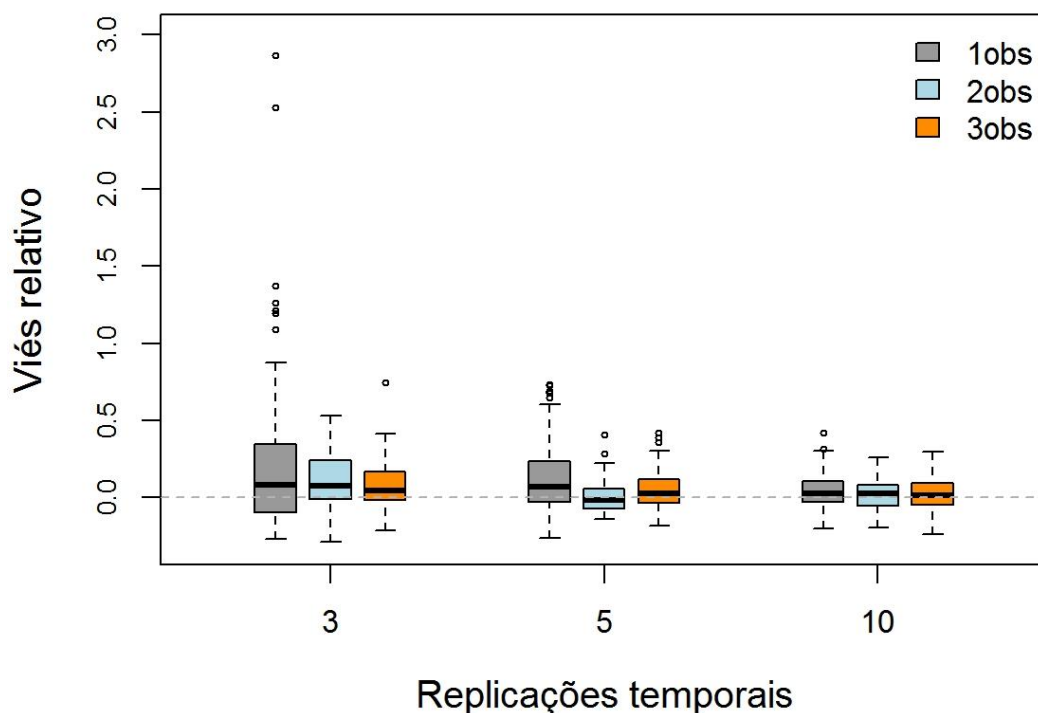


Figura 3. Viés relativo das estimativas de abundância local ($\lambda=4$) de um *multinomial N-mixture model* com emigração temporária, em cenários simulados de número de observadores e de réplicas temporais. Os boxplots indicam a distribuição de 999 vieses relativos, calculados a partir da diferença entre as médias posteriores e o valor real utilizado para gerar o conjunto de dados. Os limites inferior e superior da caixa indicam, respectivamente, os quartis de 25% e

75%, a linha preta no centro da caixa representa a mediana, os limites das hastes estendem as estimativas a 1,5 vezes os valores dos interquartis e os círculos são os valores extremos.

Quando avaliamos o número de unidades espaciais necessárias para atingir uma precisão-alvo, não foi possível alcançar a precisão desejada nos cenários de mais baixa disponibilidade de indivíduos (exceto com alta abundância local), assim como no cenário $\lambda=0,5$ e $\varphi=0,4$ (Tabela 2). Os cenários com disponibilidade mais alta ($\varphi \geq 0,6$) apresentaram números factíveis de unidades espaciais, para todas as condições de abundância local.

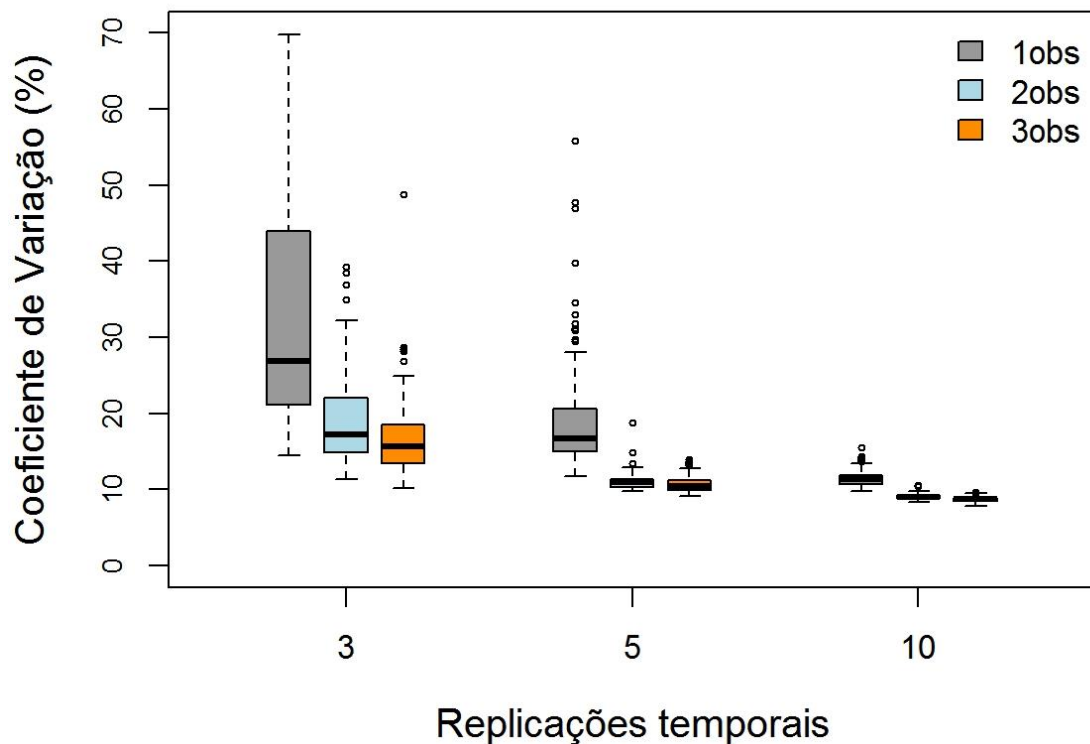


Figura 4. Coeficiente de variação das estimativas de abundância local ($\lambda=4$) de um *multinomial N-mixture model* com emigração temporária, em cenários simulados de número de observadores e de réplicas temporais. Os boxplots indicam a distribuição de 999 coeficientes de variação, calculados a partir do quociente entre o desvio padrão e a média posterior. A simbologia é a mesma da figura 1.

Quando investigamos a melhor estratégia de amostragem nos diferentes cenários, não encontramos grandes diferenças na precisão das estimativas de abundância entre as estratégias

de amostragem em cada cenário (Fig. 1 do Apêndice S2). Ainda assim, as estratégias com mais replicações temporais apresentaram uma precisão ligeiramente maior em cenários de baixa disponibilidade e abundância local. A medida que aumentamos a disponibilidade e mantivemos a abundância local baixa, a estratégia que prioriza replicações espaciais obteve estimativas mais precisas. Nenhuma das simulações atingiu a convergência no cenário com menor disponibilidade e abundância ($\lambda=0,5$; $\varphi=0,2$). De maneira geral, o coeficiente de variação foi enorme para os cenários de disponibilidade muito baixa ($\varphi=0,2$) e diminuiu bastante com o aumento da disponibilidade. A precisão aumentou levemente com o aumento da abundância local.

Tabela 1. Número de replicações espaciais (sítios amostrais) necessárias para atingir um coeficiente de variação menor que 20% em diferentes condições de abundância local (λ) e disponibilidade (φ).

		Abundância local (λ)			
		0,5	2	5	10
Disponibilidade (φ)	0,2	> 200	> 200	> 200	200
	0,4	> 200	123	118	113
	0,6	84	44	41	36
	0,8	63	20	20	20

DISCUSSÃO

Nós acessamos o desempenho de um *multinomial N-mixture model* com emigração temporária para estimar abundância com *drones* e levar em conta ambos os erros de medição falsos negativos que ocorrem em levantamentos aéreos (disponibilidade e visualização). O modelo proposto apresentou estimativas acuradas e precisas para a maioria dos cenários de amostragem simulados. Os resultados indicam que um protocolo de amostragem com *drones*, replicado espacial e temporalmente e com revisão das imagens por múltiplos observadores, pode prover informações confiáveis de abundância para suas diversas aplicações em ecologia e conservação. Os *drones* podem se tornar uma alternativa mais eficiente aos levantamentos

aéreos convencionais, podendo inclusive ser utilizados em situações onde os aviões tripulados não são comumente empregados. Devido sua versatilidade, a replicação dos voos com *drones* torna-se mais fácil e assim estimativas de abundância poderiam ser acessadas em condições de disponibilidade intermediária.

Entretanto, o desempenho do modelo em cenários simulados de baixa disponibilidade de indivíduos foi insatisfatório, principalmente com baixa abundância local. Nessas condições, as estimativas apresentaram enorme imprecisão. Como consequência, a utilização de *drones* para estimar abundância em ambientes com alta proporção de cobertura arbórea ou para espécies de hábito florestal deve ser impraticável. Outra situação em que a disponibilidade pode ser baixa é para espécies marinhas que passam pouco tempo na superfície d'água. Uma maneira de tentar contornar esse problema é planejar a amostragem para maximizar a disponibilidade dos indivíduos. Assim, deveríamos priorizar, por exemplo, horários em que os indivíduos se encontrem em áreas abertas ou na superfície d'água. Além disso, a inclusão de covariáveis de sítio (e.g., cobertura de dossel) e ocasião (e.g., temperatura, hora do dia) na disponibilidade deve ajudar a modelar a heterogeneidade na variação desse parâmetro.

O uso de mais de um observador nos permitiu particionar o processo de observação em dois níveis e aumentar consideravelmente a precisão nas estimativas de abundância. Além disso, separando o processo de observação é possível acomodar, através da inclusão de covariáveis, fontes de variação que podem ter influências diferentes em cada nível. Portanto, o emprego de observadores múltiplos permite aperfeiçoar as estimativas de abundância, ao mesmo tempo em que é pouco custoso. Por outro lado, a medida que os bancos de dados gerados pelos voos com *drones* forem sendo ampliados, o tempo de revisão das imagens pode se tornar um fator limitante. Com o desenvolvimento de métodos semiautomatizados de procura (Chrétien *et al.* 2016), o nível de visualização do modelo pode ser modificado para acomodar esse novo processo de observação (Conn *et al.* 2013).

Dado que a baixa autonomia é uma das principais limitações dos *drones* (Christie *et al.* 2016), definir o melhor desenho amostral deve ser importante para obter as estimativas de modo eficiente, sem desperdiçar tempo e recursos. Apesar de sua baixa autonomia, a versatilidade dos *drones* permite a realização de sobrevoos consecutivos, na mesma área ou não. Além das medidas para maximizar a disponibilidade dos indivíduos, também é importante determinar se

a melhor estratégia deve ser repetir mais sobrevoos sobre um menor número de unidades espaciais ou o oposto. Assim, estudos simulados como este permitem planejar previamente a amostragem e definir como concentrar os esforços (Mackenzie & Royle 2005; Bailey *et al.* 2007; Conn *et al.* 2016). Embora não tenhamos encontrado muita diferença na precisão entre as estratégias consideradas, para garantir a estimabilidade dos parâmetros em cenários de menor densidade e disponibilidade, desenhos amostrais com maior número de replicações temporais devem ser priorizados.

No contexto de amostragem com *drones* aqui apresentado, assumimos que as unidades espaciais são geograficamente fechadas durante o estudo, ou seja, o número de indivíduos permanece constante ao longo de todos os sobrevoos. Nesse caso, a disponibilidade se refere a indivíduos não escondidos presentes na unidade espacial (“*random*” *temporary emigration*). Porém, quando aplicado a dados reais de indivíduos móveis, o pressuposto de fechamento é frequentemente violado (Chandler *et al.* 2011). Desse modo, a emigração temporária também pode ser causada por indivíduos que possuem suas áreas de vida nas bordas da unidade espacial e não estão presentes no momento da ocasião amostral (“*spatial*” *temporary emigration*). Nesses casos, a disponibilidade poderia levar em conta, de maneira indistinguível, ambos processos de emigração temporária e assim relaxar o pressuposto de unidades fechadas (Kéry & Royle 2016).

O modelo aqui proposto acomoda as duas principais fontes de erro falso negativo presentes em levantamentos aéreos (Pollock *et al.* 2006). Entretanto, o modelo assume que erros do tipo falso positivo não ocorrem. Erros de identificação de espécies são fáceis de evitar quando há dúvida na revisão das imagens (Capítulo 1). Por outro lado, os planos de voo comumente empregados nas amostragens com *drones* podem gerar falsos positivos por contagens duplas de um mesmo indivíduo (Chabot & Bird 2012; Hodgson *et al.* 2013). Abordagens hierárquicas para acomodar esse tipo de erro ainda são incipientes, sobretudo para dados de contagem (Dénes *et al.* 2015). Portanto, pesquisas futuras nessa área serão úteis para modelar falsos positivos em amostragens com *drones*.

Com a rápida disseminação dos *drones* em amostragens de populações silvestres, novas abordagens estatísticas eram requisitadas para acessar tamanhos populacionais com acurácia (Linchant *et al.* 2015). Nós propusemos aqui o uso de um *multinomial N-mixture model* com

emigração temporária para estimar abundância em amostragens espacial e temporalmente replicadas com *drones*. Dessa maneira, acomodamos as fontes de erro de medição e evitamos vieses sistemáticos nas estimativas de abundância. Além disso, a estrutura extremamente flexível dos modelos hierárquicos permite estender este modelo à diversas situações como, por exemplo, acomodar falsos positivos (Guillera-Arroita *et al.* 2012) ou modelar dinâmica populacional em monitoramentos a longo prazo (Kéry *et al.* 2009; Dail & Madsen 2011). O modelo mostrou um desempenho acurado e preciso para diversas situações simuladas, onde os pressupostos foram perfeitamente contemplados. O próximo passo é colocar o modelo à prova com dados empíricos, modelando a variação nos parâmetros através de covariáveis. Dessa forma, esperamos que a modelagem hierárquica se torne uma poderosa aliada para as estimativas de abundância em levantamentos aéreos com *drones*.

REFERÊNCIAS

- Anderson, K. & Gaston, K.J. (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment*, **11**, 138–146.
- Bailey, L.L., Hines, J.E., Nichols, J.D. & MacKenzie, D.I. (2007). Sampling design trade-offs in occupancy studies with imperfect detection: Examples and software. *Ecological Applications*, **17**, 281–290.
- Barasona, J.Á., Mulero-Pázmány, M., Acevedo, P., Negro, J.J., Torres, M.J., Gortázar, C. & Vicente, J. (2014). Unmanned Aircraft Systems for Studying Spatial Abundance of Ungulates: Relevance to Spatial Epidemiology. *PLoS ONE*, **9**, e115608.
- Brooks, S.P. & Gelman, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, **7**, 434–455.
- Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2001). *Introduction to Distance Sampling*. Oxford University Press, Oxford.
- Chabot, D. & Bird, D.M. (2012). Evaluation of an off-the-shelf Unmanned Aircraft System for Surveying Flocks of Geese. *Waterbirds*, **35**, 170–174.
- Chabot, D., Craik, S.R. & Bird, D.M. (2015). Population Census of a Large Common Tern Colony with a Small Unmanned Aircraft. *PLOS ONE*, **10**, e0122588.

- Chandler, R.B., Royle, J.A. & King, D.I. (2011). Inference about density and temporary emigration in unmarked populations. *Ecology*, **92**, 1429–1435.
- Chrétien, L.-P., Théau, J. & Ménard, P. (2016). Visible and thermal infrared remote sensing for the detection of white-tailed deer using an unmanned aerial system. *Wildlife Society Bulletin*, **40**, 181–191.
- Christie, K.S., Gilbert, S.L., Brown, C.L., Hatfield, M., Biology, A., Fairbanks, A., Koyokuk, N. & Hanson, L. (2016). Unmanned aircraft systems in wildlife research: current and future applications of a transformative technology. *Frontiers in Ecology and the Environment*, **14**, 241–251.
- Conn, P.B., Ver Hoef, J.M., McClintock, B.T., Moreland, E.E., London, J.M., Cameron, M.F., Dahle, S.P. & Boveng, P.L. (2014). Estimating multispecies abundance using automated detection systems: ice-associated seals in the Bering Sea. *Methods in Ecology and Evolution*, **5**, 1280–1293.
- Conn, P.B., Moreland, E.E., Regehr, E. V, Richmond, E.L., Cameron, M.F. & Boveng, P.L. (2016). Using simulation to evaluate wildlife survey designs: polar bears and seals in the Chukchi Sea. *Royal Society Open Science*, **3**, 150561.
- Dail, D. & Madsen, L. (2011). Models for Estimating Abundance from Repeated Counts of an Open Metapopulation. *Biometrics*, **67**, 577–587.
- Dénes, F. V., Silveira, L.F. & Beissinger, S.R. (2015). Estimating abundance of unmarked animal populations: accounting for imperfect detection and other sources of zero inflation. *Methods in Ecology and Evolution*, **6**, 543–556.
- Edwards, H.H., Pollock, K.H., Ackerman, B.B., Reynolds, J.E. & Powell, J.A. (2007). Estimation of Detection Probability in Manatee Aerial Surveys at a Winter Aggregation Site. *Journal of Wildlife Management*, **71**, 2052–2060.
- Fiske, I.J. & Chandler, R.B. (2011). unmarked: An R package for fitting hierarchical models of wildlife occurrence and abundance. *Journal of Statistical Software*, **43**, 1–23.
- Guillera-Arroita, G., Ridout, M.S., Morgan, B.J.T. & Linkie, M. (2012). Models for species-detection data collected along transects in the presence of abundance-induced heterogeneity and clustering in the detection process. *Methods in Ecology and Evolution*, **3**, 358–367.

- Hodgson, J.C., Baylis, S.M., Mott, R., Herrod, A. & Clarke, R.H. (2016). Precision wildlife monitoring using unmanned aerial vehicles. *Scientific Reports*, **6**, 22574.
- Hodgson, A., Kelly, N. & Peel, D. (2013). Unmanned Aerial Vehicles (UAVs) for Surveying Marine Fauna: A Dugong Case Study. *PLoS ONE*, **8**, e79556.
- Joseph, L.N., Elkin, C., Martin, T.G. & Possingham, H.P. (2009). Modeling abundance using N-mixture models: The importance of considering ecological mechanisms. *Ecological Applications*, **19**, 631–642.
- Kellner, K.F. (2015). jagsUI: A Wrapper Around rjags to Streamline JAGS Analyses. R package version 1.3.1.
- Kéry, M., Dorazio, R.M., Soldaat, L., Van Strien, A., Zuiderwijk, A. & Royle, J.A. (2009). Trend estimation in populations with imperfect detection. *Journal of Applied Ecology*, **46**, 1163–1172.
- Kéry, M. & Royle, J.A. (2016). *Applied hierarchical modeling in ecology, volume 1 (Prelude and Static model)*. Academic Press, New York.
- Linchant, J., Lisein, J., Semeki, J., Lejeune, P. & Vermeulen, C. (2015). Are unmanned aircraft systems (UASs) the future of wildlife monitoring? A review of accomplishments and challenges. *Mammal Review*, **45**, 239–252.
- Mackenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.L. & Hines, J.E. (2006). *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Elsevier, San Diego.
- Mackenzie, D.I. & Royle, J.A. (2005). Designing occupancy studies: General advice and allocating survey effort. *Journal of Applied Ecology*, **42**, 1105–1114.
- Martin, J., Edwards, H.H., Burgess, M.A., Percival, H.F., Fagan, D.E., Gardner, B.E., Ortega-Ortiz, J.G., Ifju, P.G., Evers, B.S. & Rambo, T.J. (2012). Estimating Distribution of Hidden Objects with Drones: From Tennis Balls to Manatees. *PLoS ONE*, **7**, e38882.
- Plummer, M. (2003). JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. Proceedings of the 3rd International Workshop in Distributed Statistical Computing, Vienna, Austria.
- Pollock, K.H., Marsh, H.D., Lawler, I.R. & W, A.M. (2006). Estimating Animal Abundance in Heterogeneous Environments: An Application to Aerial Surveys for Dugongs. *Journal of*

Wildlife Management, **70**, 255–262.

- Royle, J.A. & Dorazio, R.M. (2008). *Hierarchical Modeling and Inference in Ecology*. Academic Press, New York.
- Sardà-Palomera, F., Bota, G., Viñolo, C., Pallarés, O., Sazatornil, V., Brotons, L., Gomáriz, S. & Sardà, F. (2012). Fine-scale bird monitoring from light unmanned aircraft systems. *IBIS The International Journal of Avian Science*, **154**, 177–183.
- Vas, E., Lescroel, A., Duriez, O., Boguszewski, G. & Gremillet, D. (2015). Approaching birds with drones: first experiments and ethical guidelines. *Biology Letters*, **11**, 20140754–20140754.
- Vermeulen, C., Lejeune, P., Lisein, J., Sawadogo, P. & Bouché, P. (2013). Unmanned Aerial Survey of Elephants. *PLoS ONE*, **8**, e54700.
- Vincent, J.B., Werden, L.K. & Ditmer, M.A. (2015). Barriers to adding UAVs to the ecologist's toolbox. *Frontiers in Ecology and the Environment*, **13**, 74–75.
- Williams, B.K., Nichols, J.D. & Conroy, M.J. (2002). *Analysis and Management of Animals Populations*. Academic Press, San Diego.

MATERIAL SUPPLEMENTAR

Apêndice S1. Códigos R e JAGS para gerar e analisar dados simulados.

Apêndice S2. Resultados estendidos das simulações.

MATERIAL SUPLEMENTAR

Apêndice S1. Códigos R e JAGS para gerar e analisar dados simulados.

```
#####
#####
#####** Multinomial N-mixture model with temporary emigration **#####
#####** Model to estimate abundance in spatially replicated drone **#####
#####** surveys with independent observers, while discriminating two **#####
#####** observation processes: hidden individuals and sighting failure. **#####
#####
#####** I.V. Brack, 2016 **#####
#####
#####

# Package
library(jagsUI)

# Determine sample sizes (spatial and temporal replication and number of observers)
R <- 60 # spatial units
T <- 3 # occasions
K <- 3 # observers

# Sample abundance from a Poisson distribution
lambda <- 2 # local abundance

M <- rpois(n = R, lambda = lambda)

# Define availability and detection probabilities
phi <- 0.4
p <- 0.7

# Construct the multinomial cell probabilities (pi)
if(k==3){
  cellprobs <- c(
    p * p * p, # 111
    p * p * (1-p), # 110
    p * (1-p) * p, # 101
    p * (1-p) * (1-p), # 100
    (1-p) * p * p, # 011
    (1-p) * p * (1-p), # 010
    (1-p) * (1-p) * p, # 001
    (1-p) * (1-p) * (1-p) # 000
  )
}
if(k==2){
  cellprobs <- c(
    p * p, # 11
    p * (1-p), # 10
    (1-p) * p, # 01
    (1-p) * (1-p) # 00
  )
}
if(k!=3|k!=2){stop("k must be equal 2 or 3")}

# Create structure to contain individuals available
avails <- array(dim = c(R, T))

# Create structure to contain frequencies of capture histories
Y <- array(dim=c(R,2^K,T))
if(k==3){colnames(Y) <- c("111","110","101","100","011","010","001","000")}
if(k==2){colnames(Y) <- c("11","10","01","00")}
```

```

## Loop over sites and occasions to generate data
for (i in 1:length(M)){
  avails[i,] <- rbinom(n = T, size = M[i], prob = phi)
  for (j in 1:ncol(avails)){
    Y[i,,j] <- rmultinom(1, avails[i,j], cellprobs)
  }
}

Y <- Y[,-ncol(Y),] # excluding "000"/"00"

# Bundle data
ncap <- array(dim=c(R,T))
for (k in 1:T) {
  for (i in 1:nrow(Y[, ,k])) {
    ncap[i,k] <- sum(Y[i, ,k])
  }
}
nmax <- apply(ncap,1,max)

data1 <- list(Y = Y, ncap = ncap , R = R, T = T)

# Write BUGS model
if(k==3){
  sink("model_3obs.txt")
  cat("
  model {
    # Priors
    phi~dunif(0,1)
    p0~dunif(0,1)
    lambda~dunif(0,100)

    # Likelihood
    # Biological model for true abundance
    for (i in 1:R) {
      M[i] ~ dpois(lambda)
      # Observation model for replicated counts
      for (k in 1:T) {
        avails[i,k] ~ dbin(phi , M[i])
        # Define multinomial cell probabilities
        p[i,k] <- p0
        mu[i,1,k] <- p[i,k] * p[i,k] * p[i,k] # 111
        mu[i,2,k] <- p[i,k] * p[i,k] * (1-p[i,k]) # 110
        mu[i,3,k] <- p[i,k] * (1-p[i,k]) * p[i,k] # 101
        mu[i,4,k] <- p[i,k] * (1-p[i,k]) * (1-p[i,k]) # 100
        mu[i,5,k] <- (1-p[i,k]) * p[i,k] * p[i,k] # 011
        mu[i,6,k] <- (1-p[i,k]) * p[i,k] * (1-p[i,k]) # 010
        mu[i,7,k] <- (1-p[i,k]) * (1-p[i,k]) * p[i,k] # 001
        pi0[i,k] <- 1 - mu[i,1,k] - mu[i,2,k] - mu[i,3,k] - mu[i,4,k] - mu[i,5,k] - mu[i,6,k] -
mu[i,7,k]
        pcap[i,k] <- 1 - pi0[i,k] #Probability of sighting by at least one observer
        for (j in 1:7) {
          muc[i,j,k] <- mu[i,j,k] / pcap[i,k]
        } #j
        # Observation model for multiple observers - conditional multinomial
        ncap[i,k] ~ dbin(pcap[i,k] , avails[i,k])
        Y[i,1:7,k] ~ dmulti(muc[i,1:7,k] , ncap[i,k])
      } #k
    } #i
    # Derived quantities
    totalM <- sum(M[])
  } #model

```

```

    ",fill = TRUE)
sink()
}

if(k==2){
sink("model_2obs.txt")
cat("
  model {
    # Priors
    phi~dunif(0,1)
    p0~dunif(0,1)
    lambda~dunif(0,100)

    # Likelihood
    # Biological model for true abundance
    for (i in 1:R) {
      M[i] ~ dpois(lambda)
    # Observation model for replicated counts
    for (k in 1:T) {
      avails[i,k] ~ dbin(phi , M[i])
    # Define multinomial cell probabilities
      p[i,k] <- p0
      mu[i,1,k] <- p[i,k] * p[i,k]          # 11
      mu[i,2,k] <- p[i,k] * (1-p[i,k])     # 10
      mu[i,3,k] <- (1-p[i,k]) * p[i,k]     # 01
      pi0[i,k] <- 1 - mu[i,1,k] - mu[i,2,k] - mu[i,3,k]
      pcap[i,k] <- 1 - pi0[i,k] #Probability of sighting by at least one observer
      for (j in 1:3) {
        muc[i,j,k] <- mu[i,j,k] / pcap[i,k]
      } #j
    # Observation model for multiple observers - conditional multinomial
      ncap[i,k] ~ dbin(pcap[i,k] , avails[i,k])
      Y[i,1:3,k] ~ dmulti(muc[i,1:3,k] , ncap[i,k])
    } #k
  } #i
  # Derived quantities
  totalM <- sum(M[])
} #model
",fill = TRUE)
sink()
}

# Model to be used
model <- paste("model_",k,"obs.txt", sep="")
# Initial values
inits <- function() list(p0=runif(1,0,1),phi=runif(1,0,1),M=nmax+2,avails=ncap+1)
# Parameters monitored
params <- c("totalM", "lambda","p0","phi")
# MCMC settings
ni <- 50000
nt <- 1
nb <- 20000
nc <- 3

# Running!
out <- jags(data1, inits, params, model, n.chains = nc,
            n.thin = nt, n.iter = ni, n.burnin = nb, parallel=T)

# Get max counts
max.count <- numeric(0)

```



```

for(i in 1:dim(Y)[3]){
  max.count[i] <- sum(rowSums(Y[, ,i]))
}

# Plot posteriors
par(mfrow=c(2,2),mar=c(4,4,1,1),omi=c(.3,.3,.1,.1))

## phi
hist(out$sims.list$phi, col = "gray", main = "", xlab = "availability",breaks=60,freq=F,
      las = 1, xlim = c(0,1))
abline(v = phi, lwd = 3, col = "red")
abline(v=out$mean$phi,lwd=3,col="blue")
abline(v=c(out$q2.5$phi,out$q97.5$phi),lwd=2,col="gray",lty=3)

## p
hist(out$sims.list$p0, col = "gray", main = "", xlab = "detection prob.",
      breaks=60,freq=F,las = 1, xlim = c(0,1))
abline(v = p, lwd = 2, col = "red")
abline(v=out$mean$p0,lwd=3,col="blue")
abline(v=c(out$q2.5$p0,out$q97.5$p0),lwd=2,col="gray",lty=3)

## lambda
hist(out$sims.list$lambda, col = "gray", main = "", xlab = "local abundance",breaks=60
      ,freq=F,las = 1, xlim = c(0,10))
abline(v = lambda, lwd = 3, col = "red")
abline(v=out$mean$lambda,lwd=3,col="blue")
abline(v=c(out$q2.5$lambda,out$q97.5$lambda),lwd=2,col="gray",lty=3)
abline(v=mean(apply(ncap,1,max)),lwd=3)
legend("topright",c("Real","Estimativa","I.C.95%","Contagem"),lty=c(1,1,3,1),lwd=2,col=c(2,4
,"gray",1),bty="n",cex=.6)

## total M
hist(out$sims.list$totalM, col = "gray", main = "", xlab = "total N",breaks=60,freq=F,
      las = 1, xlim = c(0,300))
abline(v = sum(M), lwd = 3, col = "red")
abline(v=out$mean$totalM,lwd=3,col="blue")
abline(v=c(out$q2.5$totalM,out$q97.5$totalM),lwd=2,col="gray",lty=3)
abline(v=max(max.count),lwd=3)

```

Apêndice S2. Resultados estendidos das simulações.

Tabela 2. Cobertura do intervalo de credibilidade de 95% (IC95%) das estimativas de abundância local ($\lambda=4$) de um *multinomial N-mixture model* com emigração temporária, em cenários simulados de número de observadores e de réplicas temporais. A cobertura foi calculada a partir da proporção dos 999 IC95% que abrangeu o valor real utilizado para simular os dados.

Nº de observadores	3 voos	5 voos	10 voos
1	0,91	0,99	0,92
2	0,96	0,99	0,92
3	0,97	0,93	0,93

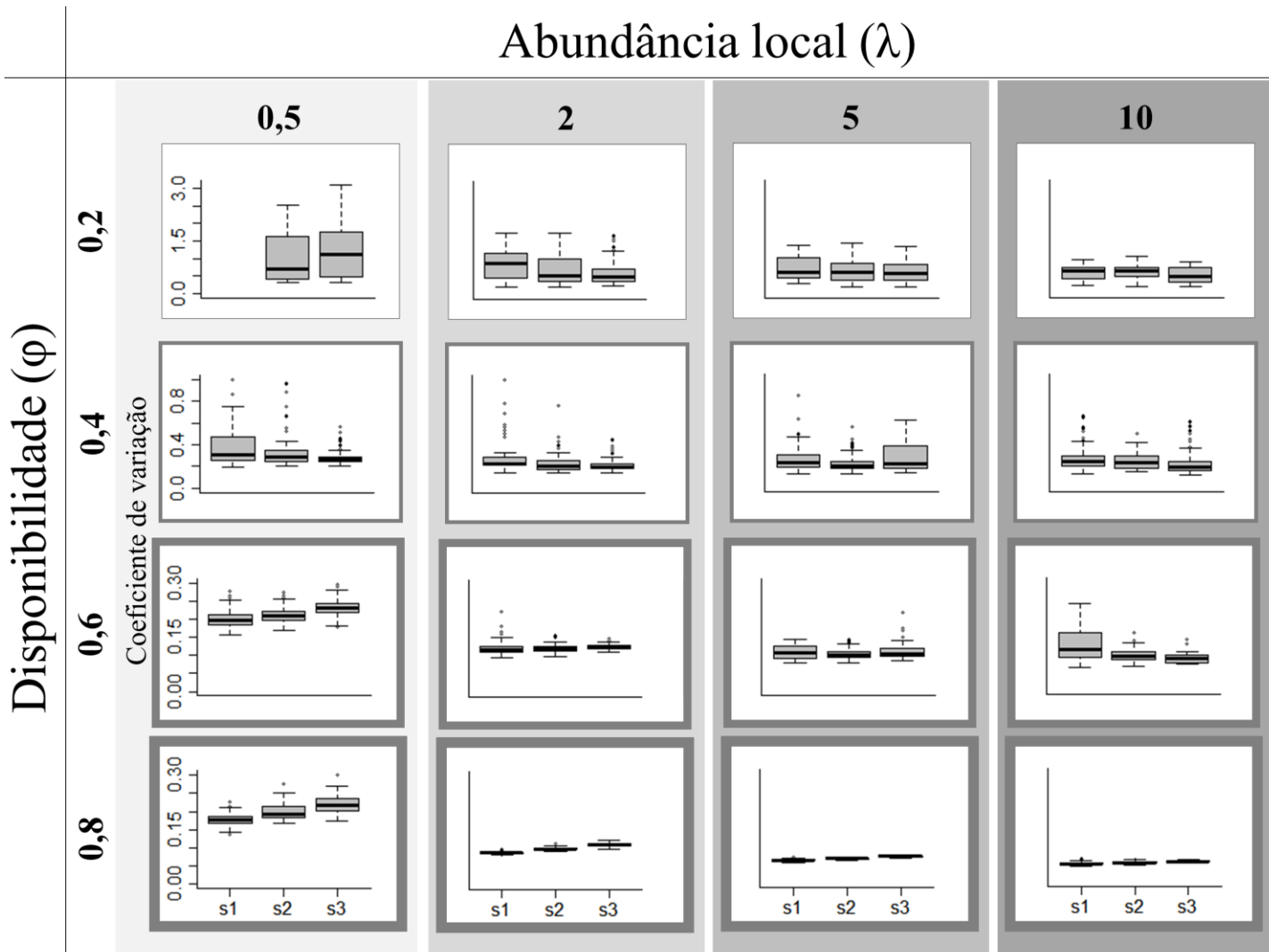


Figura 1. Coeficiente de variação das estimativas de abundância local de um *multinomial N-mixture model* com emigração temporária em diferentes cenários de disponibilidade e abundância local, sob três estratégias de amostragem. Para uma amostragem de 200 voos, **s1**: R = 67, T = 3; **s2**: R = 50, T = 4; **s3**: R = 40, T = 5. Onde R = unidades espaciais e T = replicações temporais. Os boxplots indicam a distribuição de 300 coeficientes de variação, calculados a partir do quociente entre o desvio padrão e a média posterior.

CONCLUSÕES

Com o auxílio das novas tecnologias, a coleta de dados ecológicos está atingindo condições nunca antes imaginadas (Pimm *et al.* 2015). O desenvolvimento tecnológico nos permite agora acompanhar os movimentos de um indivíduo em tempo real, realizar testes de paternidade com fezes de animais silvestres, acessar remotamente lugares inacessíveis, entre outros. Nesse contexto, os *drones* surgem como uma ferramenta que promete revolucionar a coleta de dados de populações animais (Anderson & Gaston 2013).

Apresentando-se como uma alternativa mais versátil, eficiente e segura aos levantamentos aéreos tripulados, os *drones* tem sido utilizados em estudos de diversas espécies (p. ex., aves aquáticas, grandes mamíferos terrestres, mamíferos marinhos; Linchant *et al.* 2015). Entretanto, a maioria desses estudos acessa o tamanho populacional através de simples contagens brutas, ignorando os erros de detecção que são inerentes à maioria dos estudos em ecologia animal (Mackenzie *et al.* 2006; Royle & Dorazio 2008).

Nós apresentamos as fontes de erro de medição que podem gerar vieses nas estimativas de abundância em levantamentos aéreos com *drones* (Figura 1 do Capítulo 1). A partir dessa síntese, verificamos a influência dessas fontes de erro em métodos de estimativa de abundância utilizados em estudos de levantamentos com *drones*. Realizamos um experimento em uma população de gado de tamanho real conhecido e demonstramos que todos os métodos apresentam vieses nas estimativas de abundância. A fonte de erro que apresentou maior influência foi o viés de disponibilidade, ou seja, indivíduos escondidos, não disponíveis para detecção. Assim, as estimativas geradas com o emprego de *drones* devem ser interpretadas com cautela, uma vez que podem estar enviesadas. Entretanto, nem tudo está perdido. A facilidade de replicação espacial e temporal dos *drones* permite o desenvolvimento de novas abordagens estatísticas.

E assim, no segundo capítulo, propomos uma abordagem para acomodar os erros de medição em levantamentos aéreos com *drones*. Dessa forma, avaliamos um modelo hierárquico (*multinomial N-mixture model*; Chandler *et al.* 2011; Kéry & Royle 2016) em um estudo de simulação para situações factíveis de amostragem com *drones*. O modelo apresentou estimativas acuradas e precisas na maioria dos cenários simulados, exceto em situações de baixa disponibilidade de indivíduos. Com essa abordagem, foi possível levar em conta as principais

fontes de erro em levantamentos aéreos (Pollock *et al.* 2006). Além disso, com a estrutura básica do modelo proposto, novas fontes de erro podem ser acomodadas (p.ex., falsos positivos ou falha de detecção em métodos semiautomatizados) e também podem ser incorporadas estruturas de dinâmica populacional (Kéry *et al.* 2009; Dail & Madsen 2011; Bellier *et al.* 2016) para monitorar populações em longo prazo.

Esta dissertação surgiu da preocupação em gerar estimativas confiáveis de abundância com *drones*, para assim serem aplicadas nas mais diversas situações em ecologia, conservação e manejo. Novas abordagens estatísticas já eram requisitadas (Linchant *et al.* 2015) e a consolidação dessa nova ferramenta para esses estudos depende de avaliações sobre a confiabilidade do seu uso, ou seja, da obtenção de estimativas acuradas e precisas. Com esta dissertação, esperamos ter dado algumas contribuições importantes nesse sentido.

Mas o caminho em busca da estimativa “perfeita” está longe de terminar. A abordagem aqui proposta precisa ser testada com dados empíricos, em diversos ambientes, para diversas espécies. Aliando a nova tecnologia dos *drones* e a abordagem de modelagem hierárquica, esperamos poder investigar perguntas ecológicas ainda não respondidas e monitorar com maior efetividade espécies ameaçadas.

Por fim, não podemos cair na tentação de pensar que as tecnologias resolverão todos os nossos problemas. As fontes de erro de medição ocorrem na maioria dos estudos em ecologia animal, independentemente da técnica de coleta de dados utilizada. Para cada método, para cada espécie, em cada local, diferentes fontes de erro podem apresentar influências maiores ou menores e devemos ter consciência dessas influências no momento de delinear a amostragem. O planejamento de um estudo deve passar por uma análise detalhada e criteriosa das suas perguntas/objetivos (por quê?), o que medir para respondê-las (o quê?) e como medir de maneira adequada (como?) (Yoccoz *et al.* 2001). E embora esta dissertação enfoque basicamente a terceira pergunta (como?), todo o esforço de coleta pode ser em vão se as duas perguntas anteriores não estiverem claras.

REFERÊNCIAS

- Anderson, D.R. (2001). The need to get the basics right in wildlife field studies. *Wildlife Society Bulletin*, **29**, 1294–1297.
- Anderson, K. & Gaston, K.J. (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment*, **11**, 138–146.
- Bellier, E., Kéry, M. & Schaub, M. (2016). Simulation-based assessment of dynamic N - mixture models with density-dependence and environmental stochasticity in vital rates. *Methods in Ecology and Evolution*, n/a–n/a.
- Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2001). *Introduction to Distance Sampling*. Oxford University Press, Oxford.
- Chandler, R.B., Royle, J.A. & King, D.I. (2011). Inference about density and temporary emigration in unmarked populations. *Ecology*, **92**, 1429–1435.
- Christie, K.S., Gilbert, S.L., Brown, C.L., Hatfield, M., Biology, A., Fairbanks, A., Koyokuk, N. & Hanson, L. (2016). Unmanned aircraft systems in wildlife research: current and future applications of a transformative technology. *Frontiers in Ecology and the Environment*, **14**, 241–251.
- Cooch, E.G. & White, G.C. (2016). Program MARK: A Gentle Introduction. <http://www.phidot.org/software/mark/docs/book/>.
- Dail, D. & Madsen, L. (2011). Models for Estimating Abundance from Repeated Counts of an Open Metapopulation. *Biometrics*, **67**, 577–587.
- Edwards, H.H., Pollock, K.H., Ackerman, B.B., Reynolds, J.E. & Powell, J.A. (2007). Estimation of Detection Probability in Manatee Aerial Surveys at a Winter Aggregation Site. *Journal of Wildlife Management*, **71**, 2052–2060.
- Ferreira, S.M. & Aarde, R.J. Van. (2009). Aerial Survey Intensity as a Determinant of Estimates of African Elephant Population Sizes and Trends. *South African Journal of Wildlife Research*, **39**, 181–191.
- Griffin, P.C., Lubow, B.C., Jenkins, K.J., Vales, D.J., Moeller, B.J., Reid, M., Happe, P.J., Mccorquodale, S.M., Tirhi, M.J., Schaberl, J.P. & Beirne, K. (2013). A hybrid double-observer sightability model for aerial surveys. *The Journal of Wildlife Management*, **77**, 1532–1544.

- Guillera-Arroita, G., Ridout, M.S., Morgan, B.J.T. & Linkie, M. (2012). Models for species-detection data collected along transects in the presence of abundance-induced heterogeneity and clustering in the detection process. *Methods in Ecology and Evolution*, **3**, 358–367.
- Hodgson, J.C., Baylis, S.M., Mott, R., Herrod, A. & Clarke, R.H. (2016). Precision wildlife monitoring using unmanned aerial vehicles. *Scientific Reports*, **6**, 22574.
- Kéry, M., Dorazio, R.M., Soldaat, L., Van Strien, A., Zuiderwijk, A. & Royle, J.A. (2009). Trend estimation in populations with imperfect detection. *Journal of Applied Ecology*, **46**, 1163–1172.
- Kéry, M. & Royle, J.A. (2016). *Applied hierarchical modeling in ecology, volume 1 (Prelude and Static model)*. Academic Press, New York.
- Kéry, M., Royle, J.A. & Schmid, H. (2005). Modeling avian abundance from replicated counts using binomial mixture models. *Ecological Applications*, **15**, 1450–1461.
- Kery, M. & Schaub, M. (2012). *Bayesian Population Analysis using WinBUGS: A Hierarchical Perspective*. Academic Press.
- Linchant, J., Lisein, J., Semeki, J., Lejeune, P. & Vermeulen, C. (2015). Are unmanned aircraft systems (UASs) the future of wildlife monitoring? A review of accomplishments and challenges. *Mammal Review*, **45**, 239–252.
- MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Andrew Royle, J. & Langtimm, C.A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, **83**, 2248–2255.
- Mackenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.L. & Hines, J.E. (2006). *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Elsevier, San Diego.
- Martin, T.G., Wintle, B.A., Rhodes, J.R., Kuhnert, P.M., Field, S.A., Low-Choy, S.J., Tyre, A.J. & Possingham, H.P. (2005). Zero tolerance ecology: Improving ecological inference by modelling the source of zero observations. *Ecology Letters*, **8**, 1235–1246.
- Mourão, G., Coutinho, M., Mauro, R., Campos, Z., Tomás, W. & Magnusson, W. (2000). Aerial surveys of caiman, marsh deer and pampas deer in the Pantanal Wetland of Brazil. *Biological Conservation*, **92**, 175–183.

- Murray, D.L. & Fuller, M.R. (2000). A Critical Review of the Effects of Marking. *Research Techniques in Animal Ecology: Controversies and Consequences* (eds L. Boitani & T.K. Fuller), pp. 15–64. Columbia University Press, New York.
- Pimm, S.L., Alibhai, S., Bergl, R., Dehgan, A., Giri, C., Jewell, Z., Joppa, L., Kays, R. & Loarie, S. (2015). Emerging Technologies to Conserve Biodiversity. *Trends in Ecology & Evolution*, **30**, 685–696.
- Pollock, K.H., Marsh, H.D., Lawler, I.R. & W, A.M. (2006). Estimating Animal Abundance in Heterogeneous Environments: An Application to Aerial Surveys for Dugongs. *Journal of Wildlife Management*, **70**, 255–262.
- Rodgers, J.A., Kubilis, P.S. & Nesbitt, S.A. (2005). Accuracy of Aerial Surveys of Waterbird Colonies. *Waterbirds*, **28**, 230–237.
- Royle, J.A. (2004). N-Mixture Models for Estimating Population Size from Spatially Replicated Counts. *Biometrics*, **60**, 108–115.
- Royle, J.A., Chandler, R.B., Sollmann, R. & Gardner, B. (2014). *Spatial Capture-Recapture*. Academic Press.
- Royle, J.A. & Dorazio, R.M. (2008). *Hierarchical Modeling and Inference in Ecology*. Academic Press, New York.
- Royle, J.A. & Dorazio, R.M. (2006). Hierarchical models of animal abundance and occurrence. *Journal of Agricultural, Biological, and Environmental Statistics*, **11**, 249–263.
- Royle, J.A. & Link, W. a. (2006). Generalized site occupancy models allowing for false positive and false negative errors. *Ecology*, **87**, 835–841.
- Williams, B.K., Nichols, J.D. & Conroy, M.J. (2002). *Analysis and Management of Animals Populations*. Academic Press, San Diego.
- Yoccoz, N.G., Nichols, J.D. & Boulinier, T. (2001). Monitoring of biological diversity in space and time. *Trends in Ecology and Evolution*, **16**, 446–453.