


## RESEARCH ARTICLE

# Imperfect detection and wildlife density estimation using aerial surveys with infrared and visible sensors

Zackary J. Delisle<sup>1</sup> , Patrick G. McGovern<sup>1</sup>, Brian G. Dillman<sup>2</sup> & Robert K. Swihart<sup>1</sup>

<sup>1</sup>Department of Forestry and Natural Resources, Purdue University, West Lafayette Indiana, 47907, USA

<sup>2</sup>Department of Aviation Technology, Purdue University, West Lafayette Indiana, 47907, USA

## Keywords

Density estimation, detection error, drones, infrared video, *Odocoileus virginianus*, thermal sensor

## Correspondence

Zackary J. Delisle, Department of Forestry and Natural Resources, Purdue University, West Lafayette, Indiana 47907, USA.  
[zdelisle@purdue.edu](mailto:zdelisle@purdue.edu)

## Funding Information

Funding was provided by the Indiana Department of Natural Resources Grant W-48-R-02.

Editor: Temuulen Sankey

Associate Editor: Gaia Vaglio Laurin

Received: 3 May 2022; Revised: 12 July 2022; Accepted: 29 August 2022

doi: 10.1002/rse2.305

*Remote Sensing in Ecology and Conservation* 2023, **9** (2):222–234

## Abstract

Aerial vehicles equipped with infrared thermal sensors facilitate quick density estimates of wildlife, but detection error can arise from the thermal sensor and viewer of the infrared video. We reviewed published research to determine how commonly these sources of error have been assessed in studies using infrared video from aerial platforms to sample wildlife. The number of annual articles pertaining to aerial sampling using infrared thermography has increased drastically since 2018, but past studies inconsistently assessed sources of imperfect detection. We illustrate the importance of accounting for some of these types of error in a case study on white-tailed deer *Odocoileus virginianus* in Indiana, USA, using a simple double-observer approach. In our case study, we found evidence of false negatives associated with the viewer of infrared video. Additionally, we found that concordance between the detections of two viewers increased when using a red-green-blue camera paired with the infrared thermal sensor, when altitude decreased and when more stringent criteria were used to classify thermal signatures as deer. We encourage future managers and ecologists recording infrared video from aerial platforms to use double-observer methods to account for viewer-induced false negatives when video is manually viewed by humans. We also recommend combining infrared video with red-green-blue video to reduce false positives, applying stringent verification standards to detections in infrared and red-green-blue video and collecting data at lower altitudes over snow when needed.

## Introduction

Aerial platforms are commonly used for surveying in wildlife research and management (Hone, 2008; Pollock & Kendall, 1987). Sampling animals from an aerial platform facilitates quick population estimates (McMahon et al., 2021). Traditionally, aerial sampling has been performed by human observers who ride in a plane or helicopter and detect animals out the sides of the aircraft (Pollock & Kendall, 1987). Collecting data and estimating density under various distance-sampling, mark–recapture or a combination of these two frameworks has thus been common (Barker, 2008; Fewster & Pople, 2008).

Infrared (IR) thermography, in which thermal cameras capture images from infrared radiation emitted from objects, has been adopted by wildlife researchers to

estimate population abundance (Havens & Sharp, 2015). In the context of aerial sampling, thermal sensors have been oriented two different ways when used to detect animals: (1) forward-looking infrared (FLIR) thermal sensors, which look forward from the aerial vehicle at oblique angles (Storm et al., 2011; Sudholz et al., 2021) and (2) vertical-looking infrared (VLIR) thermal sensors, which look directly beneath the aircraft (i.e. nadir orientation; Kissell & Nimmo, 2011).

Using IR thermal sensors to sample wildlife may introduce error from at least two sources. First, the thermal sensor may yield false-negative errors by failing to detect animals. Similar to Brack et al. (2018), we refer to these as availability errors. Poor or no thermal contrast (e.g. thick vegetation, lack of cloud cover (resulting in poor thermal contrast) or overhangs) induce availability errors

(Havens & Sharp, 2015). Therefore, study designs often consider weather and altitude to minimize the potential of availability errors. The use of VLIR instead of FLIR also can help to alleviate availability errors caused by vegetative obstruction because the distance and angle, and thereby the amount of vegetative or topographical obstruction, between the thermal sensor and animal is lesser (Kissell & Nimmo, 2011). In general, strategies to account for availability errors seem logistically challenging (discussed more below).

Other sources of error can arise while viewing the IR video if viewers incorrectly classify IR signatures (Stander et al., 2021). Specifically, false positives and negatives can be caused by the viewer or automated classification algorithm. Similar to Brack et al. (2018), we refer to false positives and negatives from the viewer as misidentification errors and perception errors, respectively. Because of the potential for misidentification errors, some researchers have simultaneously captured high-resolution red-green-blue (RGB) digital imagery with IR video, which is then used to confirm heat signatures as the target species (Franke et al., 2012; Schoenecker et al., 2018). An additional type of false positive can occur from double counting. Double counting can arise from overlapping transects, successive images containing the same individual, or animal movement in between neighboring transects (Brack et al., 2018; discussed more below).

We conducted a literature review to determine how frequently the aforementioned types of errors have been addressed in wildlife research using thermal sensors from aerial platforms. We then illustrated the potential importance of viewer errors in a case study where we simultaneously captured VLIR and RGB video across three different regions of Indiana, USA. To evaluate the magnitude and impact of perception errors in IR sampling, we estimated the density of white-tailed deer *Odocoileus virginianus* using the detections of two viewers and a single viewer. To assess the effects of study design on the concordance of IR detections, we computed density using video collected at multiple altitudes and differing ground conditions. Lastly, to better understand the importance of confirming IR signatures with RGB video, we estimated densities using: (1) only IR video and (2) IR and RGB video.

## Materials and Methods

### Literature review

We reviewed published studies that used thermal sensors to study wildlife from aerial vehicles by conducting the following keyword search on Web of Science™ on 1 January 2022: “(thermal OR infrared) AND wildlife AND

(aerial OR plane OR airplane OR helicopter OR unmanned OR drone)”. For each article, we recorded: (1) year of publication; (2) whether the study pertained to behavior, presence/absence, simple counts or density/abundance estimation; (3) if imperfect detection was addressed in any way, and if so, whether uniform detection across the field-of-view of the camera, availability errors, perception errors, misidentification errors or double-counting errors was addressed; (4) if photos, video or active searching was used to collect data; (5) thermal sensor orientation (forward-looking or vertical-looking); (6) if IR, RGB or IR and RGB video was recorded; (7) if the aerial vehicle was crewed or uncrewed and (8) if automated software, human viewers or automated software and human viewers were used to review video and detect the target species. We did not include articles that used data from only simulation, did not pertain to wildlife, did not use an aerial vehicle, were not fully available for review (e.g. only abstracts available), or previous review articles without a case study.

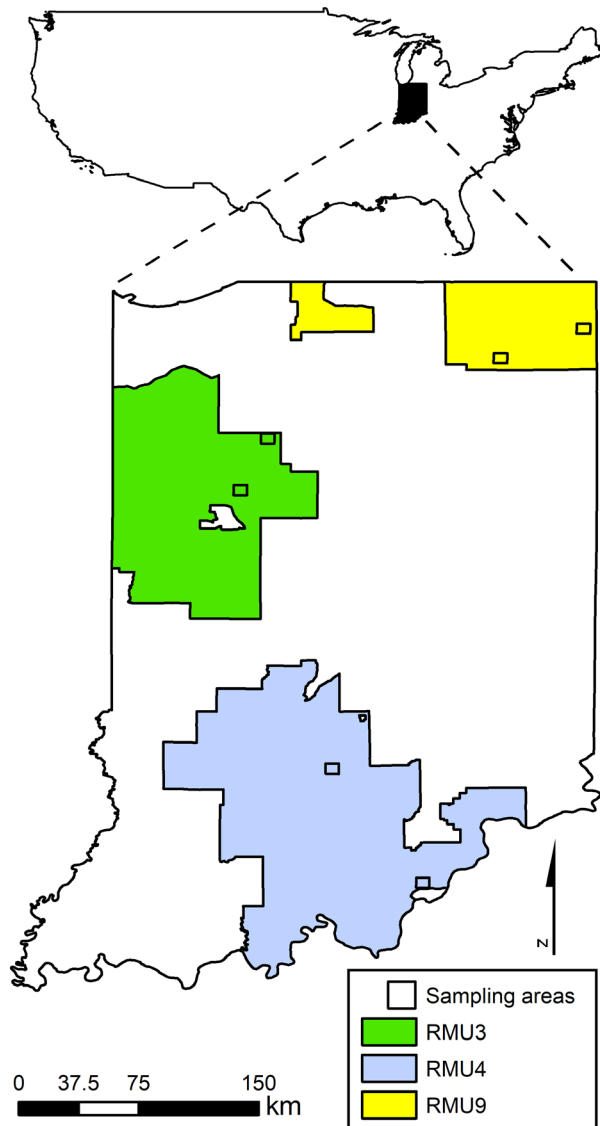
## Case study

### Study sites

We sampled deer populations in Indiana’s Deer Regional Management Units 3, 4 and 9 (Delisle et al., 2022; Swihart et al., 2020). We surveyed two different 6.4 × 6.4-km areas within each Regional Management Unit (hereafter, RMU), resulting in six total areas flown (Fig. 1). We randomly selected these areas from the deer reporting grid used by the Indiana Department of Natural Resources to collect spatially explicit harvest data. Regional Management Unit 3 is an intensively farmed region with 79% row-crop agriculture, 10% forest, 3% grassland and 1% wetland. Unlike RMU 3, RMU 4 is mainly forested with 19% row-crop agriculture, 56% forest, 16% grassland and <1% wetland. Lastly, RMU 9 is 56% row-crop agriculture, 8% forest, 11% grassland and 13% wetland. All three RMUs follow a four-season temperate weather pattern.

### Data collection

In each sampling area, we flew 16 systematically placed 6.4-km transects. Adjacent transects were separated by 400 m and aligned north to south. We flew transects in a Sky Arrow Light Sport Aircraft (Magnaghi Aeronautica S.p.A.) at speeds of ~105 kph. We flew at an altitude of 450 m in RMU 3 when there was snow cover, an altitude of 450 m in RMU 4 when there was no snow cover, and an altitude of 300 m when there was snow cover in RMU 9. Flights occurred during daytime hours (30 min after sunrise to 30 min before sunset) from 8 February to 10



**Figure 1.** Sampling areas within the Deer Regional Management Units (RMU) of Indiana, USA, that we surveyed. White-tailed deer *Odocoileus virginianus* were detected using thermal and color sensors in a crewed aerial platform during daylight hours from 8 February to 10 March 2021.

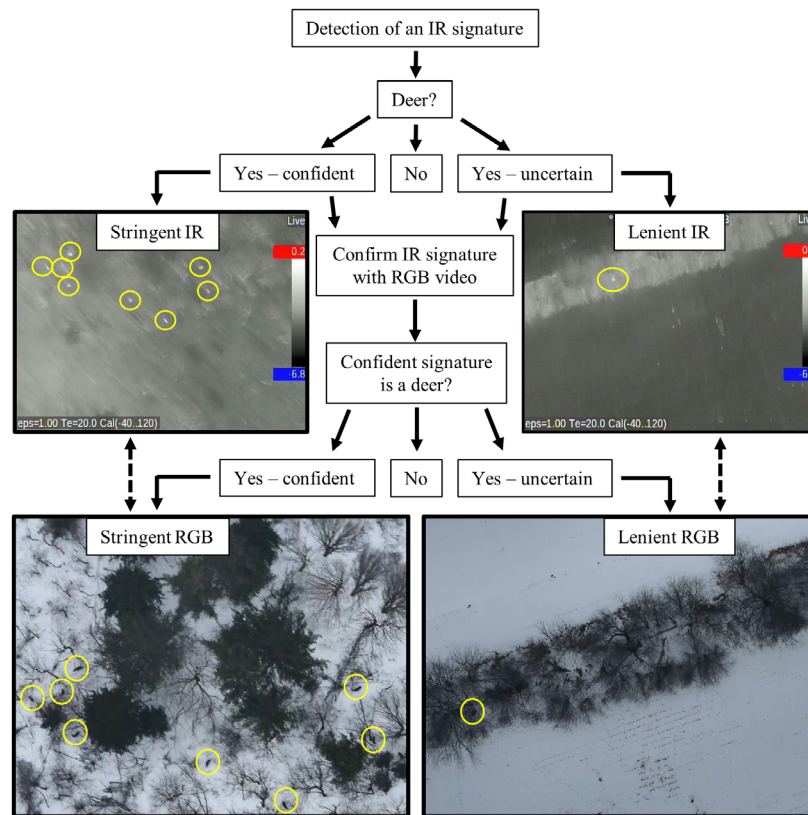
March 2021 when deciduous trees had already shed leaves. We surveyed during daytime so RGB video could be captured simultaneously. Regardless of ground condition or altitude, we only collected aerial data on overcast days while flying under cloud cover, which heightened the thermal contrast between background temperatures and body temperatures of deer.

During flights, we recorded VLIR and RGB video using an IR-TCM HD 1024 stationary thermal sensor equipped with a 60 mm lens (Jenoptik, Jena, Germany), and a

Nikon D810 DSLR camera equipped with a Nikon AF DC-NIKKOR 135 mm *f*/2D lens (Nikon Inc., Melville, NY). Cameras were affixed to opposite sides of the aircraft and focused on the ground directly beneath the aircraft. We simultaneously recorded IR and RGB video while georeferencing and digitally storing the video using a GeoDVR Mini (Remote GeoSystems, Inc.) equipped with a Garmin global positioning system (Garmin Ltd.).

We viewed IR and RGB video using the LineVision – Ultimate software (Remote GeoSystems, Inc.). While viewing video, two independent viewers that were highly trained and experienced in viewing IR video (3 sampling seasons of experience) recorded four different qualitative classes of detections (Fig. 2): (1) detections from IR video for which the heat signature had any potential to be a deer but the viewer was not confident that the heat signature was a deer (lenient IR detections); (2) detections from IR video in which the viewer was confident that the heat signature was from a deer (stringent IR detections); (3) RGB confirmations of lenient or stringent IR heat signatures in which the object in the RGB video had any potential to be a deer but the viewer was not confident the object was a deer (lenient RGB confirmations) and (4) RGB confirmations of lenient or stringent IR heat signatures in which the viewer was confident that the object in the RGB video was a deer (stringent RGB confirmations). When assigning IR heat signatures or RGB objects, stringent detections were those with a shape that was clearly defined and unambiguous so that we believed no other object than a deer could be producing such an IR heat signature or RGB object. No other species of similar shape and color to that of white-tailed deer were present in our field site. To avoid bias, the two viewers scored videos independently (i.e. did not aid each other when classifying images) and only defined an RGB confirmation as either 3 or 4 after first defining an IR detection as 1 or 2 (Fig. 2). For all detections, we recorded whether or not the detection was within concealed (forest, wetland) or open (grassland, agricultural field) habitat. We used the LineVision – Ultimate software to measure the perpendicular distance from the middle of the thermal sensor's field-of-view to the IR heat signature. Lastly, we recorded how many other lenient or stringent IR detections were within the immediate vicinity of the IR detection (henceforth referred to as group size). We considered IR detections to be in the immediate vicinity of each other if the detections could appear in the field-of-view of the thermal sensor at the same time.

We compiled four detection histories using these four detection classes: (1) detection history using lenient and stringent IR detections; (2) detection history using only stringent IR detections; (3) detection history using lenient or stringent RGB confirmations of lenient or stringent IR



**Figure 2.** Workflow for classifying infrared (IR) heat signatures and corresponding objects in red-green-blue (RGB) as either white-tailed deer *Odocoileus virginianus*, not deer or unsure. Video was captured in a crewed aircraft in Indiana, USA, from 8 February to 10 March 2021. All potential IR heat signatures are classified as either lenient IR detections (defined as detections from IR video for which the heat signature had any potential to be a deer but the viewer was not confident the heat signature was a deer) or stringent IR detections (defined as detections from IR video in which the viewer was confident that the heat signature was from a deer). All RGB confirmations were classified as either lenient RGB confirmations (defined as RGB confirmations of lenient IR heat signatures in which the object in the RGB video had any potential to be a deer but the viewer was not confident the object was a deer) or stringent RGB confirmations (defined as RGB confirmations of lenient IR heat signatures in which the viewer was confident that the object in the RGB video was a deer). When assigning IR heat signatures or RGB objects, stringent detections were those with a shape that was clearly defined and unambiguous so that we believed no other object than a deer could be producing such an IR heat signature or RGB object. No other species of similar shape and color to that of white-tailed deer were present in our field site.

detections and (4) detection history using stringent RGB confirmations of lenient or stringent IR detections. We did not assess a scenario in which RGB was used to only confirm stringent IR detections because if RGB is available, we would expect users to check all potential IR heat signatures.

### Single viewer

We estimated density from the four different types of detection histories using the observations of a single viewer (ZJD) in each of the altitudes and ground conditions we sampled. We first considered the possibility that detection probability decreased with increasing distance from the middle of the thermal sensor's field-of-view. However, we found evidence that detection probability

across the field-of-view was uniform according to Akaike's Information Criterion (AIC) and visual plots of several candidate detection functions (see Supporting Information: Distance Sampling Analysis for more details). Therefore, we used plot sampling techniques to estimate density (Buckland et al., 2015) with the following formula:

$$\hat{D} = \frac{n}{A} \quad (1)$$

where  $n$  = the total number detections,  $A$  = the total area sampled =  $Lw2$ , where  $L$  = the total length of transect sampled and  $w$  = the transect half width. Because the variation from plot sampling techniques comes from the encounter rate (i.e. there is no variation from a detection function), we used an approach modified from the 'R2'

method in Fewster et al. (2009) to estimate the standard error as

$$SE(\hat{D}) = \sqrt{\frac{K}{A^2(K-1)} \sum_{k=1}^K a_k^2 \left( \frac{n_k}{a_k} - \frac{n}{A} \right)^2} \quad (2)$$

where  $K$  = the total number of transects,  $a_k$  = the total area sampled on transect  $k$  and  $n_k$  = the total number of detections on transect  $k$ .

### Double viewer

To ensure that the detection and error rates of the double- and single-viewer methods were directly comparable, we used the detections from the single viewer as one of the double viewers. We calculated the concordance between the two viewers (ZJD and PGM) for all four detection histories at each altitude and ground condition. To assess the need for RGB video, we determined the percentage of stringent and lenient IR detections that were confirmed by RGB video to be deer, an object other than deer or unresolved. We evaluated the value of multiple viewers by calculating the probability of a single viewer detecting a deer conditional upon the other viewer, and the probability of either viewer detecting the deer using the 'mrds' package (Laake et al., 2022) in R (R Core Team, 2022). Specifically, we fit logistic conditional detection models with a logit link in the form of eq. 6.32 in Laake and Borchers (2004). We used an independent observer configuration and assumed full independence (Burt et al., 2014). We fit a mark–recapture model for each possible additive combination of the following covariates: distance from the transect line, group size and observer (viewer 1 or viewer 2). We used AIC to decide between competing mark–recapture models. We repeated this model fitting process for each of the four different types of detection histories in each of the altitudes and ground conditions we sampled. We did not test the effect of habitat type (open vs. concealed) on detection probability because we detected too few deer in open habitat. Additionally, we purposefully sampled when there was no leaf cover in the canopy to avoid reduced probability of infrared detection in wooded areas.

After selecting the best mark–recapture model, we estimated density in the sampled area using a Horvitz–Thompson-like estimator (Borchers et al., 1998) in the 'mrds' R package. Variation from the mark–recapture model was estimated using the delta method (Borchers et al., 1998), and variation from the random sample selection was estimated using the encounter-rate estimator in Innes et al. (2002) in the form of the 'R2' method in Fewster et al. (2009).

## Results

### Literature review

Our search revealed 62 articles on the use of IR thermography in aerial sampling of wildlife since 1991. The number of articles increased sharply from 2018 to 2021 (Fig. S2). Uncrewed aerial vehicles have shown a particularly pronounced increase in usage, which is consistent with past reviews focused on uncrewed aerial vehicles (Linchant et al., 2015). Although the usage of dual platforms containing IR and RGB cameras has increased since 1991, only nine of 23 articles estimating density or abundance used a dual platform; two of these articles did not use RGB to confirm IR heat signatures, but instead assessed whether or not RGB could solely be used to estimate density. Two articles did not specify if they used FLIR or VLIR. Of the 62 articles we reviewed, 19 did not address any type of error, 16 addressed one type of error, 18 addressed two types of error, eight addressed three types of error, one addressed four types of error and no articles addressed all five types of error. Twenty-four articles used photos, 32 articles used video and 16 articles used active searching methods to locate IR heat signatures.

Imperfect detection across the field-of-view of the thermal sensor was assessed by seven articles; availability errors were addressed by 19 articles; perception errors were addressed by nine articles; misidentification errors were addressed by 34 articles and double-counting errors were addressed by 11 articles. Perfect detection across the field-of-view of the thermal sensor was found in four of seven (57%) of the articles that tested for uniform detection probability. Sixteen of the 23 articles (70%) estimating density assessed imperfect detection to some degree. Of these, seven assessed uniform detection probability across the camera's field-of-view, five addressed availability errors, two addressed perception errors, nine addressed misidentification errors and six addressed double-counting errors (Table 1).

### Case study

Across the four classes of detection histories, concordance between the two viewers increased when altitude was lower and when snow covered the ground (Table 2). We were unable to confirm IR signatures using RGB video in the areas that did not have snow cover on the ground, and therefore we do not present any statistics for the lenient IR with lenient or stringent RGB confirmation histories for this sampling scenario. At altitudes of 300 m, 77.2% (SE = 4.5), 6.5% (SE = 3.6) and 16.4% (SE = 0.9) of stringent IR detections were confirmed to be deer, not



**Table 1.** Contingency table reporting the total number of articles that: used forward-looking or vertical-looking infrared thermography; did not assess imperfect detection; assessed imperfect detection in terms of uniform detection probability across the field-of-view of the thermal sensor (Uniform), availability errors (AE), perception errors (PE), misidentification errors (ME) or double-counting errors (Double); used infrared (IR), red-green-blue (RGB) or IR and RGB cameras; used crewed or uncrewed aircraft and used automated viewing software (AI), manual human viewing (human) or both AI and human.

Platform orientation			Forward looking						Vertical looking							
Aircraft	Sensor	Review	Assess detection?						Assess detection							
			No	Uniform	AE	PE	ME	Double	No	Uniform	AE	PE	ME	Double	Subtotal	
Crewed	IR	AI	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Human	6	0	3	0	1	2	2	2	1	0	1	2	15	
		Both	0	0	0	0	0	0	0	1	0	1	1	0	1	
	RGB	AI	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Human	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Both	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	IR + RGB	AI	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Human	0	0	0	0	1	1	1	2	2	0	6	1	8	
		Both	1	0	0	0	0	0	0	1	0	2	2	0	3	
Uncrewed	IR	AI	0	0	0	0	0	0	0	0	0	0	0	0	0	
		Human	3	0	1	0	5	0	3	1	0	0	1	2	14	
		Both	0	0	0	0	0	0	1	0	2	2	3	2	4	
	RGB	AI	0	0	0	0	0	0	0	0	0	0	0	0	0	
		Human	0	0	0	0	0	0	0	0	0	0	0	0	0	
		Both	0	0	0	0	0	0	0	0	1	0	0	0	1	
	IR + RGB	AI	0	0	0	0	1	0	0	0	0	0	0	0	1	
		Human	1	0	2	1	2	0	0	0	5	1	8	0	11	
		Both	0	0	0	0	0	0	0	0	1	2	2	0	2	
		Subtotal	11	0	6	1	10	3	7	7	12	8	24	7	60	

Categories are mutually exclusive (e.g. an article using IR and RGB cameras would not satisfy the individual IR and RGB categories). Articles were obtained through a Web of Science™ search conducted on 1 January 2022 for the following: “(thermal OR infrared) AND wildlife AND (aerial OR plane OR airplane OR helicopter OR unmanned OR drone)”.

deer or unresolved, respectively, and 19.7% ( $SE = 6.6$ ), 60.6% ( $SE = 13.3$ ) and 19.7% ( $SE = 6.6$ ) of lenient IR detections were confirmed to be deer, not deer or unresolved, respectively. At altitudes of 450 m, 81.8% ( $SE = 0.9$ ), 10.2% ( $SE = 4.3$ ) and 8.1% ( $SE = 5.2$ ) of stringent IR detections were confirmed to be deer, not deer or unresolved, respectively, and 21.1% ( $SE = 21.1$ ), 54.0% ( $SE = 4.0$ ) and 25.0% ( $SE = 25.0$ ) of lenient IR detections were confirmed to be deer, not deer or unresolved, respectively.

Across all ground conditions, altitudes and detection histories, the probabilities of either of two viewers detecting a heat signature were on average 3.9% ( $SE = 1.6$ ) larger than the probabilities of viewer 1 detecting a heat signature (Table 3). At 300 m altitude, the probability of detection remained fairly consistent across the different detection histories. However, at 450 m altitude, probability of detection differed between the detection histories, with increasing certainty thresholds of objects in the IR and RGB video associated with lower detection probability (Table 3).

The densities and AIC-best mark–recapture models for each of the four detection histories at each altitude and ground condition are presented in Table 3 and Table 4, respectively. Observer was the only covariate in each of the AIC-best mark–recapture models. The densities across all ground conditions, altitudes and classes of detection histories from the mark–recapture estimator were on average 9.2% ( $SE = 2.8$ ) larger than the density estimates that used the detections from a single viewer (Table 3). Unlike detection probability, the densities across different detection histories at 450 m altitude were fairly consistent, but the densities at 300 m altitude differed among classes of detection histories. Increasing certainty thresholds of objects in the IR and RGB video captured at 300 m resulted in lower density estimates.

## Discussion

We found that RGB video confirmation, lower altitudes, snow cover and increasing levels of object scrutiny substantially increased concordance between our two viewers.

**Table 2.** Concordance between two viewers' detections of white-tailed deer *Odocoileus virginianus* from aerially captured infrared (IR) and red-green-blue (RGB) video collected in Indiana, USA, from 8 February to 10 March 2021.

Ground	Altitude	A	B	C	D
Snow	300	70.81	72.98	81.68	86.02
Snow	450	56.10	56.10	61.79	55.28
Bare	450	31.30	43.48	NA	NA

Both videos were simultaneously captured on an aircraft flown at two different altitudes and over two different ground conditions. Viewers detected deer under four different scenarios: (A) detections from only IR video for which the heat signature had any potential to be a deer; (B) detections from only IR video in which the viewer was confident that the heat signature was from a deer; (C) RGB confirmations of any IR heat signatures in which the object in the RGB video had any potential to be a deer and (D) RGB confirmations of any IR heat signatures where viewers were confident that the object in the RGB video was a deer. Concordance is not reported for C or D at 450 m altitude over bare ground because we were unable to utilize RGB video in these conditions. When assigning IR heat signatures or RGB objects, stringent detections were those with a shape that was clearly defined and unambiguous so that we believed no other object than a deer could be producing such an IR heat signature or RGB object. No other species of similar shape and color to that of white-tailed deer were present in our field site.

Consequently, we strongly encourage future researchers to apply stringent verification standards and RGB confirmation to data collected from low-altitude flights over snow when surveying in similar study areas and for comparable species. Classifying IR heat signatures as deer is a viewer-dependent task, and thus contains a degree of subjectivity. Indeed, concordance between our two viewers dropped to as low as 30% in areas that did not have snow cover, regardless of detection history, even though both viewers had undergone extensive training to examine IR video that we captured from aerial platforms. Other studies have reported lack of concordance between the classifications of multiple viewers of both aerial video (Beaver et al., 2020; Preston et al., 2021) and other types of population data (Delisle et al., 2022). In instances when high certainty cannot be obtained due to the natural history of the target species (e.g. arboreal species inhabiting dense canopies or little color contrast between the animal and background; Corcoran et al., 2019), ground truthing may be required.

Similar to others, our literature review revealed that uncrewed aerial vehicles are increasingly popular in wildlife monitoring (Linchant et al., 2015). One major advantage of these vehicles is the capability of flying at much lower altitudes compared to crewed aircraft, which helps to increase video quality and thus reduce misidentification and perception errors (Linchant et al., 2015). If

battery life is of no concern, some researchers have even reduced altitude upon detection of a potential heat signature in non-forested habitats, and honed in for more certain confirmation (Smith et al., 2020). Furthermore, crewed aerial vehicle accidents account for 66% of wildlife biologist deaths while on the job (Sasse, 2003). Uncrewed aerial vehicles are a much safer alternative. Lastly, uncrewed aerial vehicles have an appealing ease of use, which facilitates quick data collection (McMahon et al., 2021). That being said, large-scale management may still struggle to efficiently sample with uncrewed aerial vehicles due to line-of-sight restrictions and short battery life (Linchant et al., 2015). Practical application over large extents likely will require improved battery life and more lenient regulations pertaining to line-of-sight operation.

We tested the effects of ground condition (snow vs. bare) and altitude on the ability to use RGB video to confirm heat signatures. Because we did not sample the same areas repeatedly at different altitudes and ground conditions, we were unable to infer how detection probability changes as a function of altitude and ground condition. Flight speed is an additional variable of interest to wildlife managers using aerial methods. Although we attempted to fly surveys at a constant speed to ensure repeatability, the ability to effectively sample at faster flight speeds would be more cost and time efficient, and may also facilitate sampling larger areas. Therefore, we encourage future researchers to examine the effects of altitude and ground condition on detection probability, and the effects of flight speed on the efficacy of RGB confirmation, concordance between independent observers and detection probability.

Simultaneously capturing IR and RGB video increases logistical difficulties and monetary costs associated with aerial sampling. Logistically, synchronizing the two video streams to ensure viewers can examine the same image in IR and RGB could prove difficult to non-experts (Bushaw et al., 2020). We used the GeoDVR Mini to simultaneously capture, georeference and store IR and RGB video streams, which required little technical expertise. Additionally, the GeoDVR Mini named and stored video files for simultaneous examination of images in IR and RGB bands within the LineVision – Ultimate software. We thus avoided the step of manually lining up two separate video streams, which can be challenging for large video files. Monetarily, the purchase of an additional RGB camera increased the cost of sampling. However, the added cost of the RGB camera (\$4089 USD) was small compared to the IR thermal sensor (\$33 488 USD). The additional cost was essential, as RGB confirmation substantially improved our object classifications. Other researchers surveying for species inhabiting more open habitat may only need RGB

**Table 3.** Density estimates ( $\hat{D}$ ) and detection probabilities (Pr(det)) of white-tailed deer (*Odocoileus virginianus*) from aerially captured infrared (IR) and red-green-blue (RGB) video collected in Indiana, USA, from 8 February to 10 March 2021.

Ground	Altitude	Method	Detection history	$\hat{D}$	se( $\hat{D}$ )	CV( $\hat{D}$ )	Pr(Det)	se(Pr[Det])
Snow	300	MR	Lenient IR	19.89	2.91	0.15	0.98	<0.01
			Stringent IR	18.02	2.76	0.15	0.99	<0.01
			Lenient RGB	16.68	3.02	0.18	0.99	<0.01
		PS	Stringent RGB	13.40	2.96	0.22	0.99	<0.01
			Lenient IR	18.79	2.92	0.16	0.98	0.01
			Stringent IR	17.32	2.72	0.16	0.98	0.01
	450	MR	Lenient RGB	16.15	2.98	0.18	0.98	0.01
			Stringent RGB	12.93	2.89	0.22	0.98	0.01
			Lenient IR	5.43	1.84	0.34	1.00	<0.01
		PS	Stringent IR	5.27	1.67	0.32	0.95	0.02
			Lenient RGB	5.60	1.89	0.34	0.94	0.02
			Stringent RGB	5.67	1.93	0.34	0.91	0.03
Bare	MR	Lenient IR	5.43	1.84	0.34	1.00	<0.01	
		Stringent IR	4.64	1.41	0.30	0.88	0.04	
		Lenient RGB	4.69	1.46	0.31	0.84	0.04	
	PS	Stringent RGB	4.59	1.44	0.31	0.81	0.05	
		Lenient IR	5.57	1.18	0.21	0.88	0.05	
		Stringent IR	3.97	0.90	0.23	1.00	<0.01	
	450	PS	Lenient IR	4.70	0.98	0.21	0.84	0.07
			Stringent IR	3.97	0.90	0.23	1.00	<0.01

Corresponding standard errors (se( $\hat{D}$ )) and coefficients of variation (CV( $\hat{D}$ )) are reported for density estimates, and standard errors (se(Pr[Det])) are reported for Pr(Det). Video was captured at differing altitudes (300 and 450 m) and ground conditions (bare ground and snow cover). Densities were estimated from two viewers using mark–recapture (MR) methods, or a single viewer using plot sampling (PS) methods. For PS, Pr(det) = the probability of viewer 1 detecting a deer, and for MR, Pr(det) = the probability of either viewer detecting a deer. Densities were estimated using four different types of detection histories: (1) detections from only IR video for which the heat signature had any potential to be a deer (Lenient IR); (2) detections from only IR video in which the viewer was confident that the heat signature was from a deer (Stringent IR); (3) RGB confirmations of any IR heat signatures in which the object in the RGB video had any potential to be a deer (Lenient RGB) and (4) RGB confirmations of any IR heat signatures where viewers were confident that the object in the RGB video was a deer (Stringent RGB). When assigning IR heat signatures or RGB objects, stringent detections were those with a shape that was clearly defined and unambiguous so that we believed no other object than a deer could be producing such an IR heat signature or RGB object. No other species of similar shape and color to that of white-tailed deer were present in our field site.

video if the target individuals are large enough or colored to be easily distinguishable from the background terrain (e.g. Edwards et al., 2021). If surveys include multiple species producing indistinguishable heat signatures (e.g. Gentle et al., 2018), RGB video may facilitate species classification (Lee et al., 2019).

False positive and negative errors can affect bias and precision of occupancy and density estimates (Miller et al., 2011; Otis et al., 1978; Royle & Link, 2006; Strickfaden et al., 2020). In the context of aerial sampling with thermal sensors, availability errors can arise from the thermal sensor when the thermal sensor fails to capture the IR heat signature of an animal within the field-of-view (Bushaw et al., 2020). Perception errors arise when the viewer, or automated viewing software, fails to detect an IR heat signature that is present in the video stream (Preston et al., 2021). Strategies to specifically account for perception errors are few and vary in terms of their validity. Manual viewing has been used to count perception

errors from automated software (e.g. Conn et al., 2021; Lethbridge et al., 2019), but this relies on the dubious assumption that human viewers do not commit perception errors—an assumption our work does not support. When counting walrus using IR thermography, Burn et al. (2009) modeled detection probability as a function of group size to help alleviate perception errors. Conducting auxiliary ground truthing by walking the flown transects and recording the coordinates of confirmed individuals can quantify perception errors (Corcoran et al., 2019). This strategy may be more effective for sedentary species, as mobile species will likely flush or move before being sampled by walkers. Double-observer methods are well known in ecological studies (Nichols et al., 2000), and enable the modeling of heterogeneity in detection probability across predictors (Laake & Borchers, 2004). Moreover, open access software for fitting such models is readily available for ecologists and managers (Laake et al., 2022). Even so, to our knowledge, we are



**Table 4.** The AIC-best mark–recapture models fit to detection histories of white-tailed deer *Odocoileus virginianus*.

Ground	Altitude	Detection history	Covariates	AIC	$\Delta$ AIC <sup>1</sup>
Snow	300	Lenient IR	Distance + Observer	−1644.62	1.88
		Stringent IR	Distance + Observer	−1500.63	1.88
		Lenient RGB	Distance + Group size + Observer	−1269.93	0.49
		Stringent RGB	Distance + Group size + Observer	−1024.25	2.22
	450	Lenient IR	Observer	−549.94	1.59
		Stringent IR	Observer	−458.13	1.06
		Lenient RGB	Observer	−440.42	1.94
		Stringent RGB	Observer	−426.68	0.70
Bare	450	Lenient IR	Group size + Observer	−504.14	12.75
		Stringent IR	Group size + Observer	−446.28	8.17

Detections were extracted from aerially captured infrared (IR) and red-green-blue (RGB) video collected in Indiana, USA, from 8 February to 10 March 2021. Video was captured at differing altitudes (300 and 450 m) and ground conditions (bare ground and snow cover). Mark–recapture models were fit using four different types of detection histories: (1) detections from only IR video for which the heat signature had any potential to be a deer (Lenient IR); (2) detections from only IR video in which the viewer was confident that the heat signature was from a deer (Stringent IR); (3) RGB confirmations of any IR heat signatures in which the object in the RGB video had any potential to be a deer (Lenient RGB) and (4) RGB confirmations of any IR heat signatures where viewers were confident that the object in the RGB video was a deer (Stringent RGB). When assigning IR heat signatures or RGB objects, stringent detections were those with a shape that was clearly defined and unambiguous so that we believed no other object than a deer could be producing such an IR heat signature or RGB object. No other species of similar shape and color to that of white-tailed deer were present in our field site. Distance = the perpendicular distance in between the transect line and the deer. Observer = the observer that detected the deer. Group size = the number of deer in the same group as the detected deer.

<sup>1</sup>Difference between the best model and next best model.

the first to utilize this framework to estimate and correct for perception errors in IR sampling. We encourage future users of IR thermal sensors to employ double-viewer methods when estimating population density from video captured with aerial platforms that is manually viewed.

Unlike perception errors, accounting for availability errors is much more logistically challenging. Such error could be induced by dense overhead cover or poor thermal contrast (Havens & Sharp, 2015). Several past studies have quantified availability errors by simultaneously conducting additional studies to determine how many individuals are available for IR sampling. These have consisted of ground surveys (Brunton et al., 2020; Kays et al., 2019; McKellar et al., 2021; Witt et al., 2020) and telemetric studies (Latham et al., 2021). Both of these strategies are problematic for study species that are highly mobile. Kissell and Tappe (2004) used human surrogates to quantify availability errors; a potentially useful strategy that nonetheless assumes: (1) IR heat signatures of the surrogates are identical to the IR heat signatures from the study animal; (2) habitat in which the surrogates are placed is representative of the actual area to be sampled and (3) conditions that affect IR heat signatures in the surrogate study replicate those in the actual study. Distance sampling is an alternative method to address availability errors by accounting for decreased detection probability associated with increasing distance from the aerial transect line (Gentle et al., 2018; Schoenecker

et al., 2018). Among other things, distance sampling assumes all objects directly on the line are detected with certainty (Buckland et al., 2001). Mark–recapture distance sampling can adjust density estimates when objects are not detected with certainty on the transect line (Burt et al., 2014). However, assessing availability errors with mark–recapture methods is challenging for aerial sampling because: (1) repeat flights typically occur immediately following initial flights and thus experience the same thermal conditions which in turn causes the thermal sensor to detect the same individuals and (2) when surveying for mobile species, animal movement in between repeat flights may confound how many individuals are within the field-of-view of the thermal sensor. For these same reasons, double-observer methods to correct for perception errors are problematic when using automated viewing algorithms to detect heat signatures. We encourage future work on procedures to estimate availability errors, especially with more mobile study species.

False positives from double counting the same individuals present an additional source of error that can positively bias estimates. Generally, double counting occurs from recording the same individual twice on overlapping images or videos, and recording the same moving individual twice on neighboring transects (Brack et al., 2018). Fortunately, strategies to account for or avoid double counting exist. Lu et al. (2022) developed a hierarchical framework that utilized entity resolution to identify the same individuals in overlapping images and thus avoid

double counting when IR images are analyzed instead of video (e.g. Chrétien et al., 2015). Double counting the same individuals on neighboring transects is not problematic when animal movement is random and spatial sampling effort is accounted for (Buckland et al., 2001). Reactive movement is problematic if individuals consistently run off the transect before the aircraft samples (due to noise) or if the aircraft continually pushes individuals and, thus, repeatedly samples the same individuals (Buckland et al., 2001). We did not document any reactive movement of deer toward our aircraft at any altitude during test flights and collection of data. While we recommend surveying at low altitudes, we also encourage future users of aerial platforms to be aware of reactive movement induced by the aerial platform (Mulero-Pázmány et al., 2017). Such reactive movement can bias density estimates if not accounted for (Buckland et al., 2005; Glennie et al., 2015, 2021). Perhaps, the easiest way to avoid double counting on neighboring transects is to space transects far enough apart to completely avoid this error (e.g. Dunn et al., 2002). Regardless, we encourage future researchers to conduct test flights to evaluate reactive movement induced by the aerial vehicle.

In wildlife modeling, accounting for false positives has received considerably less attention than false negatives (Kéry & Royle, 2016, 2020; Strickfaden et al., 2020). Small amounts of false positives can substantially bias estimates (Miller et al., 2011). Because of this, modeling false-positive rates has increased in areas such as genetics (Augustine et al., 2020), acoustic monitoring (Chambert et al., 2018) and citizen science (Clare et al., 2019). Many of the strategies to count misidentification errors in IR aerial sampling mimic those used for counting perception errors. Several studies have counted misidentification errors associated with automated viewing software by manually examining the IR video (e.g. Chrétien et al., 2016; Lethbridge et al., 2019; Lhoest et al., 2015), but this strategy assumes that the manual examiner does not commit misidentification errors. Simultaneous ground surveys have been conducted to assess misidentification errors (Bushaw et al., 2020; Corcoran et al., 2019; Stander et al., 2021). Such ground surveys appear more promising if the target species is sedentary and can thus reasonably be assumed to not move in between flights and ground truthing. Geographic coordinates of all other objects or animals can then be compared to those of the target individuals, and used to count misidentification errors (Corcoran et al., 2019). Other researchers have reduced the altitude of an uncrewed aircraft (Bushaw et al., 2019; Smith et al., 2020), or circled a crewed aircraft (Gillette et al., 2015), to confirm IR heat signatures as the target species. Circling or altitude-reduction strategies may not be feasible for large-scale spatial sampling of

common species, as time expenditures would drastically increase. Our results suggest that the use of RGB video to confirm IR heat signatures as belonging to the target species is a promising method for reducing misidentification errors when possible. Unfortunately, our review found that supplemental RGB video has inconsistently been implemented by those relying on IR thermal sensors to sample wildlife, especially in crewed vehicles. We found strong evidence for the need to use RGB video to confirm that IR heat signatures are deer—19.3% ( $SE = 4.1$ ) of lenient IR detections that would otherwise have been ignored were confirmed as deer, and 18.3% ( $SE = 2.8$ ) of stringent IR detections that would otherwise have been counted were instead confirmed as not deer or unsure. Similar to previous research, we found the efficacy of RGB confirmation to be higher at lower altitudes (Milletto et al., 2011). Additionally, we found snow to be essential when using RGB video to confirm IR heat signatures, as white-tailed deer blend in well with bare or leaf-covered ground in forested habitat. Thus, our ability to distinguish deer from debris or ground features was poor, at least at 450 m altitude. Therefore, to minimize misidentification errors, we recommend using RGB video to confirm IR heat signatures and sampling during snow cover from flights at lower altitudes. Reliance on snow and low altitude for RGB confirmations may be less important for those sampling in open habitats, when the target species has natural color contrast with the background terrain, and for endotherms with larger body masses than white-tailed deer. In instances when snow cover is unavailable, uncrewed aircraft may be a suitable alternative to crewed aircraft, as these platforms can be flown at much lower altitudes than crewed aircraft and thus may not need to rely on snow or a double-sensor platform. Nonetheless, we encourage future work to quantify and account for misidentification errors in aerial sampling.

## Acknowledgements

We thank E. A. Flaherty and P. A. Zollner for helpful comments on the paper, and E. A. Rexstad and S. T. Buckland for suggestions on double-observer methods. B. Stirm engineered the pods used to secure cameras to the aircraft and installed and removed these pods and cameras from the pods. We gratefully acknowledge the traditional homelands of the Indigenous People upon which Purdue University is built and our fieldwork was conducted. We honor and appreciate the Bodéwadmik (Potawatomi), Lenape (Delaware), Myaamia (Miami) and Shawnee People who are the original Indigenous caretakers. This paper is a contribution of the Integrated Deer Management Project, a collaborative research effort between Purdue University and the Indiana Department

of Natural Resources Division of Fish and Wildlife. Funding was provided by the Indiana Department of Natural Resources Grant W-48-R-02.

## References

- Augustine, B.C., Royle, J.A., Linden, D.W. & Fuller, A.K. (2020) Spatial proximity moderates genotype uncertainty in genetic tagging studies. *Proceedings of the National Academy of Sciences*, **117**, 17903–17912.
- Barker, R. (2008) Theory and application of mark–recapture and related techniques to aerial surveys of wildlife. *Wildlife Research*, **35**, 268–274.
- Beaver, J.T., Baldwin, R.W., Messinger, M., Newbolt, C.H., Ditchkoff, S.S. & Silman, M.R. (2020) Evaluating the use of drones equipped with thermal sensors as an effective method for estimating wildlife. *Wildlife Society Bulletin*, **44**, 434–443.
- Borchers, D.L., Buckland, S.T., Goedhart, P.W., Clarke, E.D. & Hedley, S.L. (1998) Horvitz Thompson estimators for double-platform line transect surveys. *Biometrics*, **54**, 1221–1237.
- Brack, I.V., Kindel, A. & Oliveira, L.F.B. (2018) Detection errors in wildlife abundance estimates from unmanned aerial systems (UAS) surveys: synthesis, solutions, and challenges. *Methods in Ecology and Evolution*, **9**(8), 1864–1873.
- Brunton, E.A., Leon, J.X. & Burnett, S.E. (2020) Evaluating the efficacy and optimal deployment of thermal infrared and true-colour imaging when using drones for monitoring kangaroos. *Drones*, **4**, 20.
- Buckland, S.T., Anderson, D.R., Burnham, K.P. & Laake, J.L. (2005) Distance sampling. In P. Armitage & T. Colton (Eds.), *Encyclopedia of biostatistics* (2nd ed.). Chichester, UK: John Wiley & Sons.
- Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2001) *Introduction to distance sampling*. Oxford, UK: Oxford University Press.
- Buckland, S.T., Rexstad, E.A., Marques, T.A. & Oedekoven, C.S. (2015) *Distance sampling: methods and applications*. Heidelberg, Germany: Springer.
- Burn, D.M., Udevitz, M.S., Speckman, S.G. & Benter, R.B. (2009) An improved procedure for detection and enumeration of walrus signatures in airborne thermal imagery. *International Journal of Applied Earth Observation and Geoinformation*, **11**, 324–333.
- Burt, M.L., Borchers, D.L., Jenkins, K.J. & Marques, T.A. (2014) Using mark–recapture distance sampling methods on line transect surveys. *Methods in Ecology and Evolution*, **5**, 1180–1191.
- Bushaw, J.D., Ringelman, K.M., Johnson, M.K., Rohrer, T. & Rohwer, F.C. (2020) Applications of an unmanned aerial vehicle and thermal-imaging camera to study ducks nesting over water. *Journal of Field Ornithology*, **91**, 409–420.
- Bushaw, J.D., Ringelman, K.M. & Rohwer, F.C. (2019) Applications of unmanned aerial vehicles to survey mesocarnivores. *Drones*, **3**(1), 28.
- Chambert, T., Waddle, J.H., Miller, D.A., Walls, S.C. & Nichols, J.D. (2018) A new framework for analysing automated acoustic species detection data: occupancy estimation and optimization of recordings post-processing. *Methods in Ecology and Evolution*, **9**, 560–570.
- Chrétien, L.P., Théau, J. & Ménard, P. (2015) Wildlife multispecies remote sensing using visible and thermal infrared imagery acquired from an unmanned aerial vehicle (UAV). *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, **4**, 241–248.
- 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.
- Clare, J.D., Townsend, P.A., Anhalt-Depies, C., Locke, C., Stenglein, J.L., Frett, S. et al. (2019) Making inference with messy (citizen science) data: when are data accurate enough and how can they be improved? *Ecological Applications*, **29**, e01849.
- Conn, P.B., Chernook, V.I., Moreland, E.E., Trukhanova, I.S., Regehr, E.V., Vasiliev, A.N. et al. (2021) Aerial survey estimates of polar bears and their tracks in the Chukchi Sea. *PLoS One*, **16**, e0251130.
- Corcoran, E., Denman, S., Hanger, J., Wilson, B. & Hamilton, G. (2019) Automated detection of koalas using low-level aerial surveillance and machine learning. *Scientific Reports*, **9**, 1–9.
- Delisle, Z.J., Swihart, R.K., Quinby, B.M., Sample, R.D., Kinser-McBee, K.J., Tauber, E.K. et al. (2022) Density from pellet groups: comparing methods for estimating dung persistence time. *Wildlife Society Bulletin*, **46**, e1325.
- Dunn, W.C., Donnelly, J.P. & Krausmann, W.J. (2002) Using thermal infrared sensing to count elk in the southwestern United States. *Wildlife Society Bulletin (1973–2006)*, **30**(3), 963–967.
- Edwards, H.H., Hostetler, J.A., Stith, B.M. & Martin, J. (2021) Monitoring abundance of aggregated animals (Florida manatees) using an unmanned aerial system (UAS). *Scientific Reports*, **11**(1), 1–12.
- Fewster, R.M., Buckland, S.T., Burnham, K.P., Borchers, D.L., Jupp, P.E., Laake, J.L. et al. (2009) Estimating the encounter rate variance in distance sampling. *Biometrics*, **65**, 225–236.
- Fewster, R.M. & Pople, A.R. (2008) A comparison of mark–recapture distance-sampling methods applied to aerial surveys of eastern grey kangaroos. *Wildlife Research*, **35**, 320–330.
- Franke, U., Goll, B., Hohmann, U. & Heurich, M. (2012) Aerial ungulate surveys with a combination of infrared and high-resolution natural color images. *Animal Biodiversity and Conservation*, **35**, 285–293.

- Gentle, M., Finch, N., Speed, J. & Pople, A. (2018) A comparison of unmanned aerial vehicles (drones) and manned helicopters for monitoring macropod populations. *Wildlife Research*, **45**, 586–594.
- Gillette, G.L., Reese, K.P., Connelly, J.W., Colt, C.J. & Knetter, J.M. (2015) Evaluating the potential of aerial infrared as a lek count method for prairie grouse. *Journal of Fish and Wildlife Management*, **6**(2), 486–497.
- Glennie, R., Buckland, S.T., Langrock, R., Gerrodette, T., Ballance, L.T., Chivers, S.J. et al. (2021) Incorporating animal movement into distance sampling. *Journal of the American Statistical Association*, **116**(533), 107–115.
- Glennie, R., Buckland, S.T. & Thomas, L. (2015) The effect of animal movement on line transect estimates of abundance. *PLoS One*, **10**(3), e0121333.
- Havens, K.J. & Sharp, E. (2015) *Thermal imaging techniques to survey and monitor animals in the wild: a methodology*. London, UK: Academic Press.
- Hone, J. (2008) On bias, precision and accuracy in wildlife aerial surveys. *Wildlife Research*, **35**, 253–257.
- Innes, S., Heide-Jørgensen, M.P., Laake, J.L., Laidre, K.L., Cleator, H.J., Richard, P. et al. (2002) Surveys of belugas and narwhals in the Canadian high Arctic in 1996. *NAMMCO Scientific Publications*, **4**, 169–190.
- Kays, R., Sheppard, J., Mclean, K., Welch, C., Paunescu, C., Wang, V. et al. (2019) Hot monkey, cold reality: surveying rainforest canopy mammals using drone-mounted thermal infrared sensors. *International Journal of Remote Sensing*, **40**, 407–419.
- Kéry, M. & Royle, J.A. (2016) *Applied hierarchical modeling in ecology: analysis of distribution, abundance and species richness in R and BUGS. Volume 1 – prelude and static models*. London, UK: Academic Press.
- Kéry, M. & Royle, J.A. (2020) *Applied hierarchical modeling in ecology: analysis of distribution, abundance and species richness in R and BUGS: volume 2: dynamic and advanced models*. London, UK: Academic Press.
- Kissell, R.E., Jr. & Tappe, P.A. (2004) Assessment of thermal infrared detection rates using white-tailed deer surrogates. *Journal of the Arkansas Academy of Science*, **58**, 70–73.
- Kissell, R.E. & Nimmo, S.K. (2011) A technique to estimate white-tailed deer *Odocoileus virginianus* density using vertical-looking infrared imagery. *Wildlife Biology*, **17**, 85–92.
- Laake, J., Borchers, D., Thomas, L., Miller, D. & Bishop, J. (2022) R package ‘mrds’: mark-recapture distance sampling. Version 2.2.6.
- Laake, J.L. & Borchers, D.L. (2004) Methods for incomplete detection at distance zero. In: Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (Eds.) *Advanced distance sampling*. Oxford, UK: Oxford University Press.
- Latham, A.D.M., Warburton, B., Latham, M.C., Anderson, D.P., Howard, S.W. & Binny, R.N. (2021) Detection probabilities and surveillance sensitivities for managing an invasive mammalian herbivore. *Ecosphere*, **12**, e03772.
- Lee, W.Y., Park, M. & Hyun, C.U. (2019) Detection of two Arctic birds in Greenland and an endangered bird in Korea using RGB and thermal cameras with an unmanned aerial vehicle (UAV). *PLoS One*, **14**(9), e0222088.
- Lethbridge, M., Stead, M. & Wells, C. (2019) Estimating kangaroo density by aerial survey: a comparison of thermal cameras with human observers. *Wildlife Research*, **46**, 639–648.
- Lhoest, S., Linchant, J., Quevauvillers, S., Vermeulen, C. & Lejeune, P. (2015) How many hippos (HOMHIP): algorithm for automatic counts of animals with infra-red thermal imagery from UAV. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **40**, 355–362.
- Linchant, J., Lisein, J., Semeki, J., Lejeune, P. & Vermeulen, C. (2015) Are unmanned aircraft systems (UAS s) the future of wildlife monitoring? A review of accomplishments and challenges. *Mammal Review*, **45**, 239–252.
- Lu, X., Hooten, M.B., Kaplan, A., Womble, J.N. & Bower, M.R. (2022) Improving wildlife population inference using aerial imagery and entity resolution. *Journal of Agricultural, Biological and Environmental Statistics*, **27**(2), 364–381.
- McKellar, A.E., Shephard, N.G. & Chabot, D. (2021) Dual visible-thermal camera approach facilitates drone surveys of colonial marshbirds. *Remote Sensing in Ecology and Conservation*, **7**, 214–226.
- McMahon, M.C., Dittmer, M.A. & Forester, J.D. (2021) Comparing unmanned aerial systems with conventional methodology for surveying a wild white-tailed deer population. *Wildlife Research*, **49**, 54–65.
- Miller, D.A., Nichols, J.D., McClintock, B.T., Grant, E.H.C., Bailey, L.L. & Weir, L.A. (2011) Improving occupancy estimation when two types of observational error occur: non-detection and species misidentification. *Ecology*, **92**, 1422–1428.
- Millette, T.L., Slaymaker, D., Marciano, E., Alexander, C. & Richardson, L. (2011) AIMS-thermal-A thermal and high resolution color camera system integrated with GIS for aerial moose and deer census in northeastern Vermont. *Aces: A Journal Devoted to the Biology and Management of Moose*, **47**, 27–37.
- Mulero-Pázmány, M., Jenni-Eiermann, S., Strebel, N., Sattler, T., Negro, J.J. & Tablado, Z. (2017) Unmanned aircraft systems as a new source of disturbance for wildlife: a systematic review. *PLoS One*, **12**(6), e0178448.
- Nichols, J.D., Hines, J.E., Sauer, J.R., Fallon, F.W., Fallon, J.E. & Heglund, P.J. (2000) A double-observer approach for estimating detection probability and abundance from point counts. *The Auk*, **117**, 393–408.
- Otis, D.L., Burnham, K.P., White, G.C. & Anderson, D.R. (1978) Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, **62**, 3–135.

- Pollock, K.H. & Kendall, W.L. (1987) Visibility bias in aerial surveys: a review of estimation procedures. *The Journal of Wildlife Management*, **51**, 502–510.
- Preston, T.M., Wildhaber, M.L., Green, N.S., Albers, J.L. & DeBenedetto, G.P. (2021) Enumerating White-tailed deer using unmanned aerial vehicles. *Wildlife Society Bulletin*, **45**, 97–108.
- R Core Team. (2022) *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Royle, J.A. & Link, W.A. (2006) Generalized site occupancy models allowing for false positive and false negative errors. *Ecology*, **87**, 835–841.
- Sasse, D.B. (2003) Job-related mortality of wildlife workers in the United States, 1937–2000. *Wildlife Society Bulletin*, **31**, 1015–1020.
- Schoenecker, K.A., Doherty, P.F., Jr., Hourt, J.S. & Romero, J.P. (2018) Testing infrared camera surveys and distance analyses to estimate feral horse abundance in a known population. *Wildlife Society Bulletin*, **42**, 452–459.
- Smith, T.S., Amstrup, S.C., Kirschhoffer, B.J. & York, G. (2020) Efficacy of aerial forward-looking infrared surveys for detecting polar bear maternal dens. *PLoS One*, **15**, e0222744.
- Stander, R., Walker, D.J., Rohwer, F.C. & Baydack, R.K. (2021) Drone nest searching applications using a thermal camera. *Wildlife Society Bulletin*, **45**, 371–382.
- Storm, D.J., Samuel, M.D., Van Deelen, T.R., Malcolm, K.D., Rolley, R.E., Frost, N.A. et al. (2011) Comparison of visual-based helicopter and fixed-wing forward-looking infrared surveys for counting white-tailed deer *Odocoileus virginianus*. *Wildlife Biology*, **17**, 431–440.
- Strickfaden, K.M., Fagre, D.A., Golding, J.D., Harrington, A.H., Reintsma, K.M., Tack, J.D. et al. (2020) Dependent double-observer method reduces false-positive errors in auditory avian survey data. *Ecological Applications*, **30**, e02026.
- Sudholz, A., Denman, S., Pople, A., Brennan, M., Amos, M. & Hamilton, G. (2021) A comparison of manual and automated detection of rusa deer (*Rusa timorensis*) from RPAS-derived thermal imagery. *Wildlife Research*, **49**(1), 46–53.
- Swihart, R.K., Caudell, J.N., Brooke, J.M. & Ma, Z. (2020) A flexible model-based approach to delineate wildlife management units. *Wildlife Society Bulletin*, **44**, 77–85.
- Witt, R.R., Beranek, C.T., Howell, L.G., Ryan, S.A., Clulow, J., Jordan, N.R. et al. (2020) Real-time drone derived thermal imagery outperforms traditional survey methods for an arboreal forest mammal. *PLoS One*, **15**, e0242204.

## Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Table S1.** AIC-based model selection results for fit detection functions. The top four AIC-best detection functions are shown for each altitude (m), observer and type of infrared detections used (lenient = Len, stringent = Strin).

**Table S2.** The number of articles using infrared thermal sensors from aerial platforms that addressed different error types. Error types considered were detection probability across the field-of-view of the sensor (Uniform), availability errors (AE), perception errors (PE), misclassification errors (ME) and double counting errors (Double).

**Figure S1.** Plots of the AIC-best detection function for each altitude (m), observer and type of infrared detections used (lenient = Len, stringent = Strin).

**Figure S2.** The number of articles published annually using aerial sampling techniques combined with infrared thermography to monitor wildlife.