

RESEARCH ARTICLE

Estimating animal size or distance in camera trap images: Photogrammetry using the pinhole camera model

Scott Leorna¹  | Todd Brinkman¹  | Timothy Fullman² 

¹University of Alaska Fairbanks, Institute of Arctic Biology, Fairbanks, AK, USA

²The Wilderness Society, Anchorage, AK, USA

Correspondence

Scott Leorna

Email: sleorna@alaska.edu

Funding information

Office of Polar Programs, Grant/Award

Number: 1839192

Handling Editor: Patrick Jansen

Abstract

1. As camera trapping has become a standard practice in wildlife ecology, developing techniques to extract additional information from images will increase the utility of generated data. Despite rapid advancements in camera trapping practices, methods for estimating animal size or distance from the camera using captured images have not been standardized. Deriving animal sizes directly from images creates opportunities to collect wildlife metrics such as growth rates or changes in body condition. Distances to animals may be used to quantify important aspects of sampling design such as the effective area sampled or distribution of animals in the camera's field-of-view.
2. We present a method of using pixel measurements in an image to estimate animal size or distance from the camera using a conceptual model in photogrammetry known as the 'pinhole camera model'. We evaluated the performance of this approach both using stationary three-dimensional animal targets and in a field setting using live captive reindeer *Rangifer tarandus* ranging in size and distance from the camera.
3. We found total mean relative error of estimated animal sizes or distances from the cameras in our simulation was -3.0% and 3.3% and in our field setting was -8.6% and 10.5% , respectively. In our simulation, mean relative error of size or distance estimates were not statistically different between image settings within camera models, between camera models or between the measured dimension used in calculations.
4. We provide recommendations for applying the pinhole camera model in a wildlife camera trapping context. Our approach of using the pinhole camera model to estimate animal size or distance from the camera produced robust estimates using a single image while remaining easy to implement and generalizable to different camera trap models and installations, thus enhancing its utility for a variety of camera trap applications and expanding opportunities to use camera trap images in novel ways.

KEYWORDS

depth estimation, effective area sampled, effective detection distance, monocular camera, morphometrics, trail camera

1 | INTRODUCTION

The use of time-lapse and remotely triggered cameras has become a widely popular method of non-invasively collecting information on wildlife and is often referred to as camera trapping in wildlife ecology (O'Connell et al., 2011; Rowcliffe & Carbone, 2008). With camera trap usage rapidly expanding in recent decades, researchers continue to explore and test new methods of data collection and analysis to answer novel ecological questions using this technology (Sollmann, 2018; Trollet et al., 2014). The main type of data collected from camera trap images is often animal presence or absence in an image. This information is typically used for assessments of species diversity, richness, distribution or habitat use (Burton et al., 2015). Other common types of data collected from camera trap images include animal behaviour (e.g. foraging, sleeping, mating), demographics (e.g. sex, age class) or environmental characteristics (e.g. weather, phenology; Caravaggi et al., 2017; Hofmeester et al., 2020). Statistical methods have been developed to use detection–non-detection data to estimate the absolute abundance and density of unmarked animals; however, these methods require estimates of the area sampled and animal movement rates (Gilbert et al., 2020; Moeller et al., 2018; Rowcliffe et al., 2008). Though these inputs could be potentially acquired from other data sources (e.g. using telemetry), several methods have been used to derive them directly from camera trap images (e.g. Rowcliffe et al., 2011, 2016). Evaluating techniques and expanding the types of data that can be collected from camera trap images enhance the potential to use camera trapping in novel ways and can help inform standardized methods, maximizing the value of data generated (Scotson et al., 2017; Thomson et al., 2018).

Several methods have been described in the camera trap literature for estimating animal size or distance to the camera directly from captured images (Berger, 2012; Caravaggi et al., 2016; Cui et al., 2020; Hofmeester et al., 2017; Rowcliffe et al., 2011; Tarugara et al., 2019; Willisch et al., 2013; Xu et al., 2020). These data create the potential to complement and/or supplement existing camera trap applications as well as answer novel ecological questions. For example, estimating various size dimensions of wildlife in images can be used for monitoring physiological health, assessing growth rates, estimating trophy potential, distinguishing between individuals, or in assessments of demographics where size is an indication of sex or age class (e.g. Meise et al., 2014; Tarugara et al., 2019; Willisch et al., 2013; Zhang et al., 2018; Zheng et al., 2016). Additionally, estimated distances to wildlife in images can be used to approximate the detection range of camera motion sensors or object detection models used to process images, quantify sampling areas, derive movement rates from consecutive images or to support distance sampling approaches (e.g. Corlatti et al., 2020; Hofmeester et al., 2017; Howe et al., 2017; Rowcliffe et al., 2008, 2011, 2016). While many techniques have been developed to extract pertinent information from camera trap images, they often arise to address specific project objectives under specific conditions which may limit their generalizability and hinder comparisons among studies (Thomson et al., 2018; Young et al., 2018). For example, many studies in the camera trap

literature describing methods for estimating animal size or distance from the camera have often focused on enumerating one or the other (e.g. Hofmeester et al., 2017; Tarugara et al., 2019) and few studies have used a standardized method (e.g. Rowcliffe et al., 2011, 2016). Based on the diversity of camera trap models, field applications, and project goals and objectives, identifying a generalizable method for determining animal size or distance from the camera may contribute to more efficient data collection and help establish a standardized method to facilitate study comparisons.

Our main goal was to identify and describe a simple and generalizable method of deriving either animal size or distance from the camera directly from camera trap images to support growing analytical approaches and camera trap applications. Broadly, extracting information about the physical three-dimensional scene captured in two dimensions of an image is often referred to as the field of photogrammetry (Aleixo et al., 2020; Kannala et al., 2008). We applied a conceptual model in photogrammetry known as the 'pinhole camera model' to estimate the physical size of objects at known distances or distances to objects of known size using pixel measurements taken directly from an image (Kannala et al., 2008; Megalingam et al., 2016). Specifically, our objectives were to (a) determine whether the pinhole camera model can produce reliable estimates of animal size or distance using different camera traps, models and image settings under ideal conditions using stationary life-like animal targets and (b) evaluate the performance of the pinhole camera model in a practical field setting using live captive reindeer *Rangifer tarandus*. We compare the pinhole camera model approach with other methods published in the camera trap literature and provide recommendations on how to maximize its utility for wildlife camera trap projects. Our application of the pinhole camera model can be easily implemented in a variety of camera trap applications and provides information that may be used to complement and/or supplement detection–non-detection datasets.

2 | MATERIALS AND METHODS

2.1 | Pinhole camera model

The pinhole camera model (Figure 1) relates a three-dimensional scene to the two-dimensional image captured on the camera sensor using similar triangles where light from a point in the physical scene passes through the camera (i.e. pinhole or aperture) and is transferred to a point on the camera sensor (Kannala et al., 2008; Megalingam et al., 2016). The relationship between a physical object and its projection on an image is summarized by the equation,

$$\frac{S_i}{d_i} = \frac{S_o}{d_o},$$

where S_i is the size of the object on the image (e.g. in pixels), d_i is the distance from the camera sensor to the aperture (i.e. focal length expressed in pixels), S_o is the physical size (e.g. height, width in meters) of

FIGURE 1 Conceptual model and equations derived from the 'pinhole camera model' used to estimate physical dimensions of objects based on their measured pixel dimension in an image

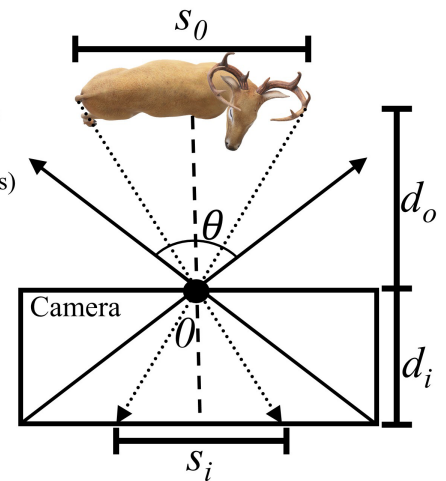
O = Optical center (pinhole or aperture)
 θ = Field-of-View (FOV)
 d_o = Distance to object (physical dimension)
 S_o = Size of object (physical dimension)
 d_i = Distance to image (focal length in pixels)
 S_i = Size of object on image (in pixels)

$$\text{eq.1} \quad S_o = \frac{S_i * d_o}{d_i}$$

$$\text{eq.2} \quad d_o = \frac{S_o * d_i}{S_i}$$

$$\text{eq.3} \quad S_i = \frac{S_o * d_i}{d_o}$$

$$\text{eq.4} \quad d_i = \frac{S_i * d_o}{S_o}$$



the object and d_o is the physical distance (e.g. meters) from the camera aperture to the object. As long as three out of the four inputs are known or assumed, this equation can be rearranged to solve for the physical size of an object (Figure 1, Equation 1), the physical distance to an object from the camera (Figure 1, Equation 2), the size of the object on the image (Figure 1, Equation 3) or the camera's focal length (Figure 1, Equation 4).

The first step to applying the pinhole camera model to estimate animal size or distance is identifying the focal length of the camera used to capture images. Since camera trap manufacturers typically do not report a camera's focal length (e.g. in mm), we used a calibration procedure to derive an estimate of each camera's focal length expressed in pixels (Megalingam et al., 2016). Deriving the camera focal length in pixels rather than millimetres conveniently results in the units in Equations 1 and 2 cancelling out with the measured pixel size of the object in the image, resulting in the calculated value having physical units (e.g. meters) of the inputted distance to the object from the camera (Equation 1) or size of the object (Equation 2).

2.2 | Camera calibration

We estimated camera focal lengths expressed in pixels of lower cost (Campark T20 ~\$40 USD) and higher cost (Reconyx Hyperfire 2 HF2X ~\$400 USD) commercially available camera traps, each with two different image settings. To do so, we collected five images of a 0.25x0.25m piece of white paper on a black poster board at 1 m intervals from the camera up to 5 m; however, any object of known size could have been used. Distances were measured using a tape measure. We captured the white paper near the centre of the image and approximately perpendicular to the face of the camera. We measured the height of the paper in pixels using the straight-line tool in the open-source software ImageJ (Schneider et al., 2012). Then, following Equation 4, we estimated each camera's focal length expressed in pixels (Table 1). To determine whether the derived focal length varied between cameras of the same make, model and image setting, we followed this calibration procedure with three separate cameras for each camera model and image setting and used a

one-way ANOVA to compare mean focal lengths between similar cameras (Table 1). We felt comparing derived focal lengths between similar cameras was important to consider since it would have consequences on whether each individual camera used in a study would have to be calibrated or whether a single derived focal length could be safely applied to all cameras of identical make, model and image setting in a study (i.e. reducing effort in applying the pinhole camera method to a large number of cameras).

2.3 | Estimating physical dimensions using pixel measurements

To determine whether estimates of animal size or distance under ideal and controlled conditions were influenced by camera model or image settings (Objective 1), we used five life-like three-dimensional animal targets ranging in size, shape and distance from the camera (Table 2). Using stationary fixed objects, we could ensure that accurate and precise measurements of animal size and distance to the camera were recorded (i.e. minimizing measurement error). One camera for each camera model and image setting was used to collect one image of each animal at distances of 5, 15 and 25m. This distance range represents a typical motion sensor detection range of most wildlife camera traps (Trollet et al., 2014). Images were taken with animals approximately perpendicular to the camera face and in the centre of the image. We recorded the distance, physical nose-to-tail length (contour length along back from tip of nose to base of tail) and shoulder height (ground to maximal height of front shoulder) for each animal using a surveyor's measuring tape. Images were reviewed in ImageJ and the segmented line tool was used to measure nose-to-tail length and the straight-line tool was used to measure shoulder height in pixels. We note that the dimensions of the animal targets do not necessarily represent the actual size or proportions of the represented animal species. However, our intent was to include objects varying in size and shape, and these targets were used to replicate realistic animal shapes and contours in a natural outdoor setting. We used each known distance from the camera to estimate each animal's nose-to-tail length and shoulder height

TABLE 1 Estimated camera trap and image setting focal lengths

Camera model	Image setting	Image resolution	Camera ID	Derived focal length d_i (px)				ANOVA		
				n	Mean	95% CI		df	F	p
						LL	UL			
Campark T20	3mp	2304 × 1296	1	5	1745.2	1735.0	1755.4	2	2.65	0.11
			2	5	1756.7	1743.4	1770.0			
			3	5	1757.2	1746.1	1768.4			
			Total	15	1753.1	1747.3	1758.8			
	16mp	5376 × 3024	1	5	4095.5	4073.4	4117.5	2	1.02	0.39
			2	5	4087.5	4068.1	4106.9			
			3	5	4103.9	4078.2	4129.6			
			Total	15	4095.6	4085.6	4105.7			
Reconyx HF2X	Wide Angle	2048 × 1152	1	5	2923.3	2908.0	2938.6	2	0.90	0.43
			2	5	2910.1	2888.2	2931.9			
			3	5	2916.3	2895.7	2936.8			
			Total	15	2916.6	2907.9	2925.2			
	Standard	2048 × 1440	1	5	2931.9	2926.9	2936.9	2	0.15	0.86
			2	5	2931.9	2919.7	2944.1			
			3	5	2934.1	2925.9	2942.4			
			Total	15	2932.7	2928.9	2936.4			

Note: Estimated focal lengths are based on a calibration procedure using an object of known size (e.g. 0.25 × 0.25 m piece of paper) at a known distance (e.g. 1–5 m) from the camera and are expressed in pixels. For each camera model and image setting, camera ID 1 was used in our field test and the associated focal length was used in calculations.

(Figure 1, Equation 1) and each animal's known nose-to-tail length and shoulder height to estimate their distance from the camera at each distance range (Figure 1, Equation 2). The specific focal length of the camera used to collect images was used in calculations (i.e. camera ID 1, Table 1).

2.4 | Captive reindeer field test

To evaluate our application of the pinhole camera model and its performance in estimating animal size or distance from the camera under more realistic field study conditions (Objective 2), we recorded images of captive reindeer at the Large Animal Research Station (LARS) at the University of Alaska Fairbanks in Alaska, USA (IACUC #1370779-1). With permission from LARS, six individuals (4 females, 2 males) were photographed using a Reconyx Hyperfire 2 HF2X camera with standard image setting (i.e. 2048 × 1440 image resolution) at various distances (ranging between 5.4 and 96.0 m) in various positions and locations in the image (Figure 2b; Table 3). For this field test, the specific camera we used was different from those we estimated focal lengths for in our camera calibrations. Therefore, we used the mean focal length we derived for this camera model and image setting (i.e. 2932.7px, Table 1) in calculations. For each individual reindeer, morphological dimensions consistent with those available in published literature (e.g. Klein et al., 1987; Nieminen & Helle, 1980) were measured by LARS staff

TABLE 2 Physical size of animal targets used for simulation

Animal	Nose-to-tail length (cm)	Shoulder height (cm)
Pig	109.7	54.4
Impala	149.3	63.6
Axis deer	186.0	79.4
White-tailed deer	201.2	104.5
Moose	208.5	85.9

Note: Actual sizes (i.e. nose-to-tail length and shoulder height) do not necessarily accurately reflect typical sizes or proportions of the represented animal species. For example, the shoulder height of the white-tailed deer was greater than the moose.

using a cloth tape measure. These included hind-leg length, fore-leg length, shoulder height, back length and nose-to-tail length (Figure 2a; Table 3). For each animal captured in images, we subjectively chose and measured the pixel lengths of morphological dimensions we felt were mostly unobstructed and least affected by their orientation to the camera (e.g. shoulder height would be used rather than back length if the animal was severely angled towards or away from the camera, Figure 3) using ImageJ. However, for all animals captured in images, at least one morphological dimension was measured in pixels to ensure we included at least one distance estimate for each animal (i.e. no images were discarded if there were no perfectly visible morphological dimensions; in such

TABLE 3 Physical size of individual captive reindeer *Rangifer tarandus* used for field test and the minimum and maximum distance at which they were photographed

Reindeer ID	Sex	Morphometric (cm)					Distance (m)	
		Hind-leg length	Fore-leg length	Shoulder height	Back length	Total length	Min	Max
580	F	45.0	55.0	105.5	87.8	171.5	7.0	77.0
854	F	41.5	55.5	102.0	80.5	161.5	12.0	28.0
273	F	43.7	55.0	100.0	86.0	183.5	10.0	47.0
003	F	45.0	60.0	104.0	86.6	159.5	17.0	96.0
608	M	45.0	58.0	115.5	98.1	187.0	17.5	44.0
859	M	44.5	66.0	117.0	98.5	184.0	5.4	26.6

cases, we subjectively chose the best one available). Due to logistical constraints and safety concerns, we were unable to enter the pastures where animals were located to set up distance markers and there were no physical landmarks (e.g. trees, rocks) that were easily distinguishable that we could use as distance references. Therefore, distances to animals at the time of photo capture were recorded using a laser rangefinder (Vortex Ranger 1300). We made no attempts to capture animals in any particular orientation (i.e. angle to the camera) or position in the image (e.g. centred or at the edges of the image; e.g. Figure 2b). We did, however, not record images when animals were bedded (i.e. laying down) which would prevent being able to measure all our available morphometrics (also, bedded animals would be unlikely to trigger the motion sensor if it were being used as is the case for many camera trapping projects). Given these constraints, we did our best to generate an image set as representative as possible of what would be expected with animals moving in the camera field-of-view in a natural and uncontrolled way, while still collecting the necessary data to apply and evaluate the pinhole camera method. Estimates of animal size (i.e. morphological dimensions) or distance from the camera were calculated using Equations 1 and 2, respectively (Figure 1).

2.5 | Data analysis

To evaluate the pinhole camera model's accuracy for estimating animal size or distance from the camera for both our simulated and field application, we calculated the percent relative error (*RE*) for each estimate, a metric commonly used in similar studies to evaluate measurement accuracy (e.g. Berger, 2012; Cui et al., 2020; Meise et al., 2014; Willisch et al., 2013). However, rather than using the absolute difference between the true and estimated value in the *RE* calculation, we subtracted the true value from the estimated value resulting in negative (i.e. representing an underestimate) or positive (i.e. representing an overestimate) *RE* values (i.e. $RE = ([Estimated - Actual]/Actual) \times 100$). For example, if the true value of a dimension were 10, estimated values of 9 and 11 would correspond with a *RE* of -10% and 10% , respectively. Since *RE* provides a percent error relative to the true value, the mean *RE* is intended to be a standardized

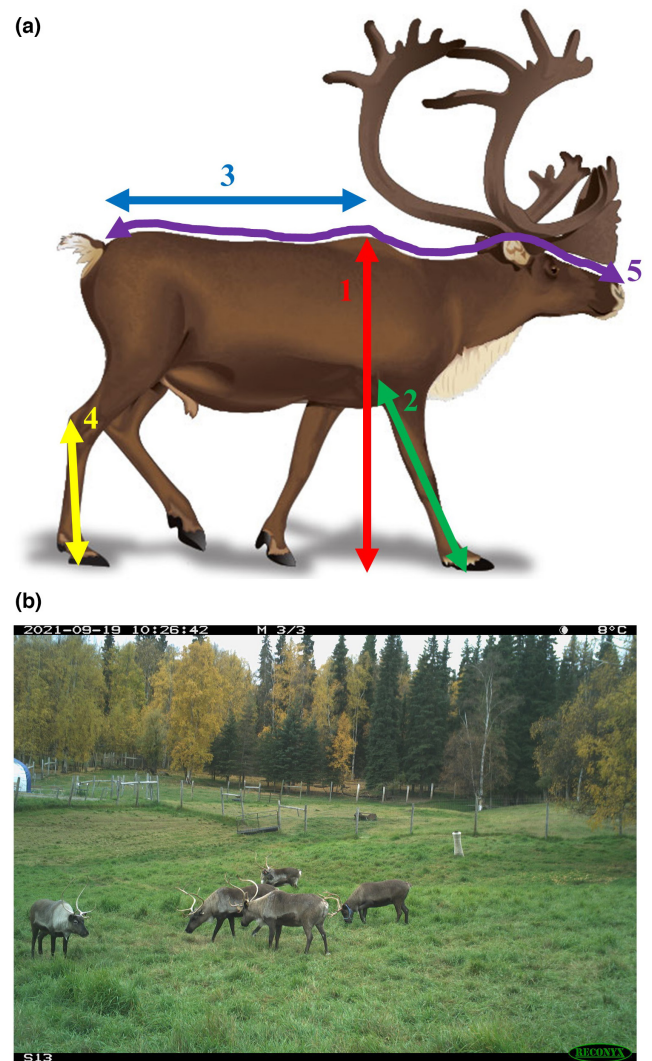


FIGURE 2 (a) Reindeer morphological dimensions measured in field test including (1) shoulder height (ground to maximal height of front shoulder), (2) fore-leg length (ground to top of radio-ulna), (3) back length (top of front shoulder to base of tail), (4) hind-leg length (ground to calcaneum) (5) and nose-to-tail length (contour length along back from tip of nose to base of tail), and (b) example image with reindeer captured at various distances, locations in the image and positions

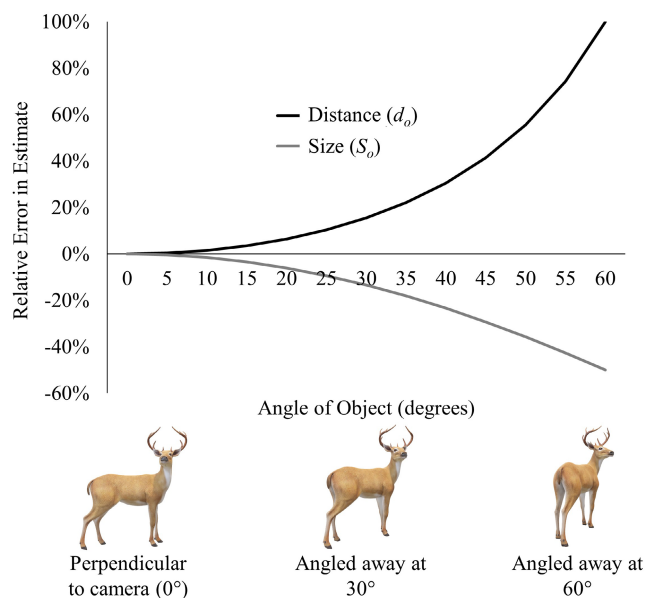


FIGURE 3 Theoretical errors in estimated size or distance of an object when using the pinhole camera model associated with object orientation to the camera (i.e. perspective distortions). The apparent size (i.e. what would be measured in the image and represents S_i in Figure 1 equations) of an object can be calculated as the actual size of the object multiplied by cosine of its angle (in degrees) to the camera. Plotted relative errors for estimated physical size (S_o) and distance (d_o) are based on substituting the apparent size for the actual size in calculations and do not account for any other sources of error (e.g. optical distortions, measurement error). The deer example provided is based on a measured dimension in the x-axis (e.g. animal length) since the animal is being rotated on the y-axis

estimate applicable to any range of animal sizes or distances from the camera. The 95% confidence interval (CI) of the mean RE is reported as a measure of precision.

In our simulation dataset, for each camera model and image setting, there were a total of 15 images from which animal size or distance from the camera were estimated (i.e. five different animals captured at three different distances). To determine whether the accuracy of estimated animal size or distance from the camera varied between image settings within camera models, between camera models or between measured dimensions used in calculations (i.e. nose-to-tail length or shoulder height measured in pixels), we used independent samples *t*-tests and report *t*-statistics and *p*-values for each comparison. For all tests, a significance level of $\alpha = 0.05$ was used. To report an overall measure of accuracy for each estimated physical dimension, we pooled all data (i.e. 120 unique measurements taken from images) and report the pooled mean RE and 95% CI. For our field test dataset, we pooled estimates of animal sizes and distances to the camera based on the selected morphological dimension measured in the image and report the mean RE and 95% CI for estimates based on each measured dimension as well as pooled across dimensions as an overall summary of performance.

3 | RESULTS

3.1 | Camera calibration

Combining images from all three cameras of the same model and image setting, the total pooled mean derived focal length for the Campark T20 camera with a 3mp image setting was 1753.1px ($n = 15$; 95% CI: 1747.3, 1758.8) and for the 16mp setting was 4095.6px ($n = 15$; 95% CI: 4085.6, 4105.7). The total pooled mean derived focal length for the Reconyx HF2X camera with a wide-angle image setting was 2916.6px ($n = 15$; 95% CI: 2907.9, 2925.2) and for the standard image setting was 2932.7px ($n = 15$; 95% CI: 2928.9, 2936.4; Table 1). For each camera model and image setting, we found no statistical differences in the derived focal lengths among the three cameras tested (all $p > 0.05$; Table 1).

3.2 | Three-dimensional animal target simulation

Pooling all physical dimension estimates among camera models and image settings, the total mean RE for size estimates was -3.0% ($n = 120$, 95% CI: -3.7 , -2.2) and for distance estimates was 3.3% ($n = 120$, 95% CI: 2.4 , 4.1 ; Table 4). Individual animal size and distance estimates from each camera model and image setting are provided in Tables S1 and S2. There were no statistically significant differences (all $p > 0.05$) in the mean RE of size or distance estimates between image settings within camera models, between camera models or between the measured dimension used in calculations (Table 4).

3.3 | Captive reindeer field test

Pooling all physical dimension estimates among morphological dimensions measured in images, the mean RE in estimated sizes was -8.6% ($n = 153$, 95% CI: -10.0 , -7.1) and for distance estimates was 10.5% ($n = 153$, 95% CI: 8.7 , 12.2 ; Table 5), representing a decrease in RE of -5.6% for size estimates and increase in RE of 7.2% for distance estimates from our animal target simulation (Tables 4 and 5). All estimates of individual morphological dimensions for each reindeer and individual distances are provided in Tables S3 and S4.

4 | DISCUSSION

Determining animal size or distance from the camera directly from camera trap images expands opportunities to use camera trapping in novel ways and complements advancements in statistical methods for analysis of detection–non-detection data (Gilbert et al., 2020; Moeller et al., 2018; Rowcliffe et al., 2008, 2011). We applied and evaluated a method based on the pinhole camera model which uses pixel measurements of wildlife in an image to estimate the size of wildlife at known distances or distance to wildlife of known sizes (Figure 1; Kannala et al., 2008; Megalingam et al., 2016). In our

TABLE 4 Summary of results to evaluate performance of the pinhole camera model in controlled simulation

Physical dimension being estimated	Measured dimension used for calculation	Camera model	Image setting	n	Relative error			t-test for image resolution			t-test for camera model			t-test for measured dimension		
					Mean	LL	UL	df	t	p	df	t	p	df	t	p
					95% CI											
Distance (d_0)	Nose-to-tail length	Campark T20	3mp	15	3.9	1.0	6.8	28	0.22	0.82	58	0.07	0.95	118	0.88	0.38
			16mp	15	3.4	0.4	6.5									
	Shoulder height	Reconyx HF2X	Wide angle	15	3.5	1.5	5.6	28	-0.08	0.93						
			Standard	15	3.6	1.6	5.7									
		Campark T20	3mp	15	1.7	-1.3	4.6	28	-1.04	0.31	58	-0.54	0.59			
			16mp	15	3.5	1.2	5.7									
Size (S_0)	Nose-to-tail length	Campark T20	3mp	15	-3.5	-6.2	-0.8	28	-0.24	0.81	58	0.05	0.96	118	-0.91	0.37
			16mp	15	-3.1	-5.9	-0.2									
	Shoulder height	Reconyx HF2X	Wide angle	15	-3.3	-5.2	-1.4	28	0.08	0.93						
			Standard	15	-3.4	-5.3	-1.5									
		Campark T20	3mp	15	-1.4	-4.1	1.3	28	1.13	0.27	58	0.58	0.56			
			16mp	15	-3.2	-5.3	-1.1									
	Reconyx HF2X	Wide angle	15	-3.2	-5.9	-0.5	28	-0.36	0.72							
		Standard	15	-2.7	-4.4	-0.9										

TABLE 5 Summary of results from captive reindeer *Rangifer tarandus* field test

Measured dimension used for calculation	n	Relative error in estimated distance			Relative error in estimated size		
		Mean	95% CI		Mean	95% CI	
			LL	UL		LL	UL
Hind-leg length	35	16.1%	12.8%	19.4%	-13.3%	-15.7%	-10.9%
Fore-leg length	29	15.3%	12.3%	18.3%	-12.8%	-15.1%	-10.6%
Shoulder height	41	11.0%	8.9%	13.1%	-9.6%	-11.2%	-7.9%
Back length	28	8.3%	3.5%	13.1%	-6.6%	-10.4%	-2.7%
Total length	20	-4.5%	-8.1%	-0.9%	5.3%	1.8%	8.8%
Total	153	10.5%	8.7%	12.2%	-8.6%	-10.0%	-7.1%

controlled three-dimensional animal target simulation, we found the mean RE of estimated animal sizes and distances from the camera (-3.0% and 3.3%, respectively; Table 4) were comparable with other studies in which the physical specifications of the cameras were known and careful image selection was used (Berger, 2012; Meise et al., 2014), with only slightly increased error in our more practical field assessment using captive reindeer (mean RE of estimated animal sizes and distances from the camera were -8.6% and 10.5%, respectively; Table 5). We suspect the increased error observed in our reindeer field test was largely a result of increased human error in measuring actual animal size and distance to the camera (i.e. parameter inputs assumed to be true) because measuring morphological dimensions of stationary animal targets and their distance to the camera was much easier than taking equally accurate and precise measurements on live reindeer. Additionally, there was likely increased instrument error in measuring actual distances because we used a surveyor's measuring tape (more accurate/precise) in our simulation as opposed to a rangefinder (less accurate/precise) in our field test. Also, vegetation often obstructed where the animal's hooves made contact with the ground in our field test but not in our simulation (Figure 2b). This forced us to make our best guess at where this point was when measuring the pixel length of vertical morphological dimensions used in the field test (i.e. hind-leg length, fore-leg length and shoulder height), and may also contribute to why estimates based on these dimensions were slightly less accurate (i.e. greater RE) than those based on horizontal dimensions not obstructed by vegetation (i.e. back length and total length; Table 5). These points highlight the importance of carefully measuring parameter inputs, which, in turn, will maximize the accuracy of estimates produced by the pinhole camera model.

Other currently described methods in the camera trap literature estimate animal size or distance from the camera using reference objects such as flags or other markers placed at known distances from the camera or spaced at known distances apart (e.g. Corlatti et al., 2020; Hofmeester et al., 2017; Tarugara et al., 2019; Willisch et al., 2013), by calibrating the field-of-view at each sampling site (e.g. Caravaggi et al., 2016), or by recreating the position of animals in images (Cui et al., 2020). Additional limitations of these methods include not being able to estimate continuous distances (e.g. if distance markers are used to bin animals to distance intervals), animals

being attracted to or avoiding artificial objects placed in the field-of-view of the camera (Corlatti et al., 2020; Hofmeester et al., 2017), increased human disturbance (e.g. trampling vegetation, spreading human scent) in the area directly in front of the camera at each field site, shifts in camera position (e.g. due to wind, leaning posts) potentially leading to increased errors or the need for re-calibration of the camera field-of-view (Caravaggi et al., 2016), or that images be processed before uninstalling the camera so researchers can identify the location of wildlife in images relative to landmarks in the camera field-of-view (Cui et al., 2020). However, one important benefit these methods have compared to the pinhole camera method is that animal size or distance can be estimated without the requirement of having to know (or assume) one variable to estimate the other. This may be particularly important for species that do not have published morphological data to estimate distances or in landscapes that do not provide a sufficient arrangement of naturally occurring distinguishable reference objects (e.g. such as trees, logs, rocks, topographical features) to estimate animal size. In such cases, it may be ideal to use the pinhole camera method in combination with one or more alternative approaches to satisfy project needs and capitalize on each methods' strengths to overcome their inherent limitations.

We found the pinhole camera method was simple to conceptualize and implement, generalizable to different camera traps and field conditions, and yielded accurate and precise estimates over a range of animal sizes and distances from the camera (Tables 2–5; Tables S1–S4) while addressing several limitations and challenges of other approaches used in the camera trap literature. First, we demonstrated that no prior knowledge of the physical specifications (e.g. focal length, sensor size) of the camera trap used to collect images is necessary to estimate animal sizes or distances accurately and precisely. Other methods and tools have been used when the physical specifications of cameras are known (Aleixo et al., 2020; Berger, 2012; Meise et al., 2014); however, many camera trap manufacturers do not report this information. For example, the physical sensor size and focal length of the Reconyx Hyperfire 2 HF2X cameras that we used were not reported by the manufacturer nor available in image metadata. Through a simple calibration procedure, we approximated each camera's focal length expressed in pixels which were later used to estimate sizes and distances of animals using equations derived from the pinhole camera model (Figure 1;

Table 1; Megalingam et al., 2016). Second, we found that the derived focal length was similar among three cameras of the same model and image setting (Table 1), suggesting that our calibration procedure produced reliable estimates and that once a pixel focal length has been established for a particular camera trap model and image setting, then it could be used to estimate physical dimensions in any image collected from that specific camera model and setting as we did in our reindeer field test (i.e. we did not estimate the focal length of the actual camera used to take photographs, but instead, applied the mean focal length determined from our initial camera calibrations). Third, we showed that the pinhole camera method can be used to estimate animal size or distance from the camera using a single camera and image. An alternative method often referred to as stereovision uses multiple cameras and the relationship between the position of an object simultaneously captured in two or more images from different perspectives to estimate size or distance (Cavagna et al., 2015; Xu et al., 2020). The obvious disadvantages of using stereovision methods for wildlife camera trapping are that multiple cameras must be used, resulting in decreased sampling sites or increased cost, battery usage, memory card storage and image processing time, as well as increased complexity of the field setup to ensure targeted wildlife are captured simultaneously from two or more perspectives.

As an example of the potential value of the pinhole camera method, consider a common scenario where a wildlife camera trap study was implemented which deployed many cameras with the initial intent to only document animal detections–non-detections for occupancy analyses. After all fieldwork was completed, researchers decide they want to quantitatively account for differences in sampling areas among camera sites (e.g. due to differences in habitat type or detection ranges of animals varying in size) or are interested in exploring other analyses for metrics such as animal abundance or density. Without having collected any additional information in the field, these researchers could use the camera calibration procedure we described along with any available information on the targeted species morphometrics to estimate the distances to animals in their images allowing for estimates of the effective area sampled or animal movement rates (Gilbert et al., 2020; Hofmeester et al., 2017; Rowcliffe et al., 2016). Note, however, careful consideration should be used when selecting animal morphometrics to estimate distances as any discrepancies between the known (or assumed) sizes used in calculations and the actual sizes of the animals captured will influence errors in estimates (see Figure 4). Previously described methods could not be used in this scenario without having collected additional information at each sampling site at the time of camera deployment or without revisiting sites and recreating the field-of-view of each camera. Therefore, the pinhole camera method adds utility to the enormously large camera trap datasets that have already been collected and also highlights the value of making images from wildlife camera trap projects publicly available whenever possible so that additional information may be collected retrospectively as data extraction methods continue to expand. We note, however, that application of the pinhole camera method is not as useful for

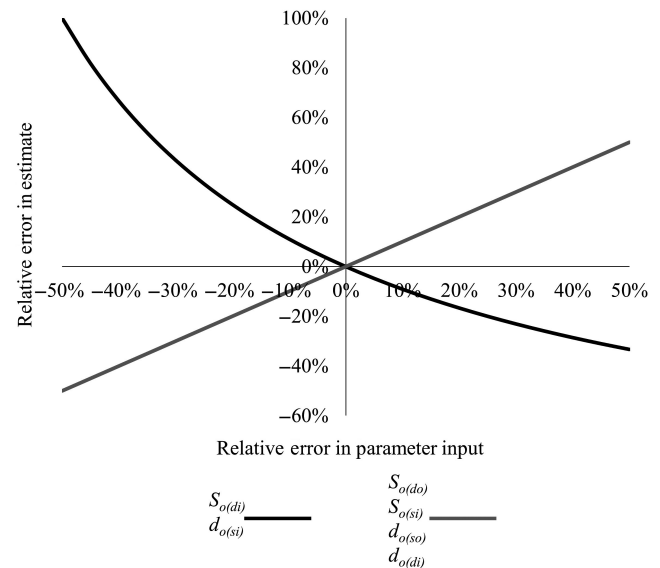


FIGURE 4 Theoretical errors in estimated object size (S_o) or distance from the camera (d_o) associated with variation in parameter inputs. In the figure legend, the parameter in parentheses represents the parameter being varied in calculations and corresponds with the x-axis. For example, $S_{o(di)}$ indicates that the size of the object (S_o) is being estimated based on variation in the focal length of the camera (d_i)

estimating animal size retrospectively because distances to reference objects in the camera field of view would need to be established to estimate animal sizes. Therefore, to use the pinhole camera method to estimate sizes, distances to reference objects should be recorded after the camera is set up or before camera removal.

4.1 | Limitations and considerations

While the pinhole camera method performed well under our experimental conditions and addresses several limitations and challenges of other methods for estimating animal size or distance from the camera, there are two general limitations of using the pinhole camera model that users should be keenly aware of when applying this method in a wildlife camera trap setting. First, as discussed above, it is necessary to know either the physical size of objects or their distance to the camera to estimate the other (Equations 1 and 2, Figure 1). Any errors or variation unaccounted for in known (or assumed) physical dimension inputs would likely result in increased errors or reduced reliability of estimates (e.g. the RE in the inputted physical distance is directly proportional to the RE in the estimated physical size of the animal and vice versa; Figure 4). Therefore, the error relationships summarized in Figure 4 may be used to help researchers anticipate how well this method will perform under specific project objectives and conditions or help inform the specificity of information needed to meet desired goals.

Second, the pinhole camera model as used here does not account for any type of image distortion that can result from different

types of lenses used in cameras (i.e. optical distortions) or the relative position of objects in relation to the camera or within the image (i.e. perspective distortions; Aleixo et al., 2020; Kannala et al., 2008). Optical distortions are unique to each specific set of technical camera specifications and lens characteristics, but in general, error size will increase as the position of the object moves away from the centre of the image and as the dimension being measured covers a larger portion of the image (e.g. if the animal is very close to the camera such that the dimension being measured covers most of the image; Aleixo et al., 2020; Neale et al., 2011). Typically, the wider the camera field-of-view, the greater effect these factors have towards the edges of the image (Figure S1; Kannala et al., 2008). Perspective distortions will generally increase errors in estimates when the apparent size of the dimension being measured is highly variable depending on the object's orientation relative to the camera (Figure 3). For example, if an animal were in the centre of an image, the apparent nose-to-tail length would vary considerably depending on whether the animal was perpendicular or directly facing away from the camera when the image is taken, whereas shoulder height would remain relatively unchanged (Figure 3; Berger, 2012; Meise et al., 2014; Zhang et al., 2018). We attempted to minimize these sources of error in our animal target simulation by capturing images with animal targets approximately perpendicular to the camera face and centred in the image to evaluate the performance of this method under ideal study conditions. However, we only observed a slight increase in errors in our field test where optical distortions and animals positions relative to the camera were not controlled for, suggesting that the pinhole camera model can provide robust estimates despite these potential sources of error.

Given the limitations identified above, we provide several considerations and suggestions to minimize errors in estimated physical dimensions of animals in camera trap images when using the pinhole camera model. First, when selecting a morphological dimension, attempt to use dimensions with low variation among individuals to produce the most accurate results. For example, if shoulder height has less variation than nose-to-tail length for a particular species, then using shoulder height for pixel measurements would be preferred. Second, be aware that the apparent size of an object dimension is influenced by its orientation relative to the camera (Figure 3). Therefore, selecting a dimension that is most consistent regardless of its orientation relative to the camera is ideal. Third, when selecting objects to measure the distance to, consider natural paths animals may take and attempt to use distances corresponding with where animals are most likely to be captured in the image. If a sufficient arrangement of natural objects is unavailable in the camera field-of-view, consider removal methods so artificial objects are not left in the field (Caravaggi et al., 2016; Cui et al., 2020). Lastly, to minimize errors associated with optical distortions, attempt to use images where animals are near the centre of the image and measure dimensions that do not take up a considerable portion of the image while still being able to measure accurately and precisely. Based on the field-of-view of most camera traps (<50 degrees; Trolliet et al., 2014), optical distortions are unlikely to be a major issue for

most wildlife camera trap applications; however, there are multiple ways to account for this if extremely high precision and accuracy are necessary (e.g. barrel distortion correction; Aleixo et al., 2020; Kannala et al., 2008). While specific project conditions and objectives will likely determine the desired level of accuracy and precision in estimated animal sizes and distances, being aware of these limitations and considerations will help ensure the best possible results are achieved when applying the pinhole camera method.

5 | CONCLUSIONS

To our knowledge, the use of the pinhole camera model for estimating the size or distance to animals has not been documented or connected to previously described methods in the camera trap literature. Previous methods have typically focused on satisfying specific project objectives and are often not generalizable across camera trap models or installations. We found the pinhole camera model can provide multiple types of data useful to camera trap researchers and can be applied to different cameras and field setups. We described how to apply the pinhole camera model approach through a controlled simulation and more practical field test and showed how estimates can be influenced by variation in model parameters to help establish realistic expectations and inform creative uses in wildlife camera trap research. While other methods have been used to derive estimates of animal size or distance, we showed how this method addresses current limitations and shortcomings while remaining conceptually simple and easy to implement. Ultimately, since wildlife camera trapping has become an increasingly applied method of collecting information on wildlife, developing innovative and generalizable methods for extracting additional information from images will help inform standardized protocols and maximize camera trapping's utility for wildlife monitoring and conservation.

ACKNOWLEDGEMENTS

We thank the National Science Foundation for providing funding for this study [Office of Polar Programs, Arctic System Science Program Award # 1839192]. We also thank Sarah Barcalow (Lead Animal Care technician, Large Animal Research Station, University of Alaska Fairbanks, Fairbanks, Alaska, USA) for providing access to reindeer enclosures and assisting with data collection of morphological dimensions. We also thank Knut Kielland (University of Alaska Fairbanks, Fairbanks, Alaska, USA) and Shawn Crimmins (U.S. Geological Survey, Alaska Cooperative Fish and Wildlife Research Unit, University of Alaska Fairbanks, Fairbanks, Alaska, USA) for their comments and thoughtful review of this manuscript.

CONFLICT OF INTEREST

No authors had any conflict of interest.

AUTHORS' CONTRIBUTIONS

S.L. and T.B. conceived the research idea and collected the field data; S.L. designed the methodology, analysed the data and led writing

of the manuscript; T.F. contributed to design and interpretation of analyses. All authors contributed critically to drafts and gave final approval for publication.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/2041-210X.13880>.

DATA AVAILABILITY STATEMENT

The data used in this research are available through the National Science Foundation's Arctic Data Center and can be found at <https://doi.org/10.18739/A2BZ61933> (Leorna, 2022).

ORCID

Scott Leorna  <https://orcid.org/0000-0001-8689-8197>

Todd Brinkman  <https://orcid.org/0000-0001-5375-4840>

Timothy Fullman  <https://orcid.org/0000-0002-6908-9184>

REFERENCES

- Aleixo, F., O'Callaghan, S. A., Ducla Soares, L., Nunes, P., & Prieto, R. (2020). AragoJ: A free, open-source software to aid single camera photogrammetry studies. *Methods in Ecology and Evolution*, 11(5), 670–677. <https://doi.org/10.1111/2041-210X.13376>
- Berger, J. (2012). Estimation of body-size traits by photogrammetry in large mammals to inform conservation. *Conservation Biology*, 26(5), 769–777. <https://doi.org/10.1111/j.1523-1739.2012.01896.x>
- Burton, A. C., Neilson, E., Moreira, D., Ladle, A., Steenweg, R., Fisher, J. T., Bayne, E., & Boutin, S. (2015). Wildlife camera trapping: A review and recommendations for linking surveys to ecological processes. *Journal of Applied Ecology*, 52(3), 675–685. <https://doi.org/10.1111/1365-2664.12432>
- Caravaggi, A., Banks, P. B., Burton, A. C., Finlay, C. M. V., Haswell, P. M., Hayward, M. W., Rowcliffe, M. J., & Wood, M. D. (2017). A review of camera trapping for conservation behaviour research. *Remote Sensing in Ecology and Conservation*, 3(3), 109–122. <https://doi.org/10.1002/rse2.48>
- Caravaggi, A., Zaccaroni, M., Riga, F., Schai-Braun, S. C., Dick, J. T. A., Montgomery, W. I., & Reid, N. (2016). An invasive-native mammalian species replacement process captured by camera trap survey random encounter models. *Remote Sensing in Ecology and Conservation*, 2(1), 45–58. <https://doi.org/10.1002/rse2.11>
- Cavagna, A., Creato, C., Del Castello, L., Giardina, I., Melillo, S., Parisi, L., & Viale, M. (2015). Error control in the set-up of stereo camera systems for 3d animal tracking. *European Physical Journal: Special Topics*, 224, 3211–3232. <https://doi.org/10.1140/epjst/e2015-50102-3>
- Corlatti, L., Sivieri, S., Sudolska, B., Giacomelli, S., & Pedrotti, L. (2020). A field test of unconventional camera trap distance sampling to estimate abundance of marmot populations. *Wildlife Biology*, 2020(4), 1–11. <https://doi.org/10.2981/wlb.00652>
- Cui, S., Chen, D., Sun, J., Chu, H., Li, C., & Jiang, Z. (2020). A simple use of camera traps for photogrammetric estimation of wild animal traits. *Journal of Zoology*, 312(1), 12–20. <https://doi.org/10.1111/jzo.12788>
- Gilbert, N. A., Clare, J. D. J., Stenglein, J. L., & Zuckerberg, B. (2020). Abundance estimation of unmarked animals based on camera-trap data. *Conservation Biology*, 35(1), 88–100. <https://doi.org/10.1111/cobi.13517>
- Hofmeester, T. R., Rowcliffe, J. M., & Jansen, P. A. (2017). A simple method for estimating the effective detection distance of camera traps. *Remote Sensing in Ecology and Conservation*, 3(2), 81–89. <https://doi.org/10.1002/rse2.25>
- Hofmeester, T. R., Young, S., Juthberg, S., Singh, N. J., Widemo, F., Andren, H., Linnell, J. D. C., & Cromsigt, J. P. G. M. (2020). Using by-catch data from wildlife surveys to quantify climatic parameters and timing of phenology for plants and animals using camera traps. *Remote Sensing in Ecology and Conservation*, 6(2), 129–140. <https://doi.org/10.1002/rse2.136>
- Howe, E. J., Buckland, S. T., Despres-Einspenner, M.-L., & Kuhl, H. S. (2017). Distance sampling with camera traps. *Methods in Ecology and Evolution*, 8, 1558–1565. <https://doi.org/10.1111/2041-210X.12790>
- Kannala, J., Heikkilä, J., & Brandt, S. S. (2008). Geometric camera calibration. In B. W. Wah (Ed.), *Wiley encyclopedia of computer science and engineering* (pp. 1–20). John Wiley & Sons, Inc. <https://doi.org/10.1002/9780470050118.ecse589>
- Klein, D. R., Meldgaard, M., & Fancy, S. G. (1987). Factors determining leg length in Rangifer tarandus. *Journal of Mammalogy*, 68(3), 642–655. Retrieved from <https://academic.oup.com/jmammal/article/68/3/642/851199>
- Leorna, S. (2022). Camera Trap Photogrammetry Dataset, Fairbanks, Alaska, USA, Sept. 2021. Arctic Data Center. <https://doi.org/10.18739/A2BZ61933>
- Megalingam, R. K., Shriram, V., Likhith, B., Rajesh, G., & Ghanta, S. (2016). Monocular distance estimation using pinhole camera approximation to avoid vehicle crash and back-over accidents. In *Proceedings of the 10th international conference on intelligent systems and control (ISCO)* (pp. 1–5). <https://doi.org/10.1109/ISCO.2016.7727017>
- Meise, K., Mueller, B., Zein, B., & Trillmich, F. (2014). Applicability of single-camera photogrammetry to determine body dimensions of pinnipeds: Galapagos Sea lions as an example. *PLoS ONE*, 9(7), 1–7. <https://doi.org/10.1371/journal.pone.0101197>
- Moeller, A. K., Lukacs, P. M., & Horne, J. S. (2018). Three novel methods to estimate abundance of unmarked animals using remote cameras. *Ecosphere*, 9(8), 1–15. <https://doi.org/10.1002/ecs2.2331>
- Neale, W. T., Hessel, D., & Terpstra, T. (2011). Photogrammetric measurement error associated with lens distortion. *SAE Technical Papers*, 1–54.
- Nieminen, M., & Helle, T. (1980). Variations in body measurements of wild and semi-domestic reindeer (*Rangifer tarandus*) in Fennoscandia. *Annales Zoologici Fennici*, 17(4), 275–283. Retrieved from <https://www.jstor.org/stable/23734512>
- O'Connell, A., Nichols, J., & Karanth, K. (2011). In A. O'Connell, J. Nichols, & K. Karanth (Eds.), *Camera traps in animal ecology*. Springer. <https://doi.org/10.1007/978-4-431-99495-4>
- Rowcliffe, J. M., & Carbone, C. (2008). Surveys using camera traps: Are we looking to a brighter future? *Animal Conservation*, 11(3), 185–186. <https://doi.org/10.1111/j.1469-1795.2008.00180.x>
- Rowcliffe, J. M., Carbone, C., Jansen, P. A., Kays, R., & Kranstauber, B. (2011). Quantifying the sensitivity of camera traps: An adapted distance sampling approach. *Methods in Ecology and Evolution*, 2(5), 464–476. <https://doi.org/10.1111/j.2041-210X.2011.00094.x>
- Rowcliffe, J. M., Field, J., Turvey, S. T., & Carbone, C. (2008). Estimating animal density using camera traps without the need for individual recognition. *Journal of Applied Ecology*, 45(4), 1228–1236. <https://doi.org/10.1111/j.1365-2664.2008.01473.x>
- Rowcliffe, J. M., Jansen, P. A., Kays, R., Kranstauber, B., & Carbone, C. (2016). Wildlife speed cameras: Measuring animal travel speed and day range using camera traps. *Remote Sensing in Ecology and Conservation*, 2, 84–94. <https://doi.org/10.1002/rse2.17>
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH image to ImageJ: 25 years of image analysis. *Nature Methods*, 9, 671–675. <https://doi.org/10.1038/nmeth.2089>
- Scotson, L., Johnston, L. R., Iannarilli, F., Wearn, O. R., Mohd-azlan, J., Wong, W., Gray, T. N. E., Dinata, Y., Suzuki, A., Willard, C. E., Frechette, J., Loken, B., Steinmetz, R., Moßbrucker, A. M., Clements, G. R., & Fieberg, J. (2017). Best practices and software for the management and sharing of camera trap data for small and large scales

- studies. *Remote Sensing in Ecology and Conservation*, 3(3), 158–172. <https://doi.org/10.1002/rse2.54>
- Sollmann, R. (2018). A gentle introduction to camera-trap data analysis. *African Journal of Ecology*, 56(4), 740–749. <https://doi.org/10.1111/aje.12557>
- Tarugara, A., Clegg, B. W., Gandiwa, E., Muposhi, V. K., & Wenham, C. M. (2019). Measuring body dimensions of leopards (*Panthera pardus*) from camera trap photographs. *PeerJ*, 1–19, e7630. <https://doi.org/10.7717/peerj.7630>
- Thomson, R., Potgieter, G., & Bahaa-el-din, L. (2018). Closing the gap between camera trap software development and the user community. *African Journal of Ecology*, 56, 721–739. <https://doi.org/10.1111/aje.12550>
- Trolliet, F., Huynen, M., Vermeulen, C., & Hambuckers, A. (2014). Use of camera traps for wildlife studies. A review. *Biotechnology, Agronomy, Society and Environment*, 18(3), 446–454. [https://doi.org/10.1016/0308-0161\(78\)90006-6](https://doi.org/10.1016/0308-0161(78)90006-6)
- Willisch, C. S., Marreros, N., & Neuhaus, P. (2013). Long-distance photogrammetric trait estimation in free-ranging animals: A new approach. *Mammalian Biology*, 78(5), 351–355. <https://doi.org/10.1016/j.mambio.2013.02.004>
- Xu, Z., Sun, L., Wang, X., Lei, P., He, J., & Zhou, Y. (2020). Stereo camera trap for wildlife in situ observations and measurements. *Applied Optics*, 59(10), 3262–3269. <https://doi.org/10.1364/ao.389835>
- Young, S., Rode-Margono, J., & Amin, R. (2018). Software to facilitate and streamline camera trap data management: A review. *Ecology and Evolution*, 8(19), 9947–9957. <https://doi.org/10.1002/ece3.4464>
- Zhang, A. L. N., Wu, B. P., Jiang, C. X. H., Xuan, D. C. Z., Ma, E. Y. H., & Zhang, F. Y. A. (2018). Development and validation of a visual image analysis for monitoring the body size of sheep. *Journal of Applied Animal Research*, 46(1), 1004–1015. <https://doi.org/10.1080/09712119.2018.1450257>
- Zheng, X., Owen, M. A., Nie, Y., Hu, Y., Swaisgood, R. R., Yan, L., & Wei, F. (2016). Individual identification of wild giant pandas from camera trap photos – A systematic and hierarchical approach. *Journal of Zoology*, 300, 247–256. <https://doi.org/10.1111/jzo.12377>

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Leorna, S., Brinkman, T. & Fullman, T. (2022). Estimating animal size or distance in camera trap images: Photogrammetry using the pinhole camera model. *Methods in Ecology and Evolution*, 13, 1707–1718. <https://doi.org/10.1111/2041-210X.13880>