SURVEYING MIXED-SPECIES WATERBIRD COLONIES WITH UNMANNED AERIAL SYSTEMS (UAS): VISIBILITY BIAS, DISTURBANCE, AND PROTOCOL RECOMMENDATIONS

by

Jarred R. Barr, B.S.

A thesis submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Science
with a Major in Wildlife Ecology
December 2017

Committee Members:

M. Clay Green, Chair

Stephen J. DeMaso

Thomas B. Hardy

COPYRIGHT

by

Jarred R. Barr

2017

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

This work is protected by the Copyright Laws of the United States (Public Law 94-553, section 107). Consistent with fair use as defined in the Copyright Laws, brief quotations from this material are allowed with proper acknowledgement. Use of this material for financial gain without the author's express written permission is not allowed.

Duplication Permission

As the copyright holder of this work I, Jarred Barr, authorize duplication of this work, in whole or in part, for educational or scholarly purposes only.

ACKNOWLEDGMENTS

I would like to thank my advisor, M. Clay Green, for his guidance and support. I also thank Stephen DeMaso and Thomas Hardy for serving on my committee, Bill Vermillion for providing additional expertise and feedback, and Floyd "Butch" Weckerly for statistical advice.

This project would not have been possible without the contributions of many biologists, agencies, landowners, and partners: Laurie Lomas Gonzales and Stuart Marcus at Trinity River National Wildlife Refuge, Houston Daniel, Stephanie Bilodeau and Owen Fitzsimmons at Coastal Bend Bays and Estuaries Program, Audubon Texas, Larry Shriver, and King Ranch Preserve. Jacob Bilbo, Thomas Heard, Jessica Frye, and Kristy Kollaus provided technical assistance. I further thank colleagues that helped with fieldwork and counting decoys: Anna Mathews, Charlie Bintliff, Rebekah Rylander, Christina Farrell, Sarah Durham, Isabelle Walker, Thom Marshall, James Tolliver, Jacob Bilbo, Stephanie Bilodeau, and Rachel Glazner. This project was graciously funded by the Gulf Coast Joint Venture.

I would like to thank my family for their love and support, and tacos for keeping me motivated.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS iv

LIST OF TABLES vii

LIST OF FIGURES viii

LIST OF ABBREVIATIONS x

ABSTRACT xi

CHAPTER

I. DETECTABILITY AND VISIBILITY BIASES ASSOCIATED WITH USING UNMANNED AIRCRAFT TO SURVEY COLONIAL WATERBIRDS 1

STUDY AREA 5
METHODS 6
UAS surveys 7
Observer counts 8
Statistical analysis 9
RESULTS 14
Visibility bias 14
Observer error 16
DISCUSSION 18
MANAGEMENT IMPLICATIONS 23

II. DISTURBANCE TO COLONIAL WATERBIRDS CAUSED BY UNMANNED AIRCRAFT SURVEYS 25

STUDY AREA 29
Lower Laguna Madre 30
Lower Trinity River Basin 31
METHODS 32
Behavioral observations 34
Statistical analysis 35

RESULTS 39
Vigilance behavior 40
Flushing behavior.....41
DISCUSSION 42
MANAGEMENT IMPLICATIONS 47

APPENDIX SECTION.....71

LITERATURE CITED 78

LIST OF TABLES

Table **Page**

1. Detection probability and bias correction factors of waterbird decoys 48

2. Selected generalized linear mixed-effect models assessing ability of observers ($N = 6$) to estimate decoy counts from aerial imagery taken with unmanned aircraft 49

3. Competing models for the magnitude of behavioral responses displayed in nesting waterbird colonies, as influenced by UAS surveys 51

4. Parameter estimates of the selected model for vigilance 53

LIST OF FIGURES

Figure **Page**

1. Map of the Gulf Coast Joint Venture (GCJV) region 55

2. Study sites where aerial imagery of decoy "colonies" was obtained 56

3. Subsection of imagery taken with a Phantom UAS (DJI, Shenzhen, Guangdong, China; 12.4 MP camera) 58

4. Mean percent error (± 1 SE) for six independent-observer counts of decoy abundance 60

5. Linear regressions illustrating the ability of Chapman (1951) abundance estimations to estimate decoy density from image quadrats 61

6. Sites in which UAS surveys were conducted over active waterbird colonies 63

7. The Minion 2.0 fixed-wing UAS 65

8. Proportion of four heron species ($\bar{x}$ ± SE) in camera frame that exhibited a behavioral response in relation to treatment levels 66

9. Proportion of ground-nesting waterbirds ($\bar{x}$ ± SE) in camera frame that exhibited a behavioral response in relation to treatment levels 67

10. Proportion of roseate spoonbill and white ibis ($\bar{x}$ ± SE) in camera frame that exhibited a behavioral response in relation to treatment levels 69

11. Conditional modes for the random effect of colony, represented for the selected flush behavior model.....70

LIST OF ABBREVIATIONS

Abbreviation	Description
AGL	Above ground level
AIC	Akaike's information criterion
AIC_c	Akaike's information criterion, corrected for small sample size
ΔAIC	The difference in AIC units of the top model and a specified model
EFS	East Flat Spoil island
FDR	False discovery rate
GI	Green Island
GLMM	Generalized linear mixed-effect model
Lag 1	Lag 1 temporal autocorrelation
LM	Linear regression model
LRT	Likelihood ratio test
NWR	National Wildlife Refuge
PB	Parametric bootstrap
Spp.	Species, plural
TRB	Trinity River basin
UAS	Unmanned aerial system

ABSTRACT

Surveys of colonial-nesting waterbirds are necessary for assessing population trends, gaining insight into wetland ecosystem health, and even determining the impact of natural disasters and other environmental concerns. The popularity of unmanned aerial systems (UAS) for use as a surveying tool has risen in the past decade, but little research has been conducted on the effectiveness of such technology. I investigated visibility bias and disturbance impacts associated with using UAS to survey waterbird colonies in Texas, specifically in cypress-tupelo watershed and coastal island habitats. I used a stratified random design to place four waterbird decoy types (black skimmers, terns, and white- and dark-plumaged herons) in each habitat and had six observers independently count decoys from aerial imagery taken with a consumer-grade UAS (DJI Phantom). I used generalized linear mixed-effects models to estimate detection probabilities of each decoy type. Black skimmers were the only decoy type at the dredge-spoil island to have a detection probability of significantly less than 100% (0.54 [0.44–0.63 CI], $P \leq 0.001$). Detectability of both white- and dark-plumaged herons decreased considerably in the canopied cypress-tupelo habitat when compared to the dredge-spoil island (by 80 and 84%, respectively). In addition, for surveys in cypress-tupelo habitat where cloud cover was above 50%, detectability of white heron decoys decreased significantly by another 20% (0.09 [0.03–0.34 CI], $P = 0.007$). Detection rates varied among observers, but only significantly for models of white-plumaged herons ($P = 0.022$) and black skimmers ($P = 0.05$). Use of the double-sample method yielded biased-low abundance estimates for

white- and dark-plumaged herons in canopied sites, suggesting that habitat differences were a greater source of bias than observer error. I investigated disturbance to waterbirds by setting up video cameras at the periphery of active nesting colonies while surveying with unmanned aircraft. I tested the effects of two UAS platforms – and a range of altitudes flown between them – on the behavioral reactions exhibited in four active colonies in Texas. Reactions were tallied in 1-minute sampling periods at each nesting colony, which were used to estimate generalized linear mixed-effects models for vigilance and flush behavior. I found that the consumer-grade UAS (DJI Phantom) increased vigilance in mixed-species colonies for survey altitudes of 91, 61, and 46 m when compared to a baseline control. Vigilant reactions were increased in magnitude by 72, 119, and 118% for these altitudes, respectively. Flush reactions were not influenced by either platform or any altitude flown. Surveys with the fixed-wing UAS did not impact vigilance or flush behavior, likely because it was used at suggested altitudes of 300 and 200 m and was hardly detectable from the ground. My results suggest that managers should employ UAS surveys on clear days in high-visibility habitats, or otherwise use another survey method to supplement photographic counts obtained by UAS. In addition, surveys should be flown between 46–91 m only when high resolution imagery is needed (e.g. for abundance estimates) to mitigate disturbance. Even though the Phantom UAS caused increased vigilance, if surveys are done promptly and in back-and-forth transects, the impact of this increased behavior is likely negligible especially when considering the much more harmful effects of ground-based survey methods.

I. DETECTABILITY AND VISIBILITY BIASES ASSOCIATED WITH USING UNMANNED AIRCRAFT TO SURVEY COLONIAL WATERBIRDS

Conservation of colonial nesting waterbirds is reliant on the ability to estimate changes in populations over time. Surveys, conducted routinely throughout the spring and summer months, are employed to estimate the number of breeding pairs as well as identify species composition within nesting colonies. Waterbird colony presence (i.e. species composition and abundance) is an indicator of ecosystem health (Temple and Wiens 1989, Kushlan 1993, Kingsford 1999, Johnson and Krohn 2001, Smith 2002). Kingsford (1999) proposed that because they forage on fish, plants, insects, and other biota, the presence of these birds may be a useful index in measuring the health of a wetland. Waterbird abundance can also provide insight into other regional and global environmental concerns, such as the effectiveness of restoring wetland habitat (Ma et al. 2010), the effect of large-scale climatic events such as droughts or hurricanes (Leberg et al. 2007, Raynor et al. 2013), and the effects of human-caused disturbance such as oil spills and habitat loss (Paracuellos and Tellería 2004, Ma et al. 2010). Moreover, reliable colony survey estimates are especially vital for species of concern, such as the reddish egret (*Egretta rufescens*), which is listed as state threatened in Texas and near-threatened by the International Union for Conservation of Nature (IUCN 2015), and the little blue heron (*Egretta caerulea*), whose population decline in the Southeastern U.S. has increasingly received attention (Hunter et al. 2006, Green et al. 2008, Rodgers and Smith 2012).

Long-term monitoring and estimation of breeding pairs can shed light on population trends, which can have implications on management decisions put forth in the

future (Wiens et al. 1986, Temple and Wiens 1989, Thomas 1996, Coulter and Frederick 1997). However, being able to detect individuals in a population is essential to constructing reliable estimates of abundance. Imperfect detection has been recognized as problematic in multiple ecological systems, including the detectability of invasive plants (Regan et al. 2011), otter (*Lontra canadensis*) sign (Jeffress et al. 2011), and auditory counts of birds (Simons et al. 2007). Most population indices are based on raw counts, which can be highly variable and subject to bias, so evaluating detection probability is necessary for obtaining accurate and precise estimates of abundance (Pollock and Kendall 1987).

Three variables have an effect on detection probability: observer error, environmental conditions, and heterogeneity in animal behavior (Anderson 2001). With regard to waterbird detectability, these variables are typically associated with visual surveys (e.g. from fixed-wing aircraft or via colony walk-through). Aircraft surveys in particular can produce many potential sources of visibility bias, including visual occlusion of nests due to vegetative characteristics (Butler et al. 2007, Conroy et al. 2008), differences in visibility due to aircraft type (Green et al. 2008, Koneff et al. 2008), and variability of counts among seat positions in the aircraft (Conroy et al. 2008). Due to these causes of visibility bias, aerial surveys have been known to, sometimes significantly, underestimate numbers of nesting birds. Frederick et al. (2003) found that, during a simulation of real-time aerial surveys, observers undercounted by a mean of 29% when compared to true numbers of birds. Although undercounting was common during the study, variability of count indices among and within observers was extreme and unpredictable even for experienced biologists. Surveys from aircraft have

occasionally overestimated species like snowy egrets (*Egretta thula*), an error which is likely due to species misidentification (Kushlan 1979, Rodgers et al. 2005, Green et al. 2008, Williams et al. 2008).

Researchers have devised a multitude of ways to correct for bias and imperfect detection. Double-count methods, for example, have been widely used to correct observer bias and provide estimates of true abundance (Cook and Jacobson 1979, Bayliss and Yeomans 1989, Graham and Bell 1989, Nichols 2000, Green et al. 2008). These methods are related to the Lincoln-Peterson index for estimating abundance, where a detection by one observer represents a single capture and a detection by the second observer represents recapture (Seber 1982). Sightability models, another common approach, offer a means of estimating abundance by quantifying different sources of visibility bias with regression-based modeling (e.g. canopy cover and aircraft speed; Caughley 1974). Given the error associated with real-time aircraft surveys, some researchers have improved the detectability of waterbirds by using aerial photography. Both Bayliss and Yeomans (1990) and Frederick et al. (2003) observed a marked increase in detectability by using photographs, presumably because large and dense colonies are difficult to count from a moving aircraft. Aerial photography, while more accurate than real-time surveys, is still subject to the same issues of observer error, environmental covariates, and species-specific behavior. The accuracy of aerial photographic counts is likely dependent upon the situation in which they are being employed. Photographs over dense canopy, for example, might yield more error-prone estimates because the observer can only see the colony from a single angle of approach.

Since traditional survey methodology is not without error, and there is lack of consensus over which method is best, new survey techniques are frequently tested in the field. Unmanned aerial systems (UAS) have recently garnered attention in regards to wildlife surveying. In previous studies, they have been used to census large colonies of common terns (*Sterna hirundo*, Chabot et al. 2015), detect roe deer (*Capreolus capreolus*) fawns in meadows with a thermal camera (Israel 2011), monitor breeding pairs of black-headed gulls (*Chroicocephalus ridibundus*, Sardá-palomera et al. 2012), and monitor the nesting status of hooded crows (*Corvus cornix*, Weissensteiner et al. 2015). Unmanned aircraft have a few key benefits that make this new technology appealing to surveying wildlife, such as suitability for fine spatial resolution (Bakó et al. 2014), comparable coverage to traditional aerial surveys (Chabot et al. 2014, Bakó et al. 2014), researcher safety, and ability to easily switch out sensors or payloads (Watts et al. 2010). In a spatial sense, there is potential for UAS to bridge the gap between ground surveys and traditional aerial surveys (Chabot et al. 2014). While they may not have the vast range of manned aircraft (Kingsford 1999), they can certainly be applied to a variety of circumstances in the field (e.g. when a colony is inaccessible to ground observers, or when fine spatial scaling is needed). Due to the recent popularity and availability of UAS among researchers, there is a clear need to study the ramifications of their use as a survey tool.

My overall research objective was to quantify detection probability and sources of bias related to counting waterbirds from aerial imagery taken with an unmanned aircraft. In particular, I wanted to (1) estimate detectability of four unique waterbird species, (2)

identify sources of variation in detection, and (3) investigate observer error associated with imagery counts, as well as the accuracy of double-sample estimates.

STUDY AREA

There are four colonial nesting waterbird species currently listed as priority within the Monitoring, Evaluation, and Research Team (MERT) of the Gulf Coast Joint Venture (GCJV): reddish egret (*Egretta rufescens*), black skimmer (*Rynchops niger*), gull-billed tern (*Gelochelidon nilotica*), and little blue heron (*Egretta caerulea*). I selected survey sites in Texas *a priori* to estimate the detectability of these 4 species in their respective nesting habitats. The GCJV is one of 18 joint ventures in the continental US, and its coverage spans coastlines in Texas, Louisiana, Alabama, and Mississippi (Figure 1). Because of variability in nesting strata among these species, I chose study sites in two distinct regions of Texas: coastal islands in the lower Laguna Madre and forested wetlands in the lower Trinity River basin.

Lower Laguna Madre.–The Laguna Madre of Texas has a subhumid-to-semiarid subtropical climate, with extreme variability in precipitation (Fulbright et al. 1990). Geographically, it is a long and narrow hypersaline coastal lagoon that extends roughly 185 km from Corpus Christi Bay to the southern tip of the state (Tunnel 2002). The lagoon is split into two subunits – upper and lower – by a land-bridge that connects to the barrier island spit known as Padre Island (Tunnel 2002). I utilized a colony site at East Flat Spoil island (EFS), which is situated within Kenedy County (Figure 2). This small dredge-spoil island is 1.2 ha and is managed by Texas Audubon. Vegetation on EFS is mostly comprised of low-lying, highly saline-tolerant species, like sea oxeye (*Borrichia frutescens*), shoregrass (*Monanthochloe littoralis*), sea purslane (*Sesuvium*

portulacastrum), and turtleweed (*Batis maritima*). The mix of vegetated area and bare ground attracts a diverse set of ground-nesting waterbirds to the island every year, including black skimmers, gull-billed terns, reddish egrets, great blue herons (*Ardea herodius*), royal terns (*Thalasseus maximus*), sandwich terns (*Thalasseus sandvicensis*), and laughing gulls (*Leucophaeus atricilla*).

Lower Trinity River basin.–The Trinity River basin (TRB) runs from North Texas – its headwaters just west of Fort Worth – all the way to the Galveston Bay system on the Gulf of Mexico (Dahm et al. 2011). The lower portion of the basin has a humid subtropical climate, and is predominantly forested (Dahm et al. 2011). Cypress-lined bayous in the lower Trinity watershed provide ideal habitat for canopy-nesting waterbird species, including the little blue heron. Due to federal restrictions on UAS operation at the time of this study, I was unable to survey at colony sites within Trinity River NWR proper, and instead utilized a colony site on Josie lake adjacent to refuge property (Figure 2). The area is dominated by bald cypress trees (*Taxodium ascendens*), tupelo (*Nyssa* spp.), button bush (*Cephalanthus occidentalis*), and a diversity of bottomland hardwood species. Many waterbird species use the lower TRB to nest and forage, including great egrets (*Ardea alba*), snowy egrets (*Egretta thula*), anhingas (*Anhinga anhinga*), cattle egrets (*Bubulcus ibis*), and little blue herons.

METHODS

UAS surveys were conducted between 0930 and 1630 hours on clear to cloudy days, with wind conditions ranging between 0–20.4 km/hour. We flew surveys parallel to the ground in a back-and-forth transect pattern (East to West) to capture a series of overlapping photographs, which would later be processed into a single mosaicked, geo-

referenced orthophoto (Bolstad 2016). We used a consumer-grade UAS (Phantom 3, DJI, Shenzhen, Guangdong, China) that came equipped with a 12.4 mega-pixel RGB camera. The Phantom is operated by a single technician, can be launched from almost any surface, and is flown via a minimalist ground station (remote control and iPad device [Apple Inc., Cupertino, CA]). Recent software packages allow for autonomous flight (Ground Station Pro, DJI), but at the time of this study no such software was available. Thus, surveys were flown manually. I flew surveys at an altitude of 61 m above ground level (AGL), which netted an image resolution of approximately 1.7 cm/pixel.

UAS Surveys

I used decoys to estimate detection probability and visibility bias of 4 wading bird groups: black skimmers, terns, white-plumaged herons, and dark-plumaged herons. All decoys with the exception of black skimmers were constructed from plastic lawn flamingos. I modified and painted plastic flamingos to reflect the size, shape, and plumage coloration of my target species (Green et al. 2005). All 4 decoy types were utilized at EFS, but just dark- and white-plumaged herons at TRB. So as not to coincide with the breeding season, I conducted surveys with decoy birds in September and November, 2016.

Prior to UAS decoy surveys, I selected old nest bowls for decoy placement by using a stratified random design (Altmann 1974). I knew where each target species (or target group) nested based on imagery taken over active colonies earlier that same year, and, using this information I marked a random subset of nest bowls with nondescript tongue depressor sticks. For each decoy type, I used a random azimuth and distance to select nest bowls, and a random number generator (uniform distribution) to select group

size. At TRB, I only used white- and dark-plumaged heron types, so I did not stratify nest bowl selection based on species. Rather, I randomized nest bowl selection based on height of the nest, because canopy closure can have an effect on detectability (Bayliss and Yeomans 1989, Rodgers et al. 2005, Pearse et al. 2008). I used a random number generator of nest heights that ranged between 2–8 m, and selected nests (after walking a random azimuth and distance) that were closest to the selected height given by the number generator. At each selected nest bowl, I recorded canopy cover (when applicable) with a spherical densiometer (Model A, Forestry Suppliers, Inc., Jackson, MS, USA) and marked its location with a high-accuracy GPS unit (Trimble GeoXT 3000, Sunnyvale, CA, USA). I had 117 total nest bowls at EFS and 50 at TRB, and I randomly selected from these for decoy placement prior to each survey. I positioned 4–5 ground control points (GCPs) at the perimeter of each decoy "colony" to aid in geo-referencing the mosaicked imagery (Bolstad 2016). I estimated cloud cover (%) with a spherical densiometer at the take-off and landing site prior to surveys.

Observer counts

I created mosaicked images of each decoy "colony" with the photomerge feature in Adobe Photoshop (Adobe Systems, Inc., San Jose, CA, USA). I then used ArcMap 10.4 (Esri, Redlands, CA, USA) to geo-reference each image and overlay all nest bowl locations. Due to issues with fully mosaicking the imagery taken at TRB, and because I wanted to standardize the photographs for observer counts, I split each randomized decoy "colony" into a set of 2000×1500-pixel photographs (Figure 3). Thus, for each imagery quadrat, I had a known density of decoys, and, when pooled together, a known abundance for each UAS survey.

I had 6 observers independently count decoys in each of the imagery quadrats, from the same computer and in the same order. Observer count indices were done alone and without help, and so were assumed independent from one another. I gave the same instructions to each observer, as well as a sample image to establish a "search image" of each decoy type (not used in analysis). None of the 6 observers had prior experience counting birds from aerial photography, but 5 of the 6 had some experience with other avian survey methods (e.g. ground counts and auditory surveys). Some loafing birds were captured in photographs, but were excluded from analysis because I had known locations for each decoy present. Observers went through photographs in sequence, and using the paint tool in Photoshop, placed a color dot on each detected decoy (Figure 3). Different colors corresponded to different decoy types. After all observers provided their respective count indices, I used the count function and image analysis function in Adobe Photoshop to tally observer counts (Chabot et al. 2015).

Statistical analysis

I modeled the detectability of each decoy type by using methods similar to Williams et al. (2008), where observer count indices were compared to corresponding known abundances. I pooled image quadrat counts by survey to get a total index for each survey and to represent errors of omission and commission by each observer. Pooling was also done to account for misidentification of decoy types, which meant that one decoy type was overcounted while another was undercounted. I used generalized linear mixed-effect models (GLMMs) with a Poisson error distribution and log-link function to estimate detectability and test the effects of several environmental covariates (McCullagh and Nelder 1989). Models with a negative binomial error distribution were also considered,

but were not used because over-dispersion was not present in the Poisson models. I included observer as a random effect, and, to preserve correspondence among observer counts and the known decoy abundance, I modeled each survey as a random effect as well. The basic model structure, without the addition of environmental covariates, has the form

$$Y_{ijk} \sim \text{Poisson}(\mu_{ijk})$$

$$\log(\mu_{ijk}) = \beta_0 + \beta_1 x_{ijk} + \eta_j + \varepsilon_k$$

where the response variable (Y_{ijk}) is the decoy count (indices and known abundances) for the ith sample unit at the jth survey and k^{th} observer, β_0 and β_1 are regression coefficients, and x_{ijk} is an indicator variable for whether the count type is an observer index or a known abundance. Known abundance was used as the reference (dummy) category. η_j and ε_k represent the random effects of survey and observer, respectively. Given this parameterization, the ratio of the mean count index to mean known abundance (the I:A mean ratio) is estimated by exponentiating the β_1 coefficient (e^{β_1}). The I:A mean ratio estimates the percent "visual error" of decoy counts from aerial imagery, where a e^{β_1} value of 0.98, for example, indicates that mean count indices underestimated the true number of decoys by 2%. The value of 0.98, then, is an estimate of detection probability based on Lancia et al. (2005)

$$\hat{p} = c/N$$

where $\hat{p}$ is the estimated probability of detection, c represents the repeated sampling of indices, and N is the true abundance. I expanded the basic model structure to include two environmental covariates: habitat type and cloud cover. Both of these covariates were treated as categorical predictors in the model, with habitat type having two levels (spoil

island and cypress-tupelo swamp) and cloud cover having two levels (0–50% and 51–100%). Spoil island and 0–50% were used as reference (dummy) categories for habitat type and cloud cover, respectively. I investigated the effect of these covariates on detectability by including a full interaction term among count type, habitat type, and cloud cover. This expanded model has structure

$$Y_{ijk} \sim \text{Poisson}(\mu_{ijk})$$

$$\log(\mu_{ijk}) = \beta_0 + \beta_1 x_{ijk} + \beta_2 y_{ijk} + \beta_3 z_{ijk} + \beta_4 x_{ijk} y_{ijk} + \beta_5 x_{ijk} z_{ijk}$$

$$+ \beta_6 y_{ijk} z_{ijk} + \beta_7 x_{ijk} y_{ijk} z_{ijk} + \eta_j + \varepsilon_k$$

where β_{2-7} are additional regression coefficients, y_{ijk} is a categorical variable for habitat type, and z_{ijk} is a categorical variable for cloud cover. Given this expanded parameterization, I estimated the effect of habitat type on detectability by examining the coefficients β_1 and β_4, where $e^{\beta_1 + \beta_4}$ is the I:A mean ratio ($\hat{p}$) for cypress-tupelo habitat. In a similar fashion, I estimated the effect of cloud cover on detectability by $e^{\beta_1 + \beta_5}$ and the interaction between habitat type and cloud cover by $e^{\beta_1 + \beta_4 + \beta_5 + \beta_7}$. Interaction effects can be viewed as representing differences in $\hat{p}$ among levels of the predictors. It should be noted that the effect of habitat type was only included for models of heron decoys.

GLMMs were built and analyzed in the lme4 package for program R (R Version 3.4.0, www.r-project.org, accessed 4 August 2017). In GLMMs, random effects are assumed to be normally distributed (Pinheiro and Bates 2000, Bolker et al. 2009). I created diagnostic plots of random effects to test for normality by using the "qqmath" function in R package lattice. In addition, generalized linear models need to be checked for inflated variance (i.e. over-dispersion). Over-dispersion was assessed with the following equation:

$$\hat{c} = D/\mathrm{df}$$

where $\hat{c}$ is the dispersion parameter, D is the residual deviance, and df is the residual degrees of freedom. When $\hat{c}$ is roughly equal to 1, there is considered to be little to no over-dispersion apparent in the model (Burnham and Anderson 2002). I tested for significant inflated variance by using a Pearson chi-squared (χ^2) test (McCullagh and Nelder 1989).

To examine variability among observer counts, the expanded model was tested against a reduced model without the random effect of observer, by likelihood ratio test (LRT; Bolker et al. 2009). The LRT is a conservative test, so I halved the P-value (Pinheiro and Bates 2000). I then ran a model selection analysis to determine the influence of habitat type and cloud cover on detectability. I used the chosen global model (with or sans the observer random effect) to estimate 6 candidate models, which included all possible interactive effects between count type (x_{ijk}) and the environmental covariates, as well as a null model. Global models for skimmer and tern decoys did not include the effect of habitat type, so only 3 candidate models were estimated. I selected the model with the lowest AIC when all other models had a ΔAIC > 2.00 (Akaike 1974, Burnham and Anderson 2002). I used AIC instead of AIC$_c$ (corrected for small sample size) because AIC$_c$ requires estimating the degrees of freedom (df), for which there is no standard method in GLMMs (Bolker et al. 2009). If models were competing (ΔAIC $\leq$ 2.00), I chose the most parsimonious. I calculated marginal and conditional R^2 values for each model to assess goodness-of-fit (Nakagawa and Schielzeth 2013). Marginal R^2 is a measure of fit (i.e. the proportion of variance explained by the model) for fixed-effects, while conditional R^2 assesses fit for the fixed and random effects combined (Nakagawa

and Schielzeth 2013). Inference of fixed-effect predictors in GLMMs can be challenging, so I followed guidelines by Bolker et al. (2009). I specified a Laplace approximation for estimating model parameters, and used bootstrapped confidence intervals and Wald *Z*-tests to test hypotheses of the fixed-effects. I established the cut-off for statistical significance ($\alpha = 0.033$) using false discovery rate (FDR) to correct for multiple comparisons (Benjamini and Hochberg 1995). I used the "bootMer" function in R package lme4 to get parametric bootstrapped (PB) 96.7% confidence intervals for each model parameter, in which 200 iterations were specified. If 96.7% CIs did not include 1 (i.e., imperfect detection), I reported the corresponding *z*- and *P*-values from the Wald *Z*-test. I obtained bias correction factors by inverting the appropriate model coefficients ($1/\hat{p}$) and their corresponding CIs.

I further examined observer accuracy and variation by calculating the mean percent error (Frederick et al. 2003) for each observer and decoy type by $([i - a]/a) \times 100$, where i is the observer index count and a is the known abundance. Due to high variability among and within observers, I investigated using the double-sample method for estimating abundance and correcting for observer bias. I selected two observers at random for each decoy type in each habitat, and calculated an abundance estimate for each image (quadrat as the sample unit). Estimates were calculated using Chapman (1951)

$$\hat{N} = \left[\frac{(n_1 + 1)(n_2 + 1)}{m + 1}\right] - 1$$

where n_1 is observer 1, n_2 is observer 2, and m is the number of decoys counted by both observers. I compared double-sample estimates to the known density of each image by linear regression in program R, where Chapman estimates were modeled as the response

and known density as the predictor. Quadratic relationships between the Chapman estimate and known density were tested against a linear-only model by LRT. I tested null hypotheses that the slope and intercept were equal to 1 and 0, respectively, which is the expected outcome if the double-sample method estimated true abundance as group size increases. I reported R^2 values for models with a single explanatory variable and R^2_{adj} for models that included both the quadratic and linear terms. Model residuals were checked for normality and homoscedasticity with diagnostic plots in R. Models did not violate any assumptions, so no data transformations were conducted.

RESULTS

I conducted 12 total UAS surveys at EFS and TRB. Detection probability across all decoy types ranged between 54–109% for EFS surveys (n = 5), and 9–29% for TRB surveys (n = 7). Observers occasionally overcounted or falsely detected decoys, thus allowing for a possible detection probability above 100%. Canopy cover at nests in the TRB ranged between 36.87–99.09%, and cloud cover during all flights at both EFS and TRB ranged between 0–100%.

Visibility bias

Black skimmer decoys.–At EFS, black skimmers (n = 60) were the only decoy type with a detection probability significantly lower than 1 (e^{β} = 0.54; PB, 96.7% CI [0.44, 0.63]; Wald, z = –7.425, $P \leq 0.001$; Table 1). The effect of cloud cover approached significance (Wald, z = –1.93, P = 0.054), but following parametric bootstrap of CIs, it was not influential on the detectability of skimmer decoys (PB, 96.7% CI [0.21, 1.03]). The random effect of observer was influential on the total explained variance of the global model for skimmers (LRT, χ_1^2 = 2.59, P = 0.050). The global model, which

contained an interaction between count type and cloud cover, was not significantly over-dispersed and seemed to fit the data well ($\hat{c}$ = 1.13; Pearson, χ_{54}^2 = 52.5, P = 0.531; R^2_m = 0.56, R^2_c = 0.83). I selected the global model based on AIC score.

Tern decoys.–Detectability of tern decoys (n = 60), gleaned from the 2nd best competing model, was not significantly different from 1 (e^{β} = 1.02; PB, 96.7% CI [0.90, 1.15]; Wald, z = 0.356, P = 0.722). Fixed effects, including that of count type, seemingly had no influence on tern decoy detection since the null model was selected. The random effect of survey explained 92% of the variance in the null model. Candidate models for tern decoy detection did not include a random effect for observer (LRT, χ_1^2 = 0, P = 0.500). The global model was under-dispersed ($\hat{c}$ = 0.19; Pearson, χ_{54}^2 = 10.4, P = 1.000), and so was the selected (null) model (Table 2). This level of under-dispersion, in addition to an estimated detection probability close to 100%, suggests that there was little variation in the ability of observers to correctly detect and identify nearly all tern decoys. Furthermore, there was no statistical difference between known abundances and observer counts.

White-plumaged heron decoys.–Detectability of white-plumaged heron decoys (n = 144) differed drastically between sites and levels of cloud cover. At EFS, detectability was not different from 1 (e^{β} = 1.09; PB, 96.7% CI [0.89, 1.31]; Wald, z = 1.095, P = 0.274), but decreased significantly by 71% at TRB (e^{β} = 0.29; PB, 96.7% CI [0.18, 0.46]; Wald, z = –9.652, $P \leq 0.001$). Additionally, when cloud cover was >50% at TRB, white-plumaged heron decoys had a detection probability of only 9%. However, cloud cover was seemingly not influential on their detectability at EFS (PB, 96.7% CI [0.60, 1.21]; Wald, z = –1.31, P = 0.190). The random effect of observer was influential on the

explained variance of the global model for white-plumaged heron decoys (LRT, χ_1^2 = 3.99, P = 0.022). The global model was selected based on AIC score, and seemed to fit the data well with little over-dispersion (Table 2).

Dark-plumaged heron decoys.–Detectability of dark-plumaged decoys (n = 144) differed between sites, but was not affected by cloud cover. At EFS, detection probability was 95%, which was not significantly different from 1 (PB, 96.7% CI [0.79, 1.18]; Wald, z = –0.746, P = 0.456). However, detectability of dark-plumaged heron decoys decreased markedly by 84% at TRB (e^{β} = 0.11; PB, 96.7% CI [0.06, 0.18]; Wald, z = –13.47, $P \leq 0.001$). The random effect of observer had no influence on the variance of the global model (LRT, χ_1^2 = 0, P = 0.499), so a reduced model was used to build candidates for selection. The global model was slightly under-dispersed and fit the data ($\hat{c}$ = 0.70; Pearson, χ_{134}^2 = 110.72, P = 0.938; R^2_m = 0.68, R^2_c = 0.92). The selected model, based on AIC score, was similar in fit (Table 2) and included the interaction between count type and habitat.

Observer error

With few exceptions, observers tended to undercount rather than overcount. 93.3% of photograph counts for black skimmers (n = 30) were undercounts and 0.7% were overcounts. For tern counts (n = 30), 46.7% were undercounts and 33.3% overcounts. Observers undercounted white-plumaged heron decoys more often in photographic quadrats at TRB than EFS (90% and 43.3%, respectively; n = 72). Similarly, dark-plumaged heron indices (n = 72) were undercounts in 88.1% of quadrats at TRB and 46.7% of quadrats at EFS. Overcounts of white- and dark-plumaged heron decoys, however, occurred 16.2% and 9.5% more often at EFS than at TRB. There was

noticeable variation among and within observers with regard to count error (Figure 4). However, of the four detectability models, only black skimmer and white-plumaged heron detection differed significantly among observers ($LRT_{skimmer}$, $\chi_1^2 = 2.59$, $P = 0.050$; LRT_{heron}, $\chi_1^2 = 3.99$, $P = 0.022$). Of the 1,986 total decoys detected by observers, 37 (1.86%) of these were misidentifications of decoy type.

I used linear regression to determine how well double-sample abundance estimates predicted known abundance for each decoy type at each site. At EFS, known densities of tern and heron decoys explained >84% of the variation in Chapman estimates, based on coefficients of determination (Figure 5). Black skimmer photograph (n = 18) density, however, explained less variation in Chapman estimates ($R^2 = 0.67$, $F_{1,16} = 31.91$, $P \leq 0.001$) and the slope was significantly different from 1, suggesting that estimates were biased low ($\beta_1 = 0.66$, $t = 5.65$, $P \leq 0.001$, 95% CI [0.42, 0.91]). Although 98.9% of the variance was explained by known density ($R^2_{adj} = 0.99$, $F_{2,23} = 1134$, $P \leq 0.001$), tern decoy abundance estimates ($n = 26$) were biased low ($\beta_1 = 0.82$, $t = 15.70$, $P \leq 0.001$, 95% CI [0.71, 0.92]). Tern estimates had a quadratic relationship to known density (LRT, $F_1 = 6.34$, $P = 0.019$), indicating that abundance was overestimated for smaller group sizes (Figure 5). For white-plumaged heron estimates at EFS ($n = 26$), the slope and intercept were not different from 1 and 0, respectively ($\beta_0 = 0.003$, $t = 0.032$, $P = 0.974$, 95% CI [–0.18, 0.18]; $\beta_1 = 0.97$, $t = 32.13$, $P \leq 0.001$, 95% CI [0.91, 1.03]). Similarly, dark-plumaged herons (n = 28) at EFS were estimated without significant bias ($\beta_0 = 0.14$, $t = 0.612$, $P = 0.546$, 95% CI [–0.34, 0.63]; $\beta_1 = 0.96$, $t = 11.93$, $P \leq 0.001$, 95% CI [0.79, 1.12]).

At TRB, Chapman estimates for both heron decoy types had poor accuracy and precision. For white-plumaged decoy photographs (n = 30), slope and intercept was significantly different from 1 and 0 (β_0 = 1.72, t = 4.29, P = 0.002, 95% CI [0.90, 2.55]; β_1 = –0.51, t = –1.54, P = 0.135, 95% CI [–1.19, 0.17]), and most of the variation in Chapman estimates were unexplained by known density (R^2_{adj} = 0.167, $F_{2,27}$ = 3.90, P = 0.032). There was a quadratic relationship between estimates of white herons and known abundance, indicating that estimates were biased high for small group sizes (LRT, F_1 = 4.33, P = 0.047). Estimates of dark-plumaged herons at TRB (n = 19) had the poorest prediction of known density, with only 3.75% of the variation in Chapman estimates being explained by known numbers ($F_{1,17}$ = 0.66, P = 0.427). Estimates were biased significantly low for dark heron decoys (β_1 = 0.07, t = 0.814, P = 0.427, 95% CI [–0.11, 0.24]).

DISCUSSION

I found that photographic counts of white herons, dark herons, and tern spp. on a dredge-spoil island showed no indications of imperfect detection. This result can be extrapolated to include two of the three focal species (gull-billed terns and plumage dimorphic reddish egrets) that nest on islands similar to EFS. Although white decoys had a detection probability slightly greater than 100%, this error was not significant for either terns or herons. Overcounting of conspicuous white-plumaged species has been a source of bias in previous studies as well, but this has typically been attributed to species misidentification (Green et al. 2008, Williams et al. 2008). Coinciding with my own results from the island site, both Sardà-Palomera et al. (2012) and Chabot et al. (2015) found that counts of nesting waterbirds from imagery taken with UAS were highly

accurate when compared to ground counts. Although observers were able to detect and correctly identify these 3 decoy types at EFS, it is unknown whether this can translate into actual waterbird species identification. Dulava et al. (2015) recommended an image resolution of approximately 5 mm/pixel in order to correctly identify waterfowl, which can be achieved with a 12 mega-pixel camera at an altitude >46 m. At a survey altitude of 61 m, the best resolution I obtained was approximately 1.7 cm/pixel, although focal length of the camera may not have been optimal for this height. Black skimmer decoys had a detection probability significantly less than 100% and a bias correction factor that would require a doubling of raw count indices. This result is surprising, given that black skimmers should be easier to detect due to their high-contrast plumage amidst a sand or shell nesting substrate. Chabot and Bird (2012) found high variability in Canada goose (*Branta canadensis*) detection, but low variation with snow geese (*Chen caerulscens*), which they attributed to the high contrast of a white-plumaged bird on a darker background. My results suggest that the high contrast of black skimmers does not make them easier to detect from imagery. A possible explanation for this is that observers could not decipher between skimmer decoys and shadows. High-visibility nesting species (such as geese in the arctic tundra), although typically easier to detect from the air than other species, are not exempt from visibility biases that can affect population trend estimates (Walter and Rusch 1997).

Detectability of both white- and dark-plumaged birds in the forested wetland was biased significantly low, which coincides with similar studies that have investigated the effects of canopy cover on waterbird detection (Frederick et al. 2003, Williams et al. 2008, Pearse et al. 2008). Although I did not specifically include canopy cover in

detectability models (due to increased model complexity), it can be assumed that the low detection of decoys when compared to the island site was mostly due to the visual occlusion of nests. Due to logistics, I was unable to create large and numerous decoy "colonies," a limitation that could potentially explain the low detection probability of heron decoys in the cypress-tupelo habitat. Small aggregations of waterbirds, for example, decrease group detection, but large aggregations increase count error (Laursen et al. 2008, Pearse et al. 2008, Williams et al. 2008). Decoys were almost always placed in nests not at the top of the canopy, which is where larger species like great blue herons choose to nest (Rodgers et al. 2005). Thus, my results for cypress-tupelo habitat should be interpreted as if decoys were surrogates for small-to-medium sized ardeids that typically nest within or underneath the canopy, such as little blue herons, snowy egrets, and cattle egrets. Due to the low (almost zero) probability of detection I estimated for these groups, visibility bias correction factors for decoys in canopied sites should be interpreted with caution. I do not recommend using correction factors as a means to estimate actual colony counts since my colony aggregations were so small, and probably poorly reflect the nature of surveying colonies of moderate-to-large size. They were included, rather, to show the increase in visibility bias associated with different levels of covariates. I included cloud cover as a covariate in the models because I assumed that light level might influence detectability of decoy birds. Cloud cover has influenced detection of sea birds via boat surveys (Dixon 1977) and waterfowl during ground-based counts (Pagano and Arnold 2009), but there is not much literature on its impact during aerial surveys or photographic counts. I found that cloud cover had a significant negative impact on the detection of white-plumaged heron decoys, but only at the canopied

wetland habitat. This suggests that natural light may play a role in the detectability of species that nest within canopy or vegetated structure, but its effect is marginal for sites that allow for high visibility of nests. Conversely, cloud cover had no effect on dark-plumaged decoy bias at the canopied site. This could potentially be explained by the fact that dark heron detectability was close to zero (11%) among all indices, which would make it difficult to detect changes in effect size. Further testing of this environmental covariate with a larger aggregation of birds (or decoys as surrogates) could potentially reveal a greater magnitude of influence. It is interesting that cloud cover approached significance for negatively influencing the detectability of black skimmers, although this may have occurred by chance given the wide variability in confidence intervals (21–103%). It is possible that the loss of shadows (due to increased cloud cover) negatively affected the detection of skimmers, which would coincide with my previous allusion that observers may have had trouble deciphering between decoys and shadows. I would have liked to test additional environmental variables in this study, but opted for simplicity in the model structure given my methods for analysis. Angle of the sun, for example, has a substantial influence on the detectability of whooping cranes (*Grus americana*) during aircraft surveys, where detection probability increases three-fold when the sun is at the observer's back (Strobel and Butler 2014). Even though photographic counts can potentially mitigate the effects of sun glint and survey azimuth, it is possible that shadows can obscure nests at canopied sites, or even increase the detection rate of birds with low-contrast to their environment.

Observer variability was only significant for the detectability of black skimmers and white-plumaged herons. Frederick et al. (2003) found that count bias varied greatly

among experienced biologists when trying to estimate the number of white alfalfa seeds in an aerial survey simulation, but that error was reduced by more than half when colonies were counted from photographs. My results suggest that, even though aerial photography via UAS may reduce bias in certain instances, these counts are not exempt from individual observer bias. Sources of variation for individual observers is difficult to quantify (Frederick et al. 2003), and thus it may be useful for managers to correct for photographic count bias on an individual level. It should be noted, however, that my results suggest a tendency for observers to undercount white- and dark-plumaged herons in canopied habitat and overcount in a higher-visibility dredge-spoil island, potentially due to observers feeling more confident in detecting decoys unobscured by vegetation. Results from the double-sample experiment provided some insight into observer error and variation, especially with regard to tern decoys (Figure 5 [B]). In contrast to the almost perfect detection of terns that was estimated via mixed-effects modeling, a randomly chosen pair of observer counts provided biased estimates of density when using the formula by Chapman (1951). This bias suggested that, at least for those two observers, smaller aggregations of terns (i.e. groups of 0–5) were regularly overcounted. It is possible that the inclusion of zeros in the dataset (due to using quadrat as the sample unit) caused this result, but nonetheless it speaks to the variability among observers. Another way of saying this, would be that capture probability (i.e. detection) was heterogeneous among observers, which caused Chapman estimations to perform poorly (Menkens and Anderson 1988). My results for the double-sample experiment also reveal that estimates of abundance can still be heavily biased after correcting for observer error. Following observer error correction, estimates of black skimmer, white heron, and dark

heron decoys were still biased significantly lower than 100% detection. This suggests that, at least for heron decoys at TRB, visual occlusion of nests was a stronger source of bias than that of individual observer. The Chapman estimator did seem to alleviate some bias for skimmer counts (0.54 detection probability compared to a Chapman regression slope of 0.66).

Overall, my results suggest that imagery taken with a UAS can be used to get unbiased abundance estimates for herons and terns that nest in high-visibility habitats like dredge-spoil islands. Unfortunately, it seems that photographic counts provide biased estimates for black skimmers and herons that nest in cypress-tupelo habitat. Observers varied in their ability to detect some decoy types, so it is important that bias-correcting factors be estimated at an individual level.

MANAGEMENT IMPLICATIONS

Unmanned aerial systems (UAS) are becoming increasingly popular among biologists and managers as a survey tool. Their use, while beneficial in many ways, requires further study due to the importance of accurate and precise abundance estimates at both regional and global scales. My results illustrate the need to employ adaptive management schemes when surveying colonial waterbirds from unmanned aircraft. In areas where nest visibility is occluded, for example, expect that photographic counts of wading birds will be heavily biased low. This could be remedied by supplementing UAS surveys with a perimeter or ground-based count of the colony, or even by using double-counts from ground observer and photograph observer. To improve the accuracy of photographic counts, I suggest conducting UAS surveys on clear and sunny days, automating flight patterns, and using the double-sample method when possible. In general, UAS show

promise as a survey tool for species that nest in high-visibility habitats. I feel that their application can greatly benefit survey regimes, especially when terrain is difficult to traverse or when ground counts cause excess disturbance.

II. DISTURBANCE TO COLONIAL WATERBIRDS CAUSED BY UNMANNED AIRCRAFT SURVEYS

Colonial waterbirds are notably sensitive to human disturbance, in part because of their unique breeding biology and nesting behavior. They congregate in groups during the breeding season to nest, such that a single disturbance event during this period can have negative impacts on multiple individual birds. This has been a bane to waterbird biologists, given the importance of conducting long-term monitoring programs (Wiens et al. 1986, Temple and Wiens 1989, Coulter and Frederick 1997, Kingsford 1999). Nisbet (2000) defined disturbance as any human activity that alters the behavior or physiology of one or more individuals of a breeding colony. This altered state can be elicited by many different forms of intrusion, including that of vehicles, recreationists, and researchers. Thus, monitoring schemes that mitigate the negative impacts of disturbance should be prioritized. Breeding waterbirds have historically been surveyed using on-the-ground bird counts, manned aircraft, or a combination of the two.

Investigator intrusions are those that involve walking near or through a nesting colony. Ground counts are just one form of investigator intrusion, which involves walking within or around a breeding site and counting nests to get an estimate of adult pairs. While typically regarded as the most accurate surveying technique, ground counts can cause significant disturbance (Kushlan 1979, Green et al. 2008). Investigator intrusions cause waterbirds to flush, which can lead to nest abandonment, greater exposure to predation, nest failure from exposure to the elements, and spilled nest contents (Kury and Gochfeld 1975, Anderson and Keith 1980, Burger and Gochfeld 1983). There is considerable variation among species with how adverse effects of

disturbances are manifested. Gulls, terns, skimmers, and alcids, for example, tend to have lowered reproductive success following investigator disturbance (Gillett et al. 1975, Anderson and Keith 1980, Burger 1981, Safina and Burger 1983, Nisbet 2000). Yet, Goering and Cherry (1971) found that investigator disturbance did not adversely affect productivity in a mixed-species heronry. Even if a breeding colony is initially tolerant of human intrusion, nests are more vulnerable to predation by gulls, corvids, and other opportunistic foragers (Tremblay and Ellison 1979, Carney and Sydeman 1999).

The magnitude of adverse effects is likely dependent on specific survey methodology – entering a heronry to count nests is presumably more detrimental than establishing a buffer zone to count from the perimeter (Burger et al. 1995). Frequency of visits can be a causative factor of disturbance as well, with some species being more sensitive to the occurrence of intrusions. Increased frequency of researcher visits caused nest abandonment of newly constructed nests for black-crowned night herons (*Nycticorax nycticorax*) in Québec, Canada, but did not hinder the reproductive success of tricolored herons (*Egretta tricolor*) in the Florida everglades (Tremblay and Ellison 1979, Frederick and Collopy 1989). In an extreme case, repeated nest checks of an Adélie penguin (*Pygoscelis adeliae*) colony decreased hatching success by 35% and chick survival by 72% (Giese 1996). In addition to the deleterious effects, investigator intrusions (e.g. ground counts) come with their own set of logistical constraints. They are time-consuming, laborious, costly, and hardly feasible for colonies of large size and scope.

Aerial surveys, an alternative to ground monitoring, typically cause less disturbance to nesting waterbirds. Kushlan (1979) concluded that fixed-wing aircraft and helicopter overflights did not disturb a heronry because birds returned to the nest within 5

minutes of leaving. In 90% of observations during flights as low as 60 m, birds either had no reaction or merely looked up. Vigilance, however, is still considered to be a behavioral manifestation of disturbance (Nisbet 2000). It is common for waterbirds to look up, scan more, or remain vigilant during aircraft overflights (Williams et al. 2007, DeRose-Wilson et al. 2015). Increased vigilance does not necessarily cause a direct negative impact, but even this seemingly innocuous response to stimuli can divert time and energy away from activities that are needed to increase individual fitness (Gutzwiller et al. 1994, Verhulst and Oosterbeek 2001). Some waterbird species (e.g. least tern [*Sternula antillarum*], common tern [*Sterna hirundo*], gull-billed tern [*Gelochelidon nilotica*], and black skimmer [*Rynchops niger*]) do not react noticeably to either civilian or military aircraft, and thus overflights do not detract from incubation behavior (Hillman et al. 2015). This is not always the case, as Adélie penguins are more prone to abandon nests after helicopter and fixed-wing aircraft surveys (Culik et al. 1990).

Since traditional survey methodology is not without its drawbacks, and there is lack of consensus over which method is best, new survey techniques are frequently tested in the field. Unmanned aerial systems (UAS) have recently garnered attention in regards to wildlife surveying. In previous studies, they have been used to census large colonies of common terns (Chabot et al. 2015), detect roe deer (*Capreolus capreolus*) fawns in meadows with a thermal camera (Israel 2011), monitor breeding pairs of black-headed gulls (*Chroicocephalus ridibundus*, Sardá-palomera et al. 2012), and monitor the nesting status of hooded crows (*Corvus cornix*, Weissensteiner et al. 2015). Unmanned aircraft have a few key benefits that make this new technology appealing to wildlife researchers, such as suitability for fine spatial resolution (Bakó et al. 2014), comparable coverage to

traditional aerial surveys (Bakó et al. 2014), researcher safety, and ability to easily switch out sensors or payloads (Watts et al. 2010). Perhaps the most important potential benefit of using UAS for surveying waterbirds is the reduction of disturbance when compared to other survey methods. Carney and Sydeman (1999) proposed that smaller and quieter aircraft do not disturb colonially-nesting waterbirds as much, and this has been touched upon in subsequent publications concerning UAS (Jones et al. 2006, Chabot et al. 2015). Vas et al. (2015) and Dulava et al. (2015) are the only peer-reviewed publications to date that have specifically investigated waterbird disturbance caused by UAS, but neither study focused on the disturbance of waterbirds during the nesting season.

Due to the recent popularity and availability of UAS among researchers and recreationists, there is a clear need to study the ramifications of their use. Even though the technology is still in its infancy, the FAA forecasts that hobbyist UAS sales will likely more than triple over the next 5 years, from 1.1 million units in 2016 to over 3.5 million units by 2021 (Lukacs and Bhadra 2017). Commercial use of unmanned aircraft is a bourgeoning industry as well, with their use already in the fields of agriculture, photography, construction, real estate, delivery services, and utility inspections. The commercial sector of unmanned aircraft is expected to grow tenfold by 2021, from 42,000 units in 2016 to 420,000 (Lukacs and Bhadra 2017). The speed at which UAS technology has grown presents a problem for regulatory agencies, including those that deal with wildlife-related restrictions (Harrington 2015). Carney and Sydeman (1999) make note of 3 main categories of human disturbance: investigator, ecotourist (e.g. wildlife photography), and recreationist (e.g. beachgoers, anglers). Unmanned aircraft

currently fit into all 3 of these categories, in addition to the potential impact caused by the commercial sector, which highlights the concern for proper protocols and regulations.

My overarching research objective was to evaluate the potential disturbance caused by unmanned aircraft when surveying nesting colonial waterbirds. Specifically, I aimed to (1) compare behavioral effects caused by two types of UAS (a fixed-wing platform with enhanced capability for payload attachment, and an off-the-shelf quadcopter consumer grade model), (2) test the behavioral response for a variety of survey altitudes (at which fine-scale resolution could be achieved), and (3) evaluate disturbance for a host of waterbird species that nest in Texas, including species that are of priority concern to the Gulf Coast Joint Venture.

STUDY AREA

There are four colonial nesting waterbird species currently listed as priority within the Monitoring, Evaluation, and Research Team (MERT) of the Gulf Coast Joint Venture (GCJV): reddish egret (*Egretta rufescens*), black skimmer (*Rynchops niger*), gull-billed tern (*Gelochelidon nilotica*), and little blue heron (*Egretta caerulea*). I selected nesting sites in Texas *a priori* for inclusion of these 4 species. The GCJV is one of 18 joint ventures in the continental US, and its coverage spans coastlines in Texas, Louisiana, Alabama, and Mississippi (Figure 1). Due to variability in nesting strata among these species, it was necessary to choose study sites in two distinct regions of Texas: coastal islands in the lower Laguna Madre and forested wetlands in the lower Trinity River basin. I collected data at four waterbird nesting colonies during the nesting seasons occurring within May–June of 2016–2017. Three of these colonies were mixed-species

(Green Island, East Flat Spoil, and a site on the lower Trinity River in 2016), and one was single-species (Trinity River site in 2017).

Lower Laguna Madre

The Laguna Madre of Texas has a subhumid-to-semiarid subtropical climate, with extreme variability in precipitation (Fulbright et al. 1990). Average rainfall is between 2.6–14 cm per month, with a minimum and maximum recorded annual precipitation of 13.7 and 50.3 cm, respectively (Tunnel 2002). Geographically, the Laguna Madre is a long and narrow hypersaline coastal lagoon that extends roughly 185 km from Corpus Christi Bay to the southern tip of the state (Tunnel 2002). The lagoon is split into two subunits – upper and lower – by a land-bridge that connects to the barrier island spit known as Padre Island (Tunnel 2002). I studied coastal waterbird colonies at Green Island and East Flat Spoil island, situated within Cameron and Kenedy Counties, respectively (Figure 6). Both sites are managed by Texas Audubon.

Green Island (GI), with an area of roughly 12 ha, is one of the few remaining naturally occurring islands on the Texas coastline. It includes one of the largest breeding sites for the reddish egret and roseate spoonbill (*Platelea ajaja*), and is a popular hotspot for naturalists to see neotropical migrants and a host of breeding waterbirds (Pemberton 1922). Historically, the island supported ≥1,000 breeding pairs of reddish egrets (Wilson et al. 2012). Support of an immense breeding colony is likely due to the island's location, size, and dense vegetative structure. GI is thick with native Tamaulipan thornscrub (e.g. huisache [*Acacia farnesiana*], honey mesquite [*Prosopis glandulosa*], and Texas ebony [*Ebenopsis ebano*]), shrubs (e.g. snake-eyes [*Phaulothamnus spinescens*]), and cacti (e.g. ladyfinger cactus [*Echinocereus pentalophus*] and *Opuntia*

spp.). At GI, I collected data on six wading bird species: reddish egret, roseate spoonbill, black-crowned night heron, white ibis (*Eudocimus albus*), tricolored heron (*Egretta tricolor*), and great blue heron (*Ardea herodias*). Tricolored herons and black-crowned night herons were not often seen, so they were excluded from analysis.

East Flat Spoil (EFS) is a 1.2 ha dredge-spoil island situated just north of the Mansfield tidal inlet. Vegetation on the island is mostly comprised of low-lying, highly saline-tolerant species, like sea oxeye (*Borrichia frutescens*), shoregrass (*Monanthochloe littoralis*), sea purslane (*Sesuvium portulacastrum*), and turtleweed (*Batis maritima*). The mix of vegetated area and bare ground attracts a diverse set of ground-nesting waterbirds to the island every year. At EFS I collected data on 5 waterbird species: reddish egret, black skimmer, great blue heron, royal tern (*Thalasseus maximus*), and laughing gull (*Leucophaeus atricilla*). I also collected data on small terns that could not be positively identified during sample periods, but were assumed to be a mix of gull-billed terns, Forster's terns (*Sterna forsteri*), and sandwich terns (*Thalasseus sandivencis*).

Lower Trinity River Basin

The Trinity River basin (TRB) runs from North Texas – its headwaters just west of Fort Worth – all the way to the Galveston Bay system on the Gulf of Mexico (Dahm et al. 2011). The lower portion of the basin has a humid subtropical climate, and is predominantly forested (Dahm et al. 2011). Cypress-lined bayous in the lower Trinity watershed provide ideal habitat for canopy-nesting waterbird species, including the little blue heron. Due to restrictions on UAS operation at the time of this study, I was unable to survey colonies at Trinity River NWR proper, and instead utilized colonies on Josie Lake adjacent to refuge property (Figure 6). The area is dominated by bald cypress trees

(*Taxodium ascendens*), tupelo (*Nyssa* spp.), button bush (*Cephalanthus occidentalis*), and a diversity of bottomland hardwood species.

Many waterbird species use the lower TRB to nest and forage. However, at the time of my study I was only able to gather behavioral data on four species in this region. One of my target species, the little blue heron, unfortunately did not nest at any colony sites in my study area like they had in previous years. In 2016, I collected data on two species: great egret (*Ardea alba*) and snowy egret (*Egretta thula*). Anhingas (*Anhinga anhinga*) and cattle egrets (*Bubulcus ibis*) also nested in this colony, however very few were observed during sample periods so were omitted from analyses. In 2017, I collected data on a great egret colony that occupied the same site as the year prior.

METHODS

UAS surveys were conducted between 0900 and 1600 hours on clear to partly cloudy days, with wind conditions ranging between 2.5–23.1 km/hour. I flew surveys parallel to the ground in a back-and-forth transect pattern, and did not approach birds from any other angle (Vas et al. 2015). Strip transects allow for a series of overlapping photos to be captured, which can later be processed into a mosaicked, georeferenced orthophoto (Bolstad 2016). A single georeferenced aerial photo is what managers would likely use to estimate the number of breeding adult birds in a colony, and I wanted my methods to reflect a workable model for using unmanned aircraft to survey waterbirds. I tested two UAS platforms – a fixed-wing craft and a consumer-grade quadcopter – and flew a range of altitudes between them. Length of surveys (for both UAS platforms) ranged between approximately 2–7 minutes, which was dependent on survey altitude and size of the site.

The fixed-wing, dubbed the Minion 2.0 (AggieAir, Utah State University, Logan, UT; Figure 3), is a lightweight (10.2 kg) platform with a 60-minute battery life and cruising speed of 80 km/hour. It is launched via pneumatic rail, is capable of fully-autonomous flight, and requires a crew of at least 3 on-the-ground technicians for safe operation. Minion flights were pre-programmed the day prior to surveys. The Minion is equipped with a 12 mega-pixel RGB camera (Lumenera Lt1265r, Lumenera, Ottawa, ON) and thermal infrared camera (ICI 9640, Infrared Cameras Inc., Beaumont, TX), although only the RGB camera was used for this study. It was launched 8 km away from the nesting colony, brought to altitude, and flown 300 m above ground (AGL) for the first survey. After a full survey at 300 m, it was programmed to drop in altitude away from the colony to 200 m AGL and complete another survey. I selected these altitudes based on recommendations by the camera manufacturer and the developers of the platform, to obtain optimal resolution for imagery capture (4 cm/pixel when at 200 m AGL).

I also flew surveys with a consumer-grade quadcopter UAS (Phantom 3 and Phantom 4, DJI, Shenzhen, Guangdong, China) that came equipped with a 12.4 mega-pixel RGB camera. The Phantom is operated by a single technician, can be launched from almost any surface, and is flown via a minimalist ground station (remote control and iPad device [Apple Inc., Cupertino, CA]). Recent software packages allow for autonomous flight (Ground Station Pro, DJI), but only one survey in 2017 was conducted using such software. Most surveys were flown manually, but still in back-and-forth transects. I used the same flight procedure for the Phantom as for the Minion, except with a series of decreasing survey altitudes of 122, 91, 61, and 46 m AGL. I chose these altitudes because it was expected that within this range I could get optimal resolution of

aerial photographs. The Phantom was brought to altitude ≥250 m away from the nesting colony. Essentially, I wanted to compare the effect of each platform when flown at their respective optimal survey altitudes.

Behavioral observations

Behavioral observations can be used to quantify an animal's biological response to stimuli (Altmann 1974, Clemmons and Buchholz 1997). I placed 2–3 video cameras at the periphery of each nesting colony, with the exception of Green Island, to capture any behavioral changes that might occur during UAS overflights. At GI, I positioned cameras in an established bird blind. I placed cameras in areas with an open field of view and selected a random azimuth for adjusting position of the camera lens. If camera frame happened to fall outside of the nesting colony or within the same subset of birds that another camera was recording, I selected a different azimuth. It was assumed that camera setups (i.e. tripods) would not interfere with the behavior of nesting waterbirds, so they were placed 50–100 m away from the colony. Set-back distances vary among waterbird species (Erwin et al. 1989), so all researchers retreated to ≥250 m away from the colony after video cameras were set to record. I used a laser range finder to confirm distance (Nikon Aculon IK-714141, Nikon Inc., Tokyo, Japan). I should not have created excess disturbance to the birds, as the farthest set-back distance found in the literature was 178 m for black skimmers (Erwin et al. 1989). I waited ≥20 minutes following video camera setup to commence surveys, which I assumed would allow birds to get back to a baseline level of disturbance prior to UAS surveys.

Scan sampling, an instantaneous behavioral observation of a group of animals, captures the behavior of the group by placing each potential behavior of an individual in

discrete categories (Altmann 1974). I chose 5 behavioral categories *a priori* to assess the extent of disturbance caused by UAS, (1) no reaction, (2) vigilance, (3) wing flapping, (4) moving away from the nest, and (5) flush. These categories are similar to what Giese and Riddle (1999) and Vas et al. (2015) used when measuring behavioral responses in emperor penguins (*Aptenodytes forsteri*) and 3 wader spp., respectively. Since examining the minutia of waterbird nesting behavior was not the aim of this study, these categories were intentionally kept broad. One caveat to category 5 is that I tallied birds flying through camera frame as well as flushing from a nest within frame. I reasoned that if UAS surveys caused more birds to flush, there should be more birds flying within camera frame during those sample periods. Video playback was observed without sound and without the aircraft in frame, so as not to bias observations when a flyover occurred (i.e., the observer was not privy to the occurrence of a survey). One observer gathered data from video playback, to eliminate any variation that would arise among multiple observers. Every minute, the observer slowed down video playback to one-third speed and tallied the behavior and species of each bird in frame for a total of 4 seconds. If a bird exhibited more than one behavior during the sample period, the more severe behavioral reaction was recorded (e.g. if a bird was vigilant and then flushed, "flushed" was recorded).

Statistical analysis

All analyses were done in program R (R Version 3.4.0, www.r-project.org, accessed 14 June 2017). For each survey, I included sample periods that spanned 5 minutes before and after UAS overflights. The effect of survey altitude on colony behavior could then be compared to a baseline (control), similar to methods used by Hillman et al. (2015).

Flight altitude was treated as a categorical variable with 8 levels, within which I included two additional levels representing pre- and post-flight sample periods. I used the pre-flight period as a reference (dummy variable) category to compare with the other treatment levels. I only used experimental sample periods in which the aircraft was overhead of the colony; all other observations were omitted from analysis. The coloniality of nesting waterbirds is known to have anti-predator advantages, so it stands to reason that birds often react to nearby conspecifics or heterospecifics (Rodgers and Smith 1995). Thus, instead of analyzing behavior of focal individuals, I assessed the magnitude of behavioral changes within a colony. I used the aforementioned categories of disturbance (2–5) that were tallied during each sample period as response variables, but only built 2 sets of models because of the rareness of some behaviors. Vigilance and flush behaviors were modeled, and the rest were included as summary statistics. Tallies were pooled between cameras, since there were 2–3 cameras used per colony. Logistical constraints prevented us from using 3 cameras on certain colonies (i.e. vegetation obstructing view).

Because of the likely variation between mixed-species waterbird colonies, I used a repeated measures study design where "colony" was considered a random effect. I used generalized linear mixed models (GLMMs) with a Poisson error distribution and log-link function to accommodate a repeated measures design and count-metric response variable. Models with a negative binomial error distribution were also considered, but were not used because over-dispersion was not present in the Poisson models. GLMMs were built and analyzed in the lme4 package for program R. In GLMMs, random effects are assumed to be normally distributed (Pinheiro and Bates 2000, Bolker et al. 2009). I

created diagnostic plots of random effects to test for normality by using the "qqmath" function in R package lattice. In addition, generalized linear models need to be checked for inflated variance (i.e. over-dispersion). Over-dispersion was assessed with the following equation:

$$\hat{c} = D/\mathrm{df}$$

where $\hat{c}$ is the dispersion parameter, D is the residual deviance, and df is the residual degrees of freedom. When $\hat{c}$ is roughly equal to 1, there is considered to be little to no over-dispersion apparent in the model (Burnham and Anderson 2002). I tested for significant inflated variance by using a Pearson chi-squared (χ^2) test (McCullagh and Nelder 1989). Because samples were taken over a period of time (every minute), I tested for temporal autocorrelation of all selected models. I used the "acf" function in R package stats to create diagnostic plots of the Pearson correlation coefficient (ρ), and if all lag periods were not significant (i.e. within 95% CIs) I considered observations to be independent (Venables and Ripley 2002). Even so, I included a fixed-effect predictor for lag 1 autocorrelation in my global models, which was just each response observation lagged by 1.

I built a full model for each response variable (vigilance and flush) with structure:

$$Y_{ij} \sim \mathrm{Poisson}(\mu_{ij})$$

$$\log(\mu_{ij}) = \beta_0 + \beta_1 w_{ij} + \beta_2 x_{ij} + \beta_3 y_{ij} + \beta_4 z_{ij} + \eta_j + \eta_{ij}$$

Where Y_{ij} is the count of behaviors observed, μ_{ij} is the conditional mean, and $\beta_{1\text{–}4}$ are fixed-effect coefficients for the ith sample period at the jth colony site. Model predictors include w_{ij} = treatment levels, x_{ij} = platform, y_{ij} = colony size, and z_{ij} = lag 1. The random effect of colony (η) has random terms for intercept and slope. I chose to include

colony size as a predictor in the model due to previous research that suggests colony size influences disturbance of nesting birds (Rodgers and Smith 1995). I used the total number of birds within camera frame as a surrogate for colony size.

The full model was tested against a reduced random intercepts-only model by likelihood ratio test (LRT; Bolker et al. 2009). The LRT is a conservative test, so I halved the *P*-value (Pinheiro and Bates 2000). I then used that full or reduced model to estimate 16 candidates for selection. I built candidate models with all possible combinations of the above fixed-effect predictors, including a null model, and selected the best model in this set with the lowest AIC when all other models had a $\Delta AIC > 2.00$ (Akaike 1974, Burnham and Anderson 2002). I used AIC instead of AIC_c (corrected for small sample size) because AIC_c requires estimating the degrees of freedom (df), for which there is no standard method in GLMMs (Bolker et al. 2009). If models were competing ($\Delta AIC \leq 2.00$), I chose either the most parsimonious (i.e. with the fewest fixed-effect predictors), or a model with nonsignificant temporal autocorrelation and over-dispersion. I calculated marginal and conditional R^2 values for each model to assess goodness-of-fit (Nakagawa and Schielzeth 2013). Marginal R^2 is a measure of fit (i.e. the proportion of variance explained by the model) for fixed-effects, while conditional R^2 assesses fit for the fixed and random effects combined (Nakagawa and Schielzeth 2013). Inference of fixed-effect predictors in GLMMs can be challenging, so I followed guidelines by Bolker et al. (2009). I specified a Laplace approximation for estimating model parameters, and used bootstrapped confidence intervals and Wald *Z*-tests to test hypotheses of the fixed-effects. I established the cut-off for statistical significance ($\alpha = 0.033$) using false discovery rate (FDR) to correct for multiple comparisons (Benjamini

and Hochberg 1995). I used the function "bootMer" in R package lme4 to get parametric bootstrapped (PB) 96.7% confidence intervals for each model parameter, in which 200 iterations were specified. If confidence intervals did not include 0, I deemed the parameter to have significant influence and reported *z*- and *P*-values from Wald-*Z* tests. To determine the influence of my predictors, I exponentiated the model coefficients, however, the exponentiated model coefficients are the influence of a predictor given all other parameters in the model are held constant, including random effects.

RESULTS

I had a total of $n = 209$ sample observations for 9 UAS overflights. Only 1 successful Minion flight was conducted (at EFS), with 8 flights from the Phantom quadcopter. Total number of birds in camera frame ranged from 18–139, and the number displaying a behavioral response ranged from 0–29. There were no "dread" flights exhibited by nesting birds during UAS surveys. "Dreads" are when all or most individuals in a colony flush from the nest, circle, and then land (Palmer 1941, Erwin 1989). For species pooled together, across all sample periods, and as a proportion of total birds in camera frame, a mean of 5.7% displayed vigilance, 0.7% walked off the nest, 2.4% flapped their wings, 7.5% flushed, and 16.4% displayed 1 of the 4 behavioral responses. It should be noted, however, that the low percentage of birds that walked off nest was likely due to this behavior being rare among herons. Each species reacted differently to UAS flights in terms of behavioral reactions, with occasionally a noticeable positive trend in the magnitude of such reactions when decreasing survey altitudes (Figure 8–10). For example, the proportion of laughing gulls that flushed appeared to increase as UAS surveys decreased in altitude (Figure 9 [D]). And yet, other species like great blue herons

did not exhibit flushing behavior often during any sample periods, but were generally more vigilant (Figure 8 [A, D]).

Vigilance behavior

My global model for vigilance (contained all 4 fixed predictors) included both random slopes and intercepts terms (LRT, $\chi_2^2 = 4.88$, $P = 0.043$). It had low over-dispersion and seemed to fit the data well ($\hat{c} = 1.24$; Pearson, $\chi_{195}^2 = 205.66$, $P = 0.286$). Of my 16 candidates based on the global model, a selection analysis yielded two competing models ($\Delta AIC \leq 2.00$) for estimating the number of vigilant waterbirds (Table 3). From these two I chose the most parsimonious, which included the predictors of treatment, total number of birds in camera frame, and lag 1. The chosen model had nonsignificant over-dispersion (Pearson, $\chi_{196}^2 = 206.11$, $P = 0.296$), with 51% of the variance explained by the fixed and random effects combined and 37% accounted for by the fixed effects alone. There was also evidence of variation among nesting colonies ($SD_{intercept} = 0.20$, $SD_{slope} = 0.24$), but the magnitude of influence from some of the fixed effects outweighed this variation (Table 4). My selected model did not have significant temporal autocorrelation (Pearson correlation coefficient, $\rho_{lag1} = 0.09$).

Treatment levels were compared against a baseline control (the pre-flight period). Both of the Minion fixed-wing treatments had similar effects, with 18% decreased vigilance for 300 m and 7% decreased vigilance for 200 m. Even though there was decreased vigilance during these periods when compared to the baseline, the influence was not significant for 300 or 200 m (Table 4). For overflights with the Phantom quadcopter, there were marked differences in the magnitude of vigilance when compared to the pre-flight period. Surveys at 122 m had a 43% increase in vigilant behavior,

however this increase was not significant following the FDR cutoff for α (PB, 96.7% CI [–0.01, 0.72]; Wald, $z = 2.03$, $P = 0.042$). Quadcopter surveys at 91, 61, and 46 m all had a significant positive influence on waterbird vigilance when compared to the baseline control. There was a 72% rate of increase in vigilance for 91 m surveys, a 119% increase for 61 m, and a 118% increase for 46 m (Table 4). The post-flight period was also significantly different from the baseline, with a 53% increase of vigilance (PB, 96.7% CI [0.21, 0.80]; Wald, $z = 2.692$, $P = 0.007$).

Lag 1 had a significant positive influence on vigilance by 5% for every one unit change, but visual inspection of diagnostic plots showed nonsignificant temporal autocorrelation in the model, so sample periods were assumed independent of one another. Flush behavior increased by 1% for every one unit change in group size within camera frame.

Flushing behavior

The global model for flush reactions again contained the four predictor variables, but unlike the model for vigilance, was a reduced model with a random intercept-only term (LRT, $\chi_2^2 = 0.36$, $P = 0.42$). It was slightly under-dispersed, but not significant enough to warrant concern ($\hat{c} = 0.80$; Pearson, $\chi_{197}^2 = 166.65$, $P = 0.943$). I had four competing models, and selected the most parsimonious with nonsignificant temporal autocorrelation (Pearson correlation coefficient, $\rho_{lag1} = 0.10$; Table 3). Two of the four competing models did have temporal autocorrelation ($P < 0.05$), so it was assumed that sample periods in those model fits were not independent. Fixed effects of lag 1 and total birds in frame were the only predictors included in the chosen model. Like the global model, it was under-dispersed, but not significantly so (Pearson, $\chi_{205}^2 = 174.26$, $P = 0.942$). Only

5.5% of the variance was explained by the fixed effects, but 89.7% of the variance was explained by fixed and random effects combined. This indicated that the random effect of colony was highly influential on the magnitude of waterbird flushing behavior ($SD_{intercept}$ = 1.97, Figure 11).

Lag 1 did not have a significant effect on waterbird flushing behavior (PB, 96.7% CI [–0.40, 0.36]; Wald, z = 0.81, P = 0.417), and neither did the total number of birds in camera frame following the bootstrap of confidence intervals (PB, 96.7% CI [–0.74, 1.18]). Treatment was not included in any of the competing models, suggesting that survey altitude had no effect on flush behavior. For the second-best model without significant temporal autocorrelation (Table 5), platform type increased flushing behavior by 16% when birds were subjected to the Phantom UAS (e^{β} = 1.16), but this effect was insignificant (PB, 96.7% CI [–0.84, 0.88]; Wald, z = 1.61, P = 0.108).

DISCUSSION

My research objectives were to (1) compare the influence of two UAS platforms on waterbird behavior, (2) test several altitudes between platforms, and (3) survey mixed-species colonies in Texas. For objective 1, I found that one platform influenced vigilant behavioral reactions while the other did not. Both survey altitudes flown with the Minion had marginal effects on vigilance, while 3 of the 4 survey altitudes flown with the Phantom increased vigilance significantly. Previous research suggests that smaller aircraft may mitigate behavioral responses and their effect thereof (Carney and Sydeman 1999, Mulero-Pázmány et al. 2017). This was not the case in my study, but this discrepancy is likely due to the influence of survey altitude. The Minion, although a larger platform, is hardly detectable from 300 and 200 m, while the Phantom could be

easily heard and seen from 122 m. Carney and Sydeman (1999) also make note of noise being a possible causative factor in disturbance. At least for military overflights and manned aircraft surveys, fixed-wing aircraft evoke less of a behavioral response than rotary wing craft in most cases (Culik et al. 1990, Plumpton 2006, Carney and Sydeman 1999). This could be true for UAS as well, but the cause of this discrepancy has not been tested for unmanned aircraft.

Flush behavior, however, was not influenced by either platform. In fact, the magnitude of flush behavior was largely determined by colony, as most of the variance in my model was explained by the random effect. The dissimilarity between TRB 2017 and the other colonies should not be understated, because it could have greatly influenced model selection and inclusion of parameters. TRB 2017 only had 24 adult great egrets nesting at the time of surveys, with flushing birds captured on video twice. This is vastly different from both GI and EFS, which had an estimated >1,000 nesting birds at the time of surveys. Still, there was no significant effect of platform on flush behavior, which is promising. Sardà-Palomera et al. (2012) found similar results for black-headed gull (*Chroicocephalus ridibundus*) surveys at 30–40 m. Surveys that are conducted below 30 m, however, can cause nonbreeding waterbirds to flush from their roost (Dulava et al. 2015). Flushing from the nest during breeding season can spell disaster for waterbirds, as dense colonies are more prone to predation (Becker 1995, Brunton 1999). Exposed nest contents are vulnerable to predators and ambient temperature, increasing the likelihood of failure and abandonment (Kury and Gochfeld 1975, Anderson and Keith 1980, Burger and Gochfeld 1983). Results regarding the Minion fixed-wing UAS should be interpreted with caution because I only did a single overflight (at EFS), however I felt it

was important to include in the model even given the small sample size.

For objective 2, vigilance seemed to increase in magnitude as survey altitudes with the Phantom decreased, as evidenced by parameter coefficients. Rümmler et al. (2015) found similar results when surveying a colony of Adélie penguins with an octocopter UAS, although surveys in their study ranged from 20–50 m. While lower altitude surveys seem to increase vigilance, the actual effect of this behavioral response is not known. If vigilance is prolonged, it could detract from normal nesting behavior enough to reduce individual fitness, but this has not been explicitly tested with regard to UAS surveys (Gutzwiller et al. 1994, Verhulst and Oosterbeek 2001). The post-survey period, when compared to the baseline control, showed evidence of increased vigilance. This result suggests that vigilant responses were heightened up to 5 minutes after surveys were completed, but it is unknown how long this increased response lasted.

In my selected vigilance model, there was an inverse correlation between slope and intercept variance (–0.45), which means vigilance behavior rates decreased across sample periods. This suggests the possibility of short-term habituation to UAS surveys, even within a few hours of use, which coincides with observations made by Chabot et al. (2015) regarding UAS surveys of common terns. Rümmler et al. (2015) concluded that habituation was not apparent during UAS surveys, although their study involved different colonial nesting species from this study. Survey altitude also had no effect on flush behavior, suggesting that UAS surveys with the Phantom could be employed to gather high resolution imagery without immediate risk to nesting success, at least within the bounds of 46–122 m AGL.

In regard to objective 3, species varied in how they reacted to unmanned aircraft

surveys. Walking off the nest and wing flapping were not often seen among all species, but there seemed to be differences among species for flush behavior. There was a clear increase in flushing behavior for laughing gulls, for example, so my results should be viewed and interpreted with caution. Species were pooled due to logistical constraints as well as to increase statistical power, and so they reflect the nature of mixed-species waterbird colonies as a whole, not the idiosyncrasies of each species. I acknowledge the need and importance of testing hypotheses within species mentioned, as it is possible that models did not detect differences due to pooling. Habituation and tolerance varies among waterbird species, and responses to stimuli are species-specific (Mueller and Glass 1988, Nisbet 2000). The deleterious effects (e.g. nest abandonment, adult mortality, and reduced reproduction) caused by a behavioral response is species-specific as well (Stillman et al. 2007, Hillman et al. 2015). Breeding biology can dictate whether birds are prone to high disturbance, which can possibly explain the gradient of tolerance and habituation among species. For example, herons are known to establish rookeries in highly urbanized areas, and can thus tolerate high anthropogenic disturbance (Parsons and Burger 1982, Parnell et al. 1988). However, they are sensitive to disturbance during specific nesting phases and when people enter the colony (Parnell et al. 1988).

Testing unmanned aircraft for use in wildlife research requires more extensive study. The most important question to ask is what kind of effect an increased behavioral response has on individual fitness. There are many facets to anthropogenic disturbance that were not included in this study, such as incubation phase. At least with waterbird species, the egg-laying period (and early incubation phases) is when birds are most sensitive to disturbance (Hunt 1972, Gillet et al. 1975, Tremblay and Ellison 1979, Safina

and Burger 1983, Parnell et al. 1988). However, this should be examined on a per-species basis (Van de Voorde 2015). Survey frequency should also be investigated, because this can discourage late-nesters from initiating in a colony (Tremblay and Ellison 1979). If unmanned aircraft were to be used in weekly or monthly monitoring schemes, frequency could be an important factor for managerial decisions. Time length of survey can potentially have an effect on waterbird behavioral responses, although in this study no surveys lasted >7 minutes when the UAS was directly overhead of the nesting colony. Managers should be aware that surveys will last longer at larger sites in which low-altitude flights are employed. For small sites like EFS (1.2 ha), surveys lasted between 2–4 minutes for altitudes of 46–122 m. Survey time is also affected by the use of automated flight patterns when compared to manually-operated flights. Automated flights had a noticeable increase in survey time. For example, at TRB, the automated flight plan took 3.8 times as long (23 minutes, including all sample periods that were not overhead of the bird colony) than the manual flight (6 minutes).

In summary, waterbird colonies show a significant increase in vigilant behavior when surveying with a Phantom UAS between 46–91 m. The Minion fixed-wing UAS seemingly does not increase any behavioral reaction at its suggested survey altitudes, which is likely due to improbable detectability of the craft from ground level. Severe behavioral reactions (i.e. flushing from the nest) are not increased for either platform and at any of the tested altitudes, which implies that UAS surveys do not impact the colony as much as on-the-ground transects. More testing of flush reactions need to be conducted on a per-species basis, however, because my results suggest differentiation among species. My methods could likely be applied in a similar manner to get flush initiation

distances (FID) for each species of interest, which would more accurately estimate disturbance magnitude of UAS. Unmanned aircraft are becoming increasingly popular for recreation, research, ecotourism, and commercial purposes, which highlights the need to continue investigating their impact to wildlife.

MANAGEMENT IMPLICATIONS

The impact of increased vigilance was not tested in this study. However, if waterbird surveys are conducted within a reasonable time frame and between 122–46 m with a quadcopter UAS, the impact of vigilance is likely negligible. Surveying with a high-altitude fixed-wing UAS could be employed without any worry of deleterious effects, but their use comes with logistical issues and data quality trade-offs (Appendix A). The Minion needs one-quarter mile of takeoff and landing space, a crew of 3 technicians, and its cost at the time of this study was approximately 50,000USD. In contrast, the Phantom quadcopter requires only one pilot, is flexible in terms of take-off and landing (I regularly initiated take-off from a boat), and is a fraction of the cost. In addition, imagery resolution was far superior for the Phantom aircraft. For the Phantom, survey altitudes at 122, 91, 61, and 46 m yielded approximately 3.3, 2.5, 1.7, and 1.3 cm/pixel resolution, respectively, while Minion surveys yielded approximately 7.4 and 4.0 cm/pixel resolution for altitudes at 300 and 200 m. To identify species from aerial imagery, however, a resolution of about 0.5 cm would be required (Dulava et al. 2015). For monitoring efforts that wish to obtain abundance estimates, the authors recommend using unmanned aircraft similar to the Phantom, even given the propensity to increase vigilant behavior. For recreational use, or in situations where fine-scale resolution is not necessary, I recommend overflights at ≥122 m to minimize disturbance.

Table 1. Detection probability and bias correction factors of waterbird decoys. Data collected at East Flat Spoil and Trinity River basin, August and November 2016, Texas, USA.

Decoy	Habitat	Clouds (%)	$\hat{p}$ (96.7% CI)[a]	z[b]	P	Correction factor[c] (96.7% CI)
Skimmer	Spoil island	0–100	0.54 (0.44–0.63)*	–7.425	≤0.001	1.85 (1.58–2.27)
Tern	Spoil island	0–100	1.02 (0.90–1.15)	0.356	0.722	0.98 (0.87–1.11)
White heron	Spoil island	0–100	1.09 (0.89–1.31)	1.095	0.274	0.92 (0.76–1.12)
	Cypress-tupelo	0–50	0.29 (0.18–0.46)*	–9.652	≤0.001	3.45 (2.17–5.56)
	Cypress-tupelo	51–100	0.09 (0.03–0.34)*	–2.688	0.007	11.11 (2.94–33.3)
Dark heron	Spoil island	0–100	0.95 (0.79–1.18)	–0.746	0.456	1.05 (0.84–1.27)
	Cypress-tupelo	0–100	0.11 (0.06–0.18)*	–13.47	≤0.001	9.09 (5.56–16.6)

[a] Detection probability, the ratio of the mean count index to mean decoy abundance, was estimated using GLMMs with a Poisson error distribution. 96.7% confidence intervals (CIs) estimated by parametric bootstrap ($N = 200$).

[b] z- and P-values obtained by Wald Z-tests.

[c] Visibility bias correction factors were obtained by $1/\hat{p}$.

* Detection probability is significantly different from 1 (i.e. imperfect detection), based on 96.7% CIs.

Table 2. Selected generalized linear mixed-effects regression models assessing ability of observers ($N = 6$) to estimate decoy counts from aerial imagery taken with unmanned aircraft. Imagery collected August and November 2016, Texas, USA. Includes competing models for each decoy type.

Decoy type	Model[a]	K	ΔAIC	LL[b]	R^2_m	R^2_c	$\hat{c}$
Skimmer	$\log(\mu_{ijk}) = \beta_0 + \text{COUNT} + \text{CLOUD} + \text{COUNT}\times\text{CLOUD} + \eta_j + \varepsilon_k$	6	0.00	–163.42	0.56[c]	0.83[d]	1.13[e]
Tern	$\log(\mu_{ik}) = \beta_0 + \varepsilon_k$	2	0.00	–166.96	0.00	0.92	0.20
	$\log(\mu_{ik}) = \beta_0 + \text{COUNT} + \varepsilon_k$	3	1.87	–166.89	0.00	0.92	0.19
White heron	$\log(\mu_{ijk}) = \beta_0 + \text{COUNT} + \text{CLOUD} + \text{HABITAT} + \text{COUNT}\times\text{CLOUD} + \text{COUNT}\times\text{HABITAT} + \text{CLOUD}\times\text{HABITAT} + \text{COUNT}\times\text{CLOUD}\times\text{HABITAT} + \eta_j + \varepsilon_k$	10	0.00	–364.52	0.52	0.89	1.04

Table 2 continued.

Decoy type	Model[a]	K	ΔAIC	LL[b]	R^2_m	R^2_c	$\hat{c}$
Dark heron	$\log(\mu_{ik}) = \beta_0$ + COUNT + HABITAT + COUNT×HABITAT + ε_k	5	0.00	–318.15	0.66	0.92	0.70

[a] Models include predictors of (COUNT) an indicator variable designating the count type (known abundance or observer index), (CLOUD) cloud cover in two categories, and (HABITAT) habitat type (spoil island and cypress-tupelo swamp). η_j and ε_k represent the random effects of survey and observer, respectively. I tested a full model with both random effects against a reduced model without the observer effect by likelihood ratio test (LRT).

[b] Log-likelihood of the model.

[c] Marginal R^2, the percent of variance explained by the fixed effects in the model.

[d] Conditional R^2, the percent of variance explained by both the fixed effects and random effects (i.e. among-observer and among-colony) combined.

[e] Dispersion parameter.

Table 3. Competing models for the magnitude of behavioral responses displayed in nesting waterbird colonies, as influenced by UAS surveys. Data collected in Texas, USA, 2016–2017 (n = 209).

Behavioral response	Model[a]	K	ΔAIC	LL[b]	R^2_m	R^2_c	$\hat{c}$
Vigilance	$\log(\mu_{ij}) = \beta_0$ + AGL + SIZE + LAG 1 + $\eta_j + \eta_{ij}$[c]	13	0.00	–393.87	0.37[d]	0.51[e]	1.23[f]
	$\log(\mu_{ij}) = \beta_0$ + AGL + PLATFORM + SIZE + LAG 1 + $\eta_j + \eta_{ij}$	14	1.69	–393.72	0.36	0.49	1.24
Flushing	$\log(\mu_{ij}) = \beta_0$ + PLATFORM + SIZE + η_j	4	0.00	–396.21	0.08	0.89	0.82
	$\log(\mu_{ij}) = \beta_0$ + SIZE + η_j	3	0.63	–397.52	0.05	0.90	0.82

Table 3 continued.

Behavioral response	Model[a]	K	ΔAIC	LL[b]	R^2_m	R^2_c	$\hat{c}$
Flushing	$\log(\mu_{ij}) = \beta_0$	5	1.39	–395.90	0.08[d]	0.89[e]	0.82[f]
	+ PLATFORM						
	+ SIZE						
	+ LAG 1						
	$+ \eta_j$						
	$\log(\mu_{ij}) = \beta_0$ + SIZE	4	1.98	–397.19	0.06	0.90	0.82
	+ LAG 1						
	$+ \eta_j$[c]						

[a] Predictor variables are (AGL) survey altitudes, (PLATFORM) craft type, (SIZE) colony size, and (LAG 1) lagged response values by 1, for the ith sample period at the jth colony site. AGL levels include categories for all survey altitudes, as well as for the pre- and post-flight period. Random terms of (η_j) intercept and (η_{ij}) slope for colony site are included in model statements.

[b] Log-likelihood of the model.

[c] The selected model.

[d] Marginal R^2 describes the proportion of model variance explained by the fixed effects.

[e] Conditional R^2 describes the proportion of model variance explained by the fixed and random effects combined.

[f] Dispersion parameter for the model.

Table 4. Parameter estimates of the selected model for vigilance. Data collected in Texas, USA, 2016–2017. Vigilant reactions were tallied in 4 waterbird colonies during surveys with a UAS.

Parameter	β	e^{β}	96.7% CI[a]	z[b]	P
Intercept	0.43[c]	1.54[d]	[–0.05–0.77]	2.54	0.011
300 m[e]	–0.20	0.82	[–0.91–0.51]	–0.656	0.513
200 m	–0.07	0.93	[–0.90–0.57]	–0.265	0.791
122 m	0.36	1.44	[–0.01–0.72]	2.033	0.042
91 m	0.55	1.72	[0.19–0.93]*	3.086	0.002
61 m	0.78	2.19	[0.55–1.13]*	4.514	≤0.001
46 m	0.78	2.18	[0.52–1.16]*	4.744	≤0.001
Post-flight	0.43	1.54	[0.21–0.80]*	2.692	0.007
Colony size[f]	0.01	1.01	[0.16–0.68]*	3.799	≤0.001
Lag 1	0.05	1.05	[0.01–0.21]*	2.426	0.015

[a] Parametric bootstrapped 96.7% confidence intervals of each parameter (N = 200). Estimates are significant if they do not include 0.

[b] z and P-values from Wald-Z tests.

[c] Parameter estimates for fixed effects, in log-odds.

[d] Exponentiated estimates, to reflect rate-ratios (e.g., a rate-ratio of 0.82 indicates an 18% decrease).

[e] Survey altitudes and the post-flight treatment level. Estimates reflect their influence on vigilance when compared to a baseline (pre-flight). 300 and 200 m surveys were flown with the Minion fixed-wing, 122–46 m surveys with the Phantom quadcopter.

[f] Colony size and lag 1. Parameter estimates reflect their influence on the number of vigilant birds.

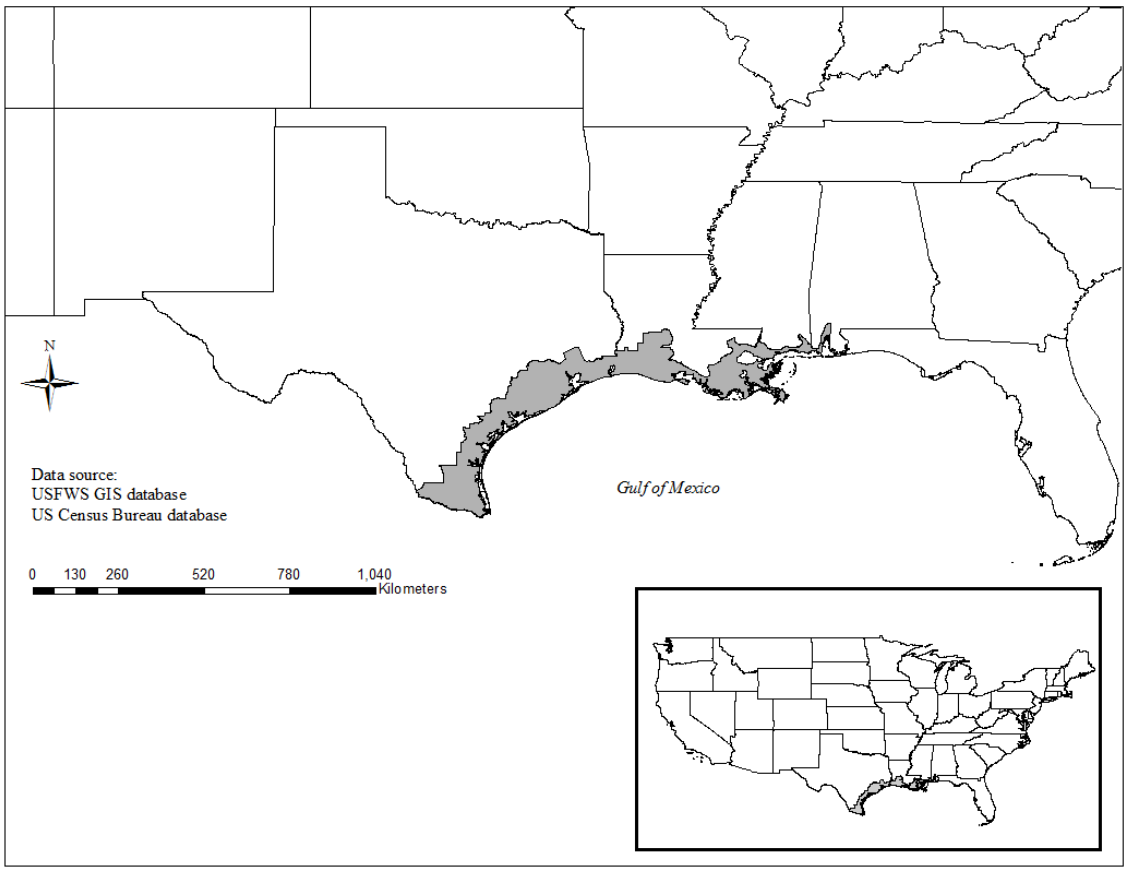

Figure 1. Map of the Gulf Coast Joint Venture (GCJV) region. The GCJV spans coastlines in 4 states: (from left to right) Texas, Lousiana, Mississippi, and Alabama.

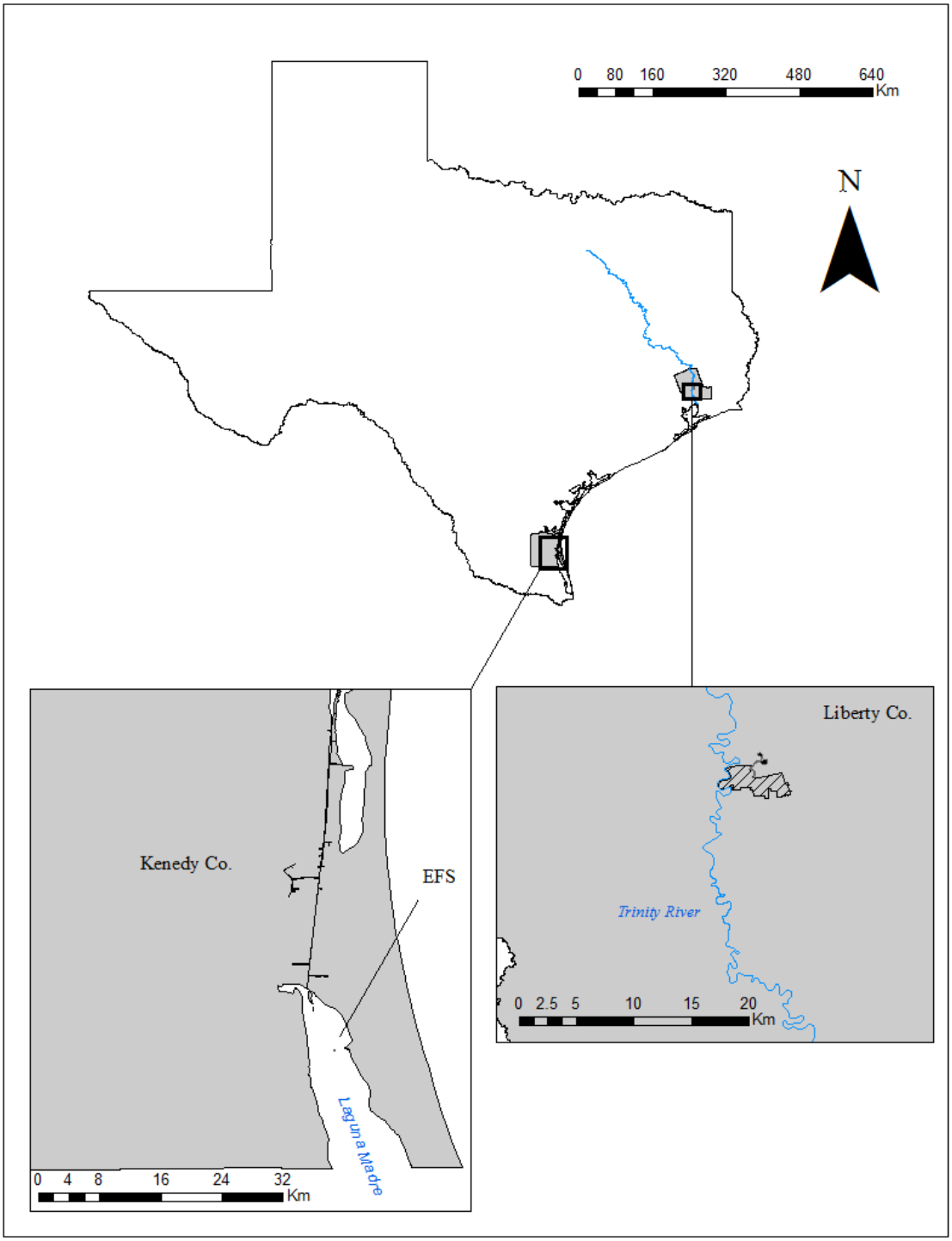

Figure 2. Study sites where aerial imagery of decoy "colonies" was obtained. Data collected August and November 2016, Texas, USA. Trinity River NWR property is indicated by hash marks within Liberty County, with the Josie Lake site located just north

of the refuge in black. East Flat Spoil island (EFS) is denoted in the inset map for Kenedy County.

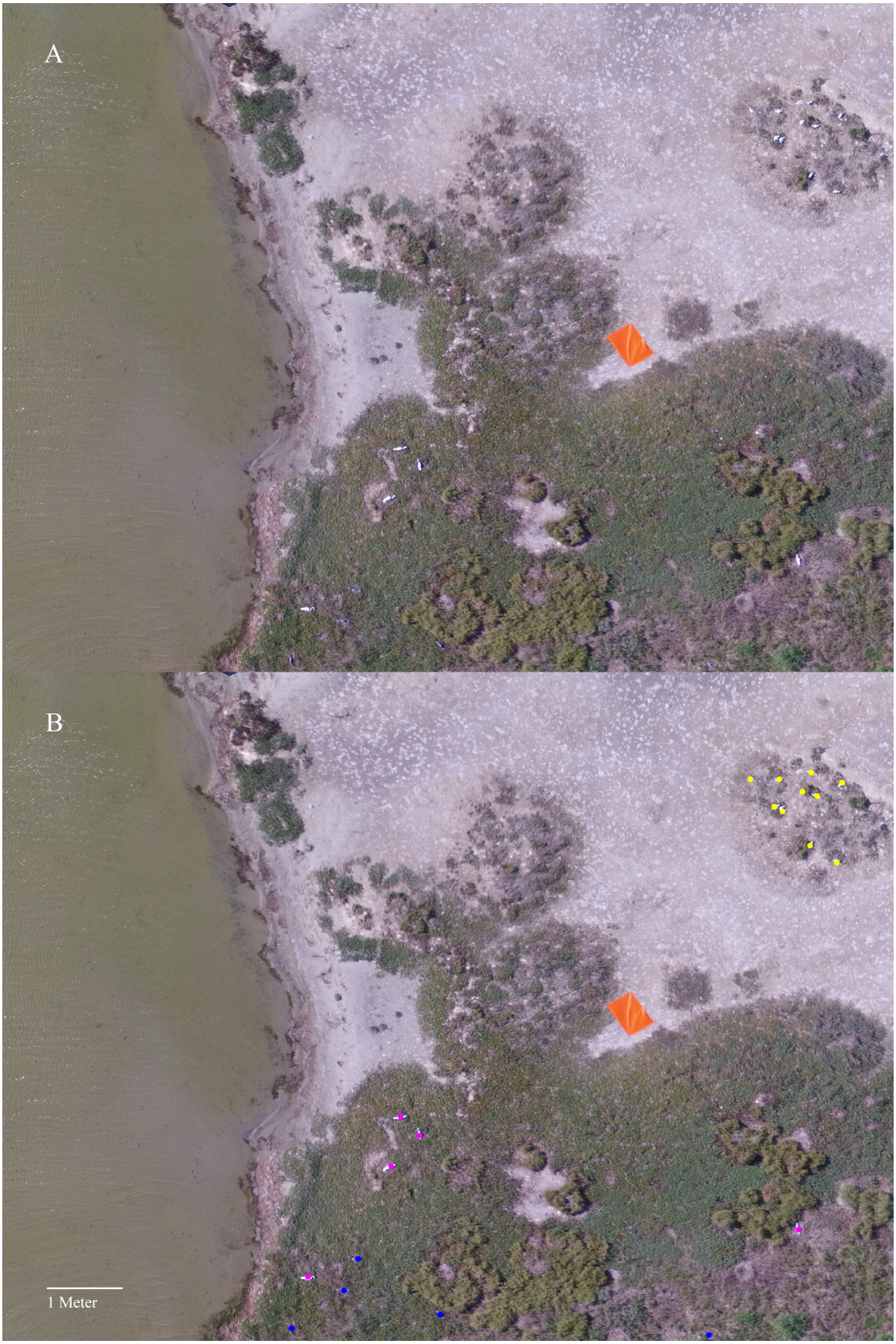
A
B
1 Meter

Figure 3. Subsection of imagery taken with a Phantom UAS (DJI, Shenzhen, Guangdong, China; 12.4 MP camera). Imagery collected in 2016 at East Flat Spoil island, Texas, USA. Imagery is shown in (A) raw form, and (B) after an observer counted decoys. Dots indicate decoy detections, with each color representing a different decoy type (pink and blue dots represent white and dark herons, respectively, and yellow dots represent terns). Aerial imagery resolution was approx. 1.7 cm/pixel.

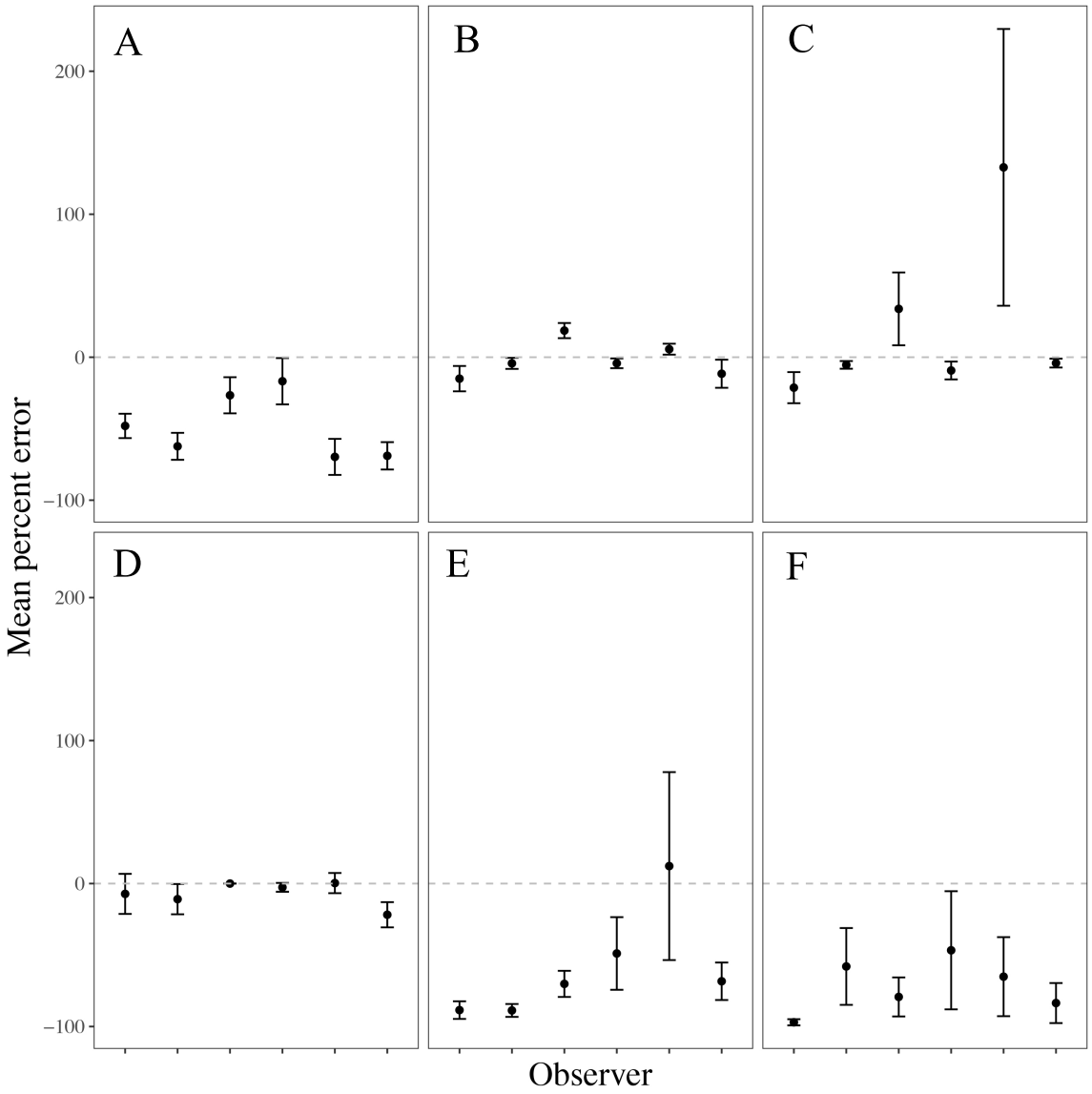

Figure 4. Mean percent error (± 1 SE) for six independent-observer counts of decoy abundance. Observers (x-axis) estimated four unique decoy types in two distinct habitats. Decoy types at East Flat Spoil island included black skimmer (A), tern (B), white-plumaged heron (C), and dark-plumaged heron (D). Decoy types at Trinity River included white-plumaged heron (E), and dark-plumaged heron (F). The dashed line demarcates a mean error of 0.

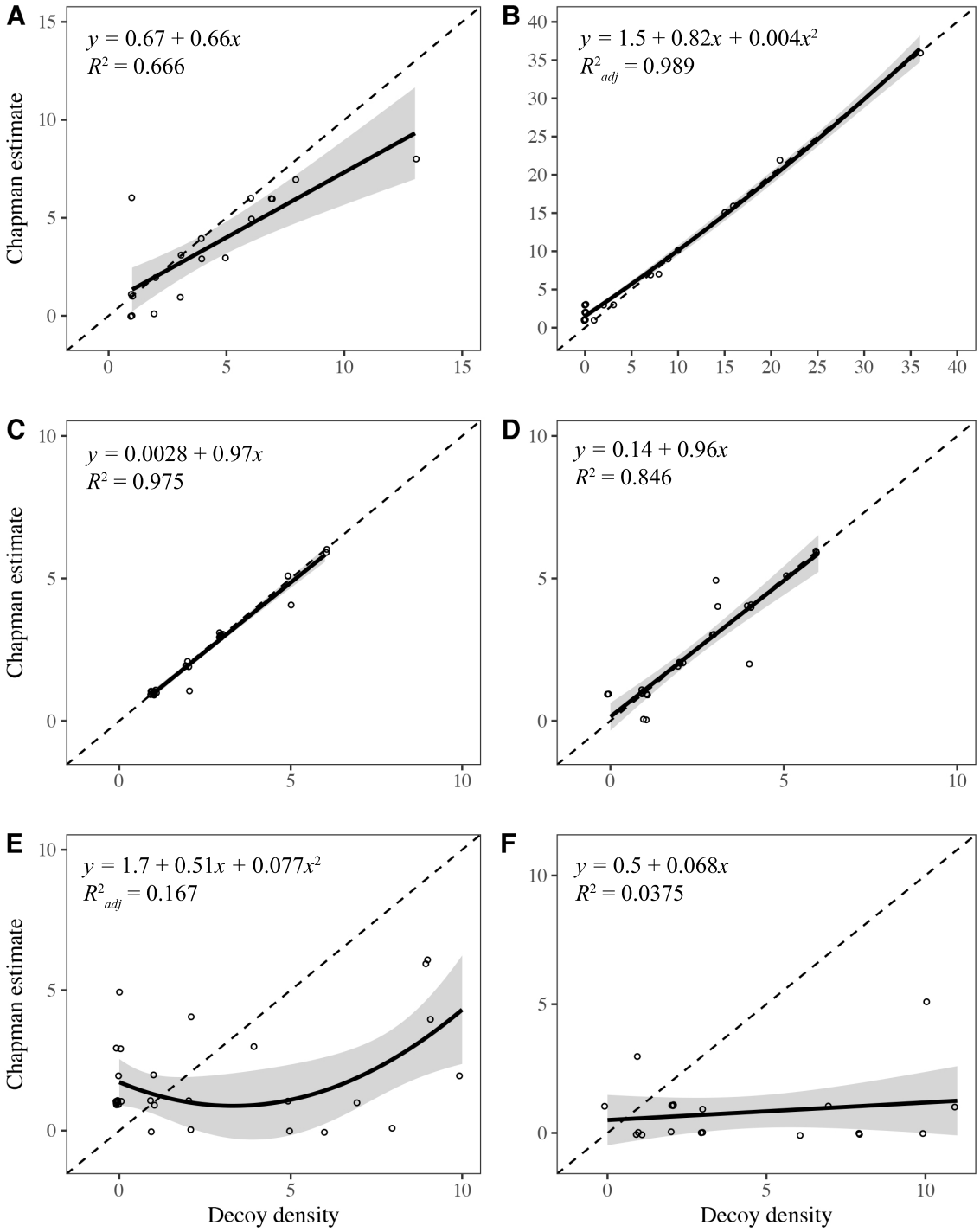

Figure 5. Linear regressions illustrating the ability of Chapman (1951) abundance estimations to estimate decoy density from image quadrats. Data collected August and November 2016, Texas, USA. Decoy types at East Flat Spoil island include black

skimmer (A), tern (B), white-plumaged heron (C), and dark-plumaged heron (D). Decoy types at Trinity River include white-plumaged heron (E), and dark-plumaged heron (F). The dashed line represents a slope equal to 1. Adjusted R^2 was reported for models that included the quadratic term (i.e. two model predictors).

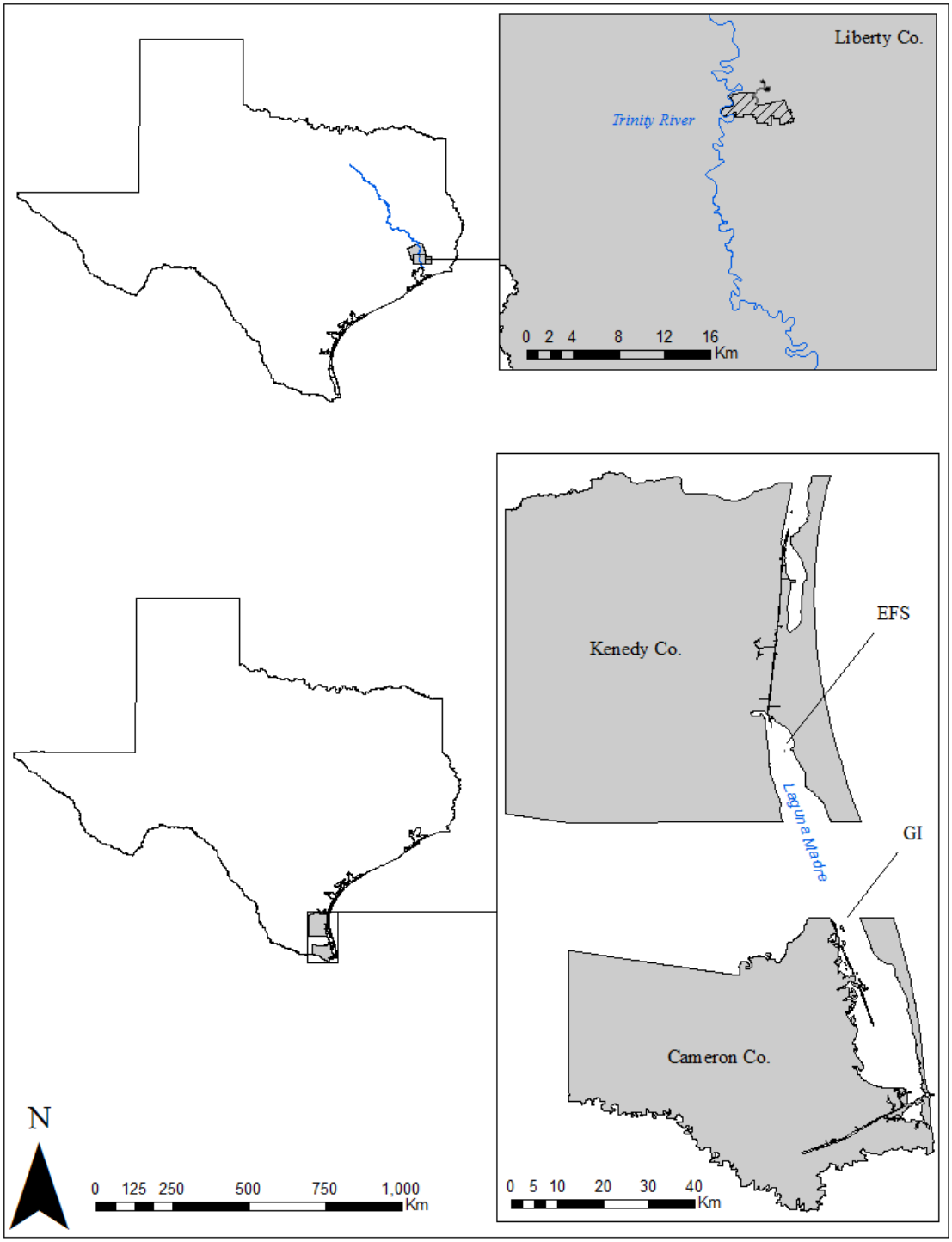

Figure 6. Sites in which UAS surveys were conducted over active waterbird colonies. Surveys took place May–June 2016 and May 2017, Texas, USA. The colonies at East

Flat Spoil (EFS) and Green Island (GI) are labeled in the inset map for Kenedy and Cameron Counties, respectively. Trinity River NWR property is indicated by hash marks within Liberty County, with the Josie Lake site just north of the refuge in black.

Figure 7. The Minion 2.0 fixed-wing UAS. Platform developed by AggieAir (Utah State University, Logan, UT, USA). Pictured is the launching system (pneumatic rail) as well as the platform itself. Surveys with the Minion were flown at 300 and 200 m.

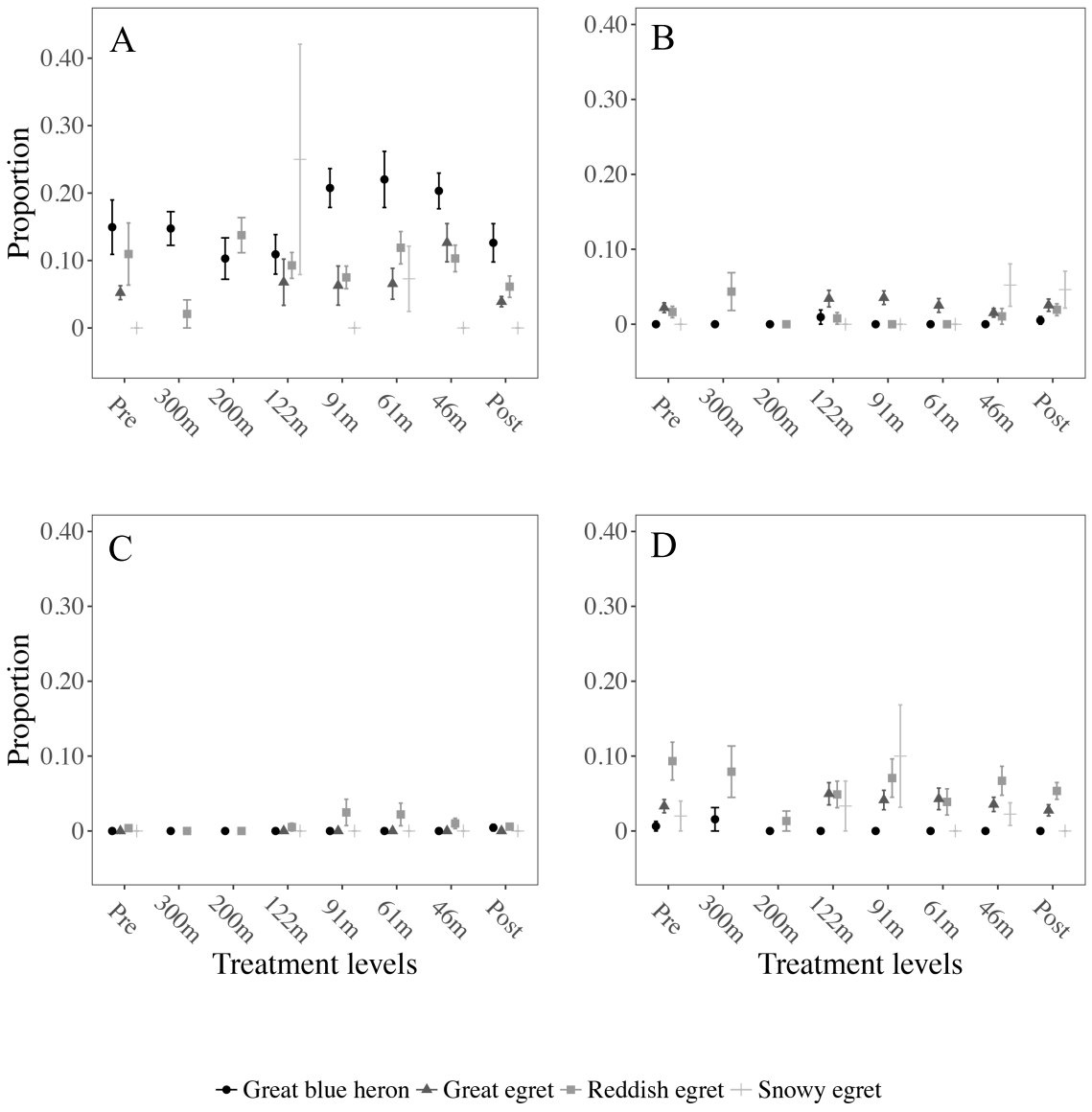

Figure 8. Proportion of four heron species ($\bar{x}$ ± SE) in camera frame that exhibited a behavioral response in relation to treatment levels. Data collected at Trinity River NWR, GI, and EFS, Texas, USA, 2016–2017. Behaviors include vigilance (A), wing flapping (B), walking off nest (C), and flushing (D). Treatment levels include pre- and post-flight periods, as well as survey altitudes from the Minion (300 and 200 m) and Phantom quadcopter (122–46 m). Note that some species were not subjected to Minion surveys.

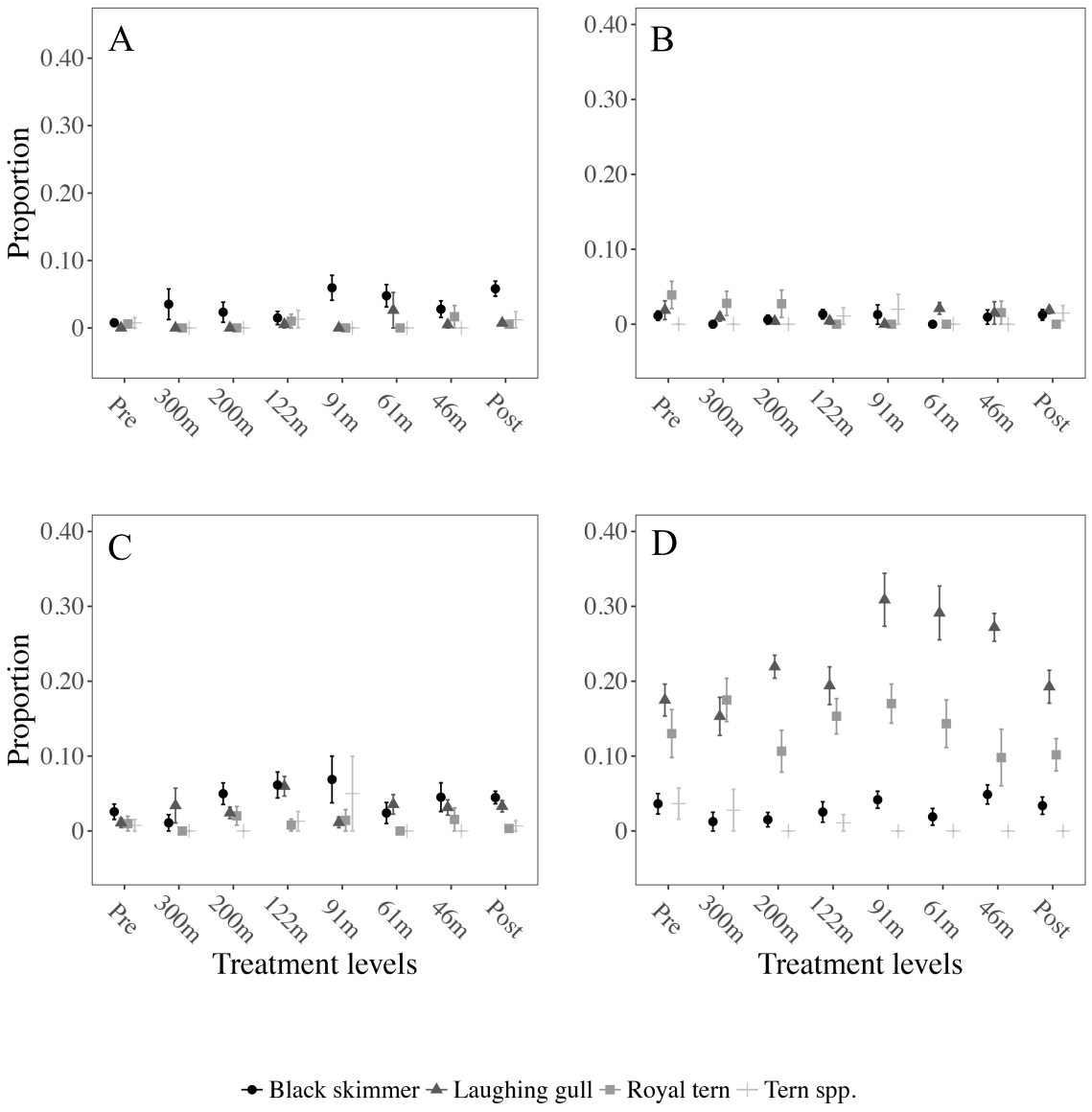

Figure 9. Proportion of ground-nesting waterbirds ($\bar{x} \pm$ SE) in camera frame that exhibited a behavioral response in relation to treatment levels. Data collected at EFS, Texas, USA, 2016–2017. Behaviors include vigilance (A), wing flapping (B), walking off nest (C), and flushing (D). Treatment levels include pre- and post-flight periods, as well as survey altitudes that were flown with the Minion (300 and 200 m) and Phantom

(122–46 m). Unidentified tern spp. were assumed to be a mix of gull-billed, sandwich, and Forster's terns.

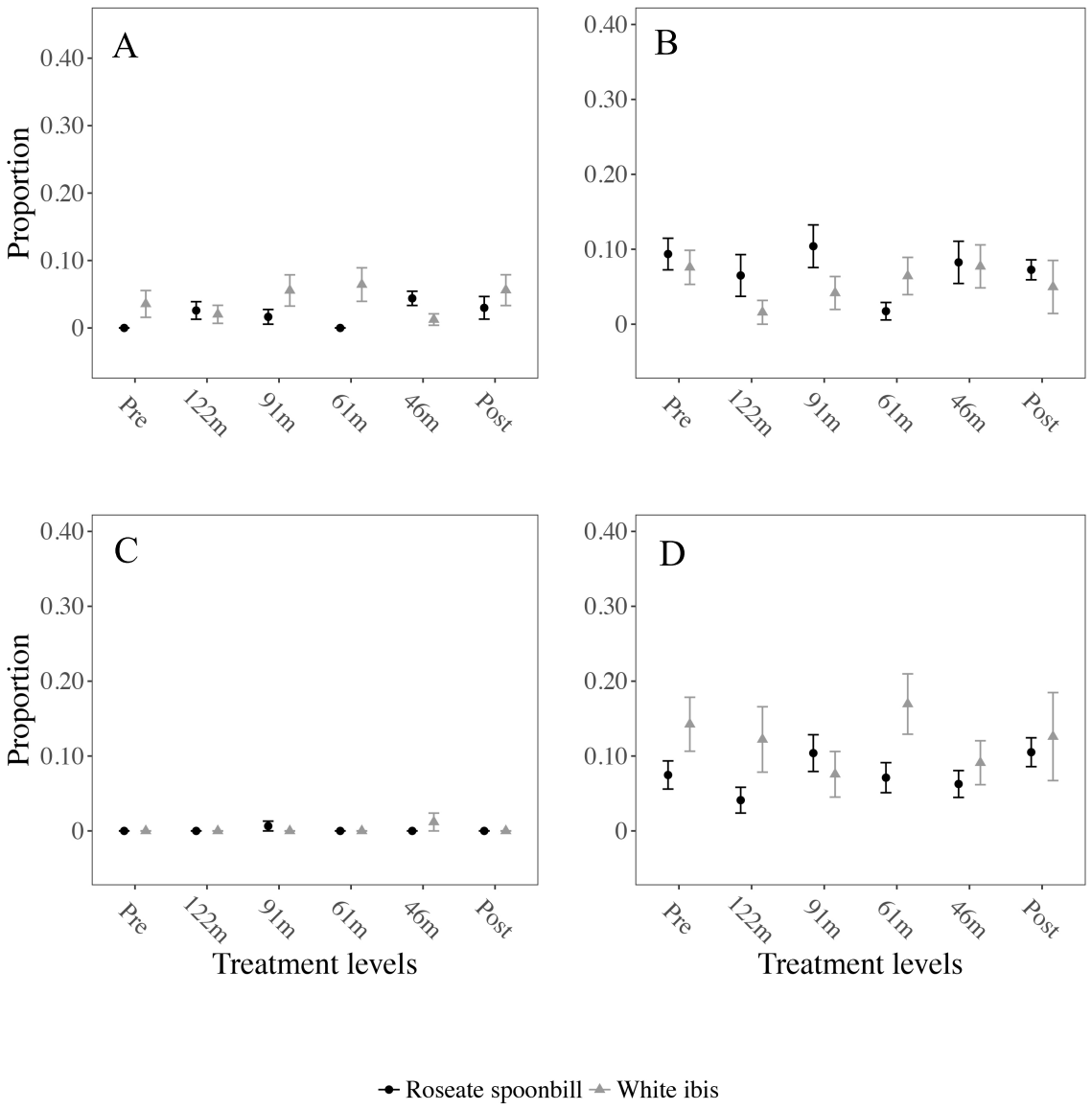

Figure 10. Proportion of roseate spoonbill and white ibis ($\bar{x} \pm$ SE) in camera frame that exhibited a behavioral response in relation to treatment levels. Data collected at GI, Texas, USA, 2016–2017. Behaviors include vigilance (A), wing flapping (B), walking off nest (C), and flushing (D). Treatment levels include the pre- and post-flight periods, as well as survey altitudes that were flown with the Phantom (122–46 m).

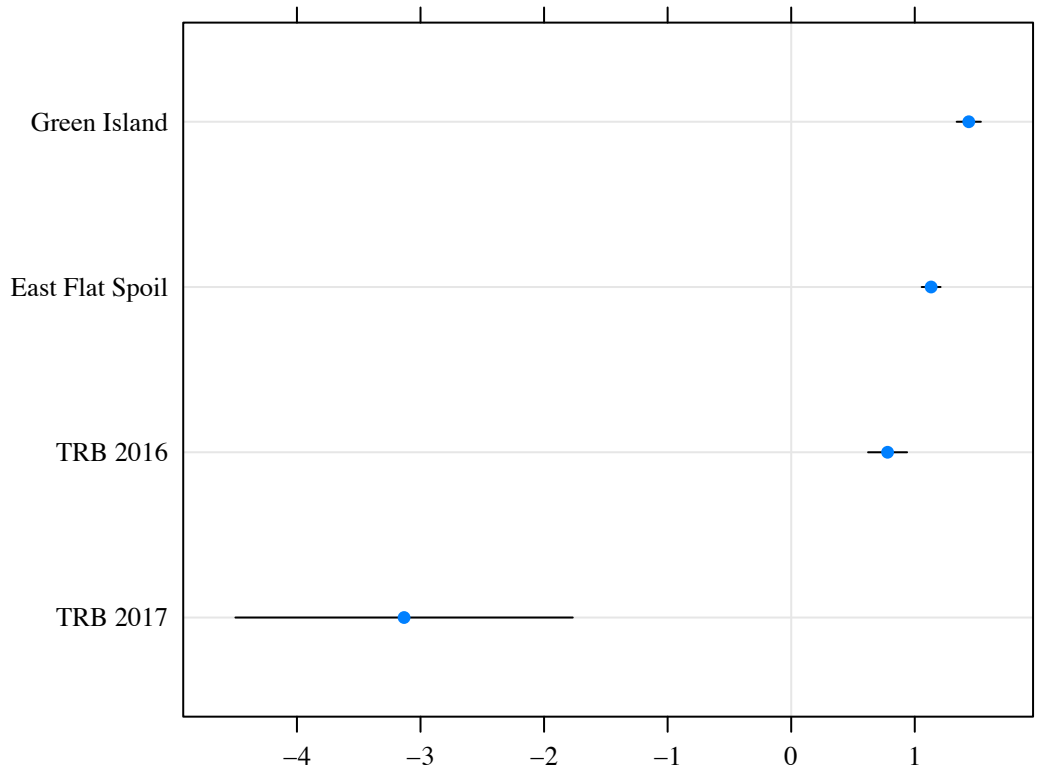

Figure 11. Conditional modes for the random effect of colony, represented for the selected flush behavior model. Colonies include Green Island, East Flat Spoil, Trinity River basin in 2016, and Trinity River basin in 2017. Note the extreme variability in the number of birds that flushed within the Trinity River basin (TRB) colony in 2017, as well as the variation among TRB 2017 and the other 3 colonies (R^2_m = 0.06, R^2_c = 0.90, $SD_{intercept}$ = 1.97).

APPENDIX SECTION

LIST OF TABLES

Table **Page**

A1. Positive and negative attributes of the Phantom quadcopter UAS for surveying mixed-species waterbird colonies 72

A2. Positive and negative attributes of the Minion fixed-wing UAS for surveying mixed-species waterbird colonies 73

B1. Recommended minimum standards for two counting techniques: automated bird counts and manual counts 74

LIST OF FIGURES

Figure **Page**

B2. Usage flowchart for two UAS platforms 76

APPENDIX A

Below is a summary of positive and negative attributes for two UAS platforms when surveying colonial waterbirds. Overall, the lightweight consumer-grade UAS (Phantom 3 and 4, DJI, Shenzhen, Guangdong, China) outperformed the custom-built fixed-wing (AggieAir, Utah State University, Logan, UT, USA) in applications for surveying.

Table A1. Positive and negative attributes of the Phantom quadcopter UAS for surveying mixed-species waterbird colonies.

Positive (+)	Negative (–)
Easily transportable.	Battery life limited to approx. 15–20 min.
Take-off and landing can be done on most terrain.	Not fully autonomous out-of-the-box, although software is available.
Inexpensive.	Limited payload (e.g., 1 sensor attachment allowed at a time).
A single technician can pilot the Phantom.	Increased vigilance of waterbirds when flown between 46–91 m (see Chapter II).
High resolution from stock camera (approx. 1.3 cm/pixel for flights at 46 m).	
Many modifications and open source software available.	
Customer service and hardware repairs available from manufacturer.	

Table A2. Positive and negative attributes of the Minion fixed-wing UAS for surveying mixed-species waterbird colonies.

Positive (+)	Negative (–)
Can carry multiple cameras and sensors.*	Bulky, difficult to transport in the field.*
Completely autonomous. Flight path is input before take-off.	Crew of 3 technicians needed for flight and ground station.
Battery life is comparatively better, at approx. 60 min. per survey.	Large landing area needed (approx. 0.5 km), with specific terrain (e.g., short grass or sand).*
High altitude flights, potentially less disturbing to birds (see Chapter II).*	Various software and hardware problems. Possibly not the case with all fixed-wing platforms.
	Stringent airspace restrictions (FAA).*
	Repairs and customer service not reliable since it was custom-built.
	Expensive.
	Every landing was a "soft" crash landing, so repairs were frequent.
	Imagery resolution not suitable for surveying mid-size birds (approx. 4 cm/pixel at 200 m altitude). This was due to the 12 MP camera used, resolution would likely suitable with a 20 MP camera or similar.

* Attributes that are common among all fixed-wing UAS models, not just the Minion.

APPENDIX B

Below is a summary of recommended usage steps for surveying with both platform types, and for using automated bird counting algorithms or manual observer counts.

Table B1. Recommended minimum standards for two counting techniques: automated bird counts and manual counts.

Attribute	Manual count[a]	Automated count[b]
Climatic conditions	Sunny, clear skies, little to no wind.	Little to no wind.
Nesting substrate	Supplement surveys with another method if nest visibility is occluded (e.g. canopy).	Supplement surveys with another method if nest visibility is occluded (e.g. canopy).
Altitude (AGL)[c]	46–91 m	46–91 m
Flight path	Back-and-forth transects, preferably automated.	Back-and-forth transects, preferably automated.
Resolution[d]	≤1.7 cm/pixel for already identified species. ≤0.5 cm/pixel to identify species via imagery.[e]	Focal species contrasts heavily with background, medium-large sized birds, high visibility habitat: ≤4.5 cm/pixel. Focal species does not contrast with background, small birds, any habitat: ≤1.7 cm/pixel.
Software	Photoshop	eCognition ArcGIS Photoshop
Analysis technique	Photoshop: color dot marking, count function, and image analysis function.	eCognition: Object Based Image Analysis. Used to more accurately detect birds that are densely clustered.

Table B1 continued.

Attribute	Manual count[a]	Automated count[b]
Analysis technique		ArcGIS: unsupervised and supervised spectral analysis. Size and adjacency filtering is recommended. Photoshop: spectral thresholding. Should only be used when birds contrast markedly with substrate and are not in dense clusters.

[a] Bird counts done manually by an observer.

[b] Bird counts automated via software.

[c] Above Ground Level.

[d] Recommended resolution, software, and analysis techniques gleaned from Chabot and Francis (2016) and Dulava et al. (2015).

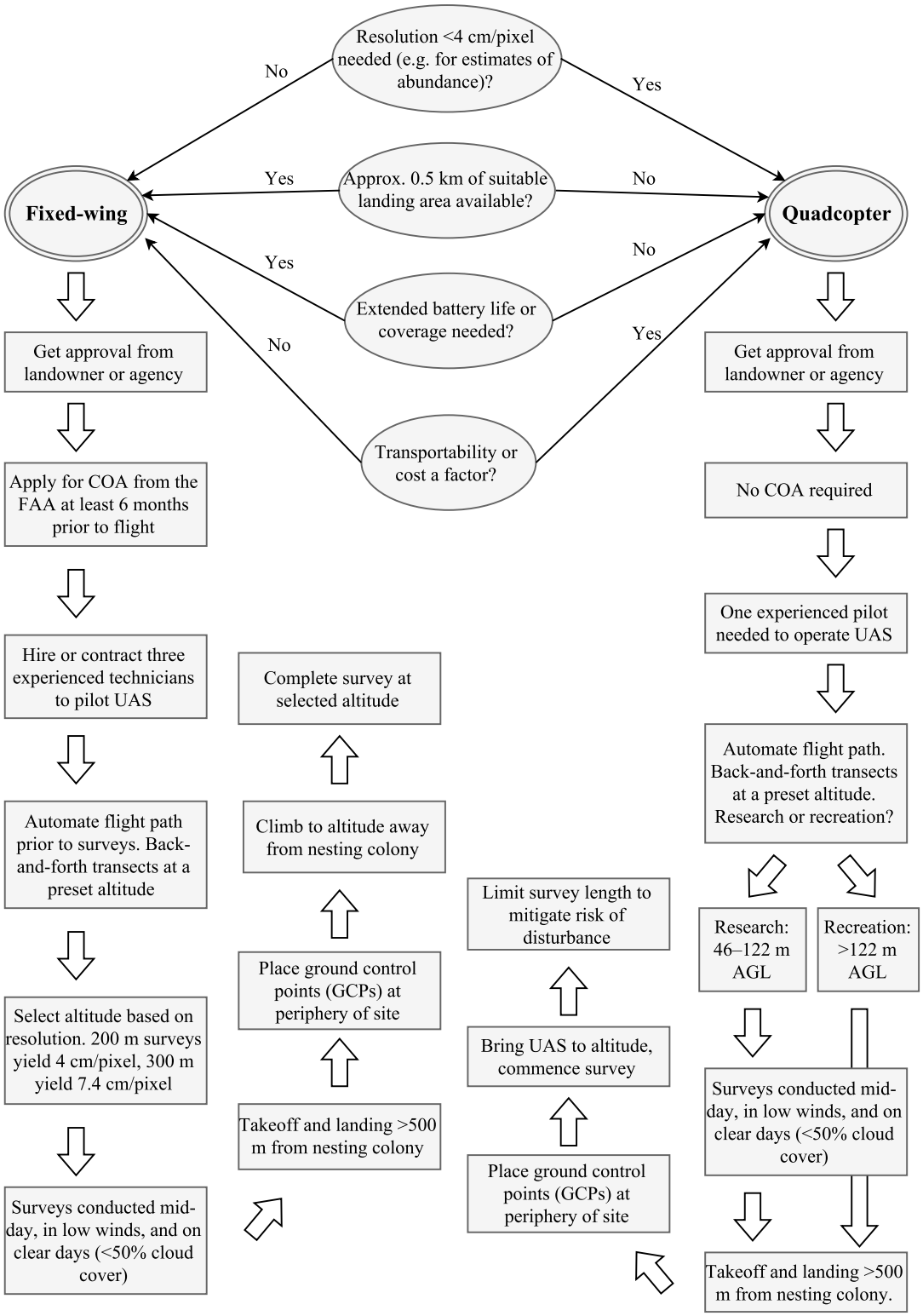
Resolution <4 cm/pixel needed (e.g. for estimates of abundance)?
No
Yes
Approx. 0.5 km of suitable landing area available?
Yes
No
Fixed-wing
Quadcopter
Extended battery life or coverage needed?
Yes
No
No
Yes
Transportability or cost a factor?
Get approval from landowner or agency
Apply for COA from the FAA at least 6 months prior to flight
Hire or contract three experienced technicians to pilot UAS
Automate flight path prior to surveys. Back-and-forth transects at a preset altitude
Select altitude based on resolution. 200 m surveys yield 4 cm/pixel, 300 m yield 7.4 cm/pixel
Surveys conducted mid-day, in low winds, and on clear days (<50% cloud cover)
Takeoff and landing >500 m from nesting colony
Place ground control points (GCPs) at periphery of site
Climb to altitude away from nesting colony
Complete survey at selected altitude
Get approval from landowner or agency
No COA required
One experienced pilot needed to operate UAS
Automate flight path. Back-and-forth transects at a preset altitude. Research or recreation?
Research: 46–122 m AGL
Recreation: >122 m AGL
Surveys conducted mid-day, in low winds, and on clear days (<50% cloud cover)
Takeoff and landing >500 m from nesting colony.
Place ground control points (GCPs) at periphery of site
Bring UAS to altitude, commence survey
Limit survey length to mitigate risk of disturbance

Figure B2. Usage flowchart for two UAS platforms. Guidelines for the fixed-wing UAS are specific to the Minion 2.0 craft (AggieAir, Utah State University, Logan, UT, USA). Survey altitude for the Phantom quadcopter (DJI, Shenzhen, Guangdong, China) is dependent upon the goal of the project and the imagery resolution needed. Surveys at 46, 61, 91, and 122 m yielded approximate imagery resolutions of 1.3, 1.7, 2.5, and 3.3 cm/pixel, respectively.

LITERATURE CITED

Akaike, H. 1974. A new look at the statistical model identification. Transactions on Automatic Control 19(6):716-723.

Altmann, J. 1974. Observational study of behavior: sampling methods. Behaviour 49(3):227–266.

Anderson, D. W., and J. O. Keith. 1980. The human influence on seabird nesting success: conservation implications. Biological Conservation 18:65–80.

Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin 29(4):1294–1297.

Bakó, G., M. Tolnai, and A. Takács. 2014. Introduction and testing of a monitoring and colony-mapping methods for waterbid populations that uses high-speed and ultra-detailed aerial remote sensing. Sensors 14:12828–12846.

Bayliss, P. and K. M. Yeomans. 1989. Correcting bias in aerial survey population estimates of feral livestock in northern Australia using the double-count technique. Journal of Applied Ecology 26(3):925–933.

Bayliss, P. and K. M. Yeomans. 1990. Use of low-level aerial photography to correct bias in aerial survey estimates of magpie goose and whistling duck density in the northern territory. Wildlife Research 17:1–10.

Becker, P. H. 1995. Effects of coloniality on gull predation on common tern chicks. Colonial Waterbirds 18(1):11–22.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 57:289–300.

Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J. S. S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology and Evolution 24(3):127–135.

Bolstad, P., editor. 2016. Aerial and satellite images. Pages 247–296 *in* GIS fundamentals: a first text on geographic information systems. Fifth edition. Eider Press, White Bear Lake, Minnesota, USA.

Brunton, D. 1999. "Optimal" colony size for least terns: an inter-colony study of opposing selective pressures by predators. The Condor 101:607–615.

Burger, J. 1981. Effects of human disturbance on colonial species, particularly gulls. Colonial Waterbirds 4:28–36.

Burger, J., and M. Gochfeld. 1983. Behavioral responses to human intruders of herring gulls (*Larus agentatus*) and great black-backed gulls (*L. marinus*) with varying exposure to human disturbance. Behavioral Processes 8:326–344.

Burger, J., M. Gochfeld, and L. J. Niles. 1995. Ecotourism and birds in coastal New Jersey: contrasting responses of birds, tourists, and managers. Environmental Conservation 22(1):56–65.

Burnham, K. P. and D. R. Anderson. 2002. Model selection and multimodal inference. Second edition. Springer-Verlag, New York, New York, USA.

Butler, M. J., W. B. Ballard, M. C. Wallace, S. J. DeMaso, and B. K. McGee. 2007. Aerial surveys for estimating wild turkey abundance in the Texas rolling plains. Journal of Wildlife Management 71(5):1639–1645.

Carney, K. M., and W. J. Sydeman. 1999. A review of human disturbance effects on nesting colonial waterbirds. Waterbirds 22(1):68–79.

Caughley, G. 1974. Bias in aerial survey. Journal of Wildlife Management 38(4):921–933.

Chabot, D., and D. M. Bird. 2012. Evaluation of an off-the-shelf unmanned aircraft system for surveying flocks of geese. Waterbirds 35(1):170–174.

Chabot, D., V. Carignan, S. R. Craik, and D. M. Bird. 2014. Evaluation and application of a small unmanned aircraft in waterbird research and conservation. Proceedings of The Wildlife Society annual conference. Pittsburgh, Pennsylvania, USA.

Chabot, D., S. R. Craik, and D. M. Bird. 2015. Population census of a large common tern colony with a small unmanned aircraft. PLoS One 10(4):e0122588.

Chabot, D., and C. M. Francis. 2016. Computer-automated bird detection and counts in high-resolution aerial images: a review. Journal of Field Ornithology 0(0):1–17.

Chapman, D. G. 1951. Some properties of the hypergeometric distribution with applications to zoological censuses. University of California Publications in Statistics 1:131–160.

Clemmons, J. R. and R. Buchholz, editors. 1997. Linking conservation and behavior. Pages 3–22 *in* Behavioral Approaches to Conservation in the Wild. Cambridge University Press, Cambridge, UK.

Conroy, M. J., J. T. Peterson, O. L. Bass, C. J. Fonnesbeck, J. E. Howell, C. T. Moore, and J. P. Runge. 2008. Sources of variation in detection of wading birds from aerial surveys in the Florida everglades. The Auk 125(3):731–743.

Cook, R. D. and J. O. Jacobson. 1979. A design for estimating visibility bias in aerial surveys. Biometrics 35(4):735–742.

Coulter, M. C. and P. C. Frederick. 1997. Movements and population dynamics of colonial waterbirds as guides for the temporal and spatial scales of conservation. Colonial Waterbirds 20(2):295–297.

Culik, B., D. Adelung, and J. Woakes. 1990. The effect of disturbance on the heart rate and behavior of Adélie penguins (*Pygoscelis adeliae*) during the breeding season. Antarctic Ecosystems 177–182.

Dahm, C. N., R. J. Edwards, F. P. Gelwick. 2011. Gulf coast rivers of the Southwestern United States. Pages 181-230 *in* Benke, A. C. and C. E. Cushing, editors. Rivers of North America. Academic Press, Waltham, Massachusetts, USA.

DeRose-Wilson, A., J. D. Fraser, S. M. Karpanty, and M. D. Hillman. 2015. Effects of overflights on incubating Wilson's plover behavior and heart rate. Journal of Wildlife Management 79(8):1246–1254.

Dixon, T. J. 1977. The distance at which sitting birds can be seen at sea. Ibis 119(3):372–375.

Dulava, S., W. T. Bean, and O. M. W. Richmond. 2015. Applications of unmanned aerial systems (UAS) for waterbird surveys. Environmental Practice 17:201–210.

Erwin, R. M. 1989. Responses to intruders by birds nesting in colonies: experimental results and management guidelines. Colonial Waterbirds 12(1):104–108.

Frederick, P. C., and M. W. Collopy. 1989. Nesting success of five Ciconiiform species in relation to water conditions in the Florida Everglades. The Auk 106(4):625–634.

Frederick, P. C., B. Hylton, J. A. Heath, and M. Ruane. 2003. Accuracy and variation in estimates of large numbers of birds by individual observers using an aerial survey simulator. Journal of Field Ornithology 74(3):281–287.

Fulbright, T. E., D. D. Diamond, J. Rappole, and J. Norwine. 1990. The coastal sand plain of south Texas. Rangelands 12:337–340.

Giese, M. 1996. Effects of human activity on adélie penguin (*Pygoscelis adeliae*) breeding success. Biological Conservation 75(2):157–164.

Giese, M. and M. Riddle. 1999. Disturbance of emperor penguin (*Aptenodytes forsteri*) chicks by helicopters. Polar Biology 22(6):366–371.

Gillett, W. H., J. L. Hayward, and J. F. Stout. 1975. Effects of human activity on egg and chick mortality in a glaucous-winged gull colony. The Condor 77(4):492–495.

Goering, D. K., and R. Cherry. 1971. Nestling mortality in a Texas heronry. The Wilson Bulletin 83(3):303–305.

Graham, A., and R. Bell. 1989. Investigating observer bias in aerial survey by simultaneous double-counts. Journal of Wildlife Management 53(4):1009–1016.

Green, M.C., and P. L. Leberg. 2005. Flock formation and the role of plumage colouration in Ardeidae. Canadian Journal of Zoology 83:683–693.

Green, M. C., M. C. Luent, T. C. Michot, C. W. Jeske, and P. L. Leberg. 2008. Comparison and assessment of aerial and ground estimates of waterbird colonies. Journal of Wildlife Management 72:697–706.

Gutzwiller, K. J., R. T. Wiedenmann, K. L. Clements, and S. H. Anderson. 1994. Effects of human intrusion on song occurrence and singing consistency in subalpine. The Auk 111(1):28–37.

Harrington, A. 2015. Who controls the drones? (regulation unmanned aircraft). Engineering and Technology 10(2):80–83.

Hillman, M. D., S. M. Karpanty, J. D. Fraser, and A. DeRose-Wilson. 2015. Effects of aircraft and recreation on colonial waterbird nesting behavior. Journal of Wildlife Management 79(7):1192–1198.

Hunt, G. L. 1972. Influence of food distribution and human disturbance on the reproductive success of herring gulls. Ecology 53(6):1051–1061.

Hunter, W. C., W. Golder, S. L. Melvin, and J. A. Wheeler. 2006. Southeastern United States regional waterbird conservation plan. U.S. Fish and Wildlife Service, Atlanta, Georgia, USA.

Israel, M. 2011. A UAV-based roe deer fawn detection system. International Archives of the Photogrammetry, Remote Sensing, and Spatial Information Sciences 38(1):51–55.

IUCN red list of threatened species. 2015. <www.iucnredlist.org>. Accessed 10 October 2015.

Jeffress, M. R., C. P. Paukert, B. K. Sandercock, and P. S. Gipson. 2011. Factors affecting detectability of river otters during sign surveys. Journal of Wildlife Management 75(1):144–150.

Johnson, C. M. and W. B. Krohn. 2001. The importance of survey timing in monitoring breeding seabird numbers. Waterbirds 24(1):22–33.

Jones IV, G. P., L. G. Pearlstine, and H. F. Percival. 2006. An assessment of small unmanned aerial vehicles for wildlife research. Wildlife Society Bulletin 34:750–758.

Kingsford, R. T. 1999. Aerial survey of waterbirds on wetlands as a measure of river and floodplain health. Freshwater Biology 41:425–438.

Koneff, M. D., J. A. Royle, M. C. Otto, J. A. Wortham, and J. K. Bidwell. 2008. A double-observer method to estimate detection rate during aerial waterfowl surveys. Journal of Wildlife Management 72(7):1641–1649.

Kushlan, J. A. 1979. Effects of helicopter censuses on wading bird colonies. Journal of Wildlife Management 43(3):756–760.

Kushlan, J. A. 1993. Colonial waterbirds as bioindicators of environmental change. Colonial waterbirds 16(2):223–251.

Kury, C. R., and M. Gochfeld. 1975. Human interference and gull predation in cormorant colonies. Biological Conservation 8:24–34.

Laursen, K., J. Frikke, and J. Kahlert. 2008. Accuracy of 'total counts' of waterbirds from aircraft in coastal waters. Wildlife Biology 14(2):165–175.

Leberg, P. L., M. C. Green, B. A. Adams, K. M. Purcell, and M. C. Luent. 2007. Response of waterbird colonies in southern Louisiana to recent droughts and hurricanes. Animal Conservation 10:502–508.

Lukacs, M., and D. Bhadra. 2017. FAA aerospace forecast, fiscal years 2017–2037: unmanned aircraft vehicles. Federal Aviation Administration, Washington, D.C., USA.

Ma, Z., Y. Cai, B. Li, and J. Chen. 2010. Managing wetlands for waterbirds: an international perpective. Wetlands 30(1):15–27.

McCullagh, P. and J. A. Nelder, editors. 1989. Monographs on statistics and applied probability: generalized linear models. Second edition. Chapman and Hall, New York, New York, USA.

Menkens, G. E., and S. H. Anderson. 1988. Estimation of small-mammal population size. Ecology 69(6):1952–1959.

Mueller, A. J., and P. O. Glass. 1988. Disturbance tolerance in a Texas waterbird colony. Colonial Waterbirds 11(1):119–122.

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.

Nakagawa, S. and Schielzeth, H. 2013. A general and simple method for obtaining R2 from generalized mixed-effects models. Methods in Ecology and Evolution 4(2):133–142.

Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from count points. The Auk 117(2):393–408.

Nisbet, I. C. T. 2000. Disturbance, habituation, and management of waterbird colonies. Waterbirds 23(2):312–332.

Pagano, A. M., and T. W. Arnold. 2009. Detection probabilities for ground-based breeding waterfowl surveys. Journal of Wildlife Management 73(3):392–398.

Palmer, R. 1941. A behavior study of the common tern. Proceedings of the Boston Society of Natural History 42:1–119.

Paracuellos, M. and J. L. Tellería. 2004. Factors affecting the distribution of a waterbird community: the role of habitat configuration and bird abundance. Waterbirds 27(4):446–453.

Parnell, J. F., D. G. Ainley, H. Blokpoel, B. Cain, T. W. Custer, J. L. Dusi, S. Kress, J. A. Kushlan, W. E. Southern, L. E. Strenzel, and B. C. Thompson. 1988. Colonial waterbird management in North America. Colonial Waterbirds 11(2):129–169.

Parsons, K. C., and J. Burger. 1982. Human disturbance and nestling behavior in black-crowned night herons. The Condor 84:184–187.

Pearse, A. T., P. D. Gerard, S. J. Dinsmore, R. M. Kaminski, and K. J. Reinecke. 2008. Estimation and correction of visibility bias in aerial surveys of wintering ducks. Journal of Wildlife Management 72(3):808–813.

Pemberton, J. R. 1922. A large tern colony in Texas. The Condor 24(2):37–48.

Pinheiro, J. C. and D. M. Bates. 2000. Linear mixed-effects models: basic concepts and examples. Pages 3–56 *in* Chambers, J., W. Eddy, W. Hardle, S. Sheather, and L. Tierney, editors. Mixed-effects models in S and S-Plus. Springer-Verlag, New York, New York, USA.

Plumpton, D. L. 2006. Review of studies related to aircraft noise disturbance in waterfowl: a technical report in support of the Supplemental Environmental Impact Statement (SEIS) for introduction of F/A-18 E/F (Super Hornet) aircraft to the East Coast of the United States. Ecology and Environment, Inc., San Francisco, California, USA.

Pollock, K.H. and W. L. Kendall. 1987. Visibility bias in aerial surveys: a review of estimation procedures. Journal of Wildlife Management 51(2):502–510.

Raynor, E. J., A.R. Pierce, T. M. Owen, C. M. Leumas, and F. C. Rohwer. 2013. Short-term demographic responses of a coastal waterbird community after two major hurricanes. Waterbirds 36(1):88–93.

Regan, T. J., I. Chadès, and H. P. Possingham. 2011. Optimally managing for imperfect detection: a method for plant invasions. Journal of Applied Ecology 48:76–85.

Rodgers, J. A. and H. T. Smith. 1995. Set-back distances to protect nesting bird colonies from human disturbance in Florida. Conservation Biology 9(1):89–99.

Rodgers, J. A., P. S. Kubilis, and S. A. Nesbitt. 2005. Accuracy of aerial surveys of waterbird colonies. Waterbirds 28(2):230–237.

Rodgers, J. A. and H. T. Smith. 2012. Little Blue Heron (*Egretta caerulea*). The Birds of North America Online. <https://birdsna.org/Species-Account/bna/species/libher>. Accessed 20 October 2015.

Rümmler, M. C., Mustafa, O., Maercker, J., Peter, H. U., Esefeld, J. 2015. Measuring the influence of unmanned aerial vehicles on Adélie penguins. Polar Biology 39:1329–1334.

Safina, C., and J. Burger. 1983. Effects of human disturbance on reproductive success in the black skimmer. The Condor 85:164–171.

Sardá-Palomera, F., G. Bota, C. Viñolo, O. Pallarés, V. Sazatornil, L. Brotons, S. Gomáriz, and F. Sardá. 2012. Fine-scale bird monitoring from light unmanned aircraft systems. Ibis 154:177–183.

Seber, G. A. F., editor. 1982. Closed population: single mark release. Pages 59–125 *in* the estimation of animal abundance and related parameters. Griffin, London, England, U.K.

Simons, T. R., M. W. Alldredge, K. H. Pollock, and J. M. Wettroth. 2007. Experimental analysis of the auditory detection process on avian point counts. The Auk 124(3): 986–999.

Smith, E. H. 2002. Colonial waterbirds and rookery islands. Pages 182–197 *in* Tunnel, J. W. and F. W. Judd, editors. The Laguna Madre of Texas and Tamaulipas. Texas A&M University Press, College Station, Texas, USA.

Stillman, R. A., A. D. West, R. W. G. Caldow, and S. E. A. Le V. Dit Durell. 2007. Predicting the effect of disturbance on coastal birds. Ibis 149(1):73–81.

Strobel, B. N., and M. J. Butler. 2014. Monitoring whooping crane abundance using aerial surveys: influences on detectability. Wildlife Society Bulletin 38(1):188–195.

Temple, S. A. and J. A. Wiens. 1989. Bird populations and environmental changes: can birds be bio-indicators? American Birds 43(2):260–270.

Thomas, L. 1996. Monitoring long-term population change: why are there so many analysis methods? Ecology 77(1):49–58.

Tremblay, J., and L. N. Ellison. 1979. Effects of human disturbance on breeding of black-crowned night herons. The Auk 96(2):364–369.

Tunnel, J. W. 2002. Geography, climate, and hydrology. Pages 7-27 *in* J. W. Tunnel and F. W. Judd, editors. The Laguna Madre of Texas and Tamaulipas. Texas A&M University Press, College Station, Texas, USA.

Van de Voorde, S., Witteveen, M., Brown, M. 2015. Differential reactions to anthropogenic disturbance by two ground-nesting shorebirds. Ostrich 1–10.

Vas, E., A. Lescroel, O. Duriez, G. Boguszewski, and D. Gremillet. 2015. Approaching birds with drones: first experiments and ethical guidelines. Biology Letters 11: 20140754.

Venables, W. N., and B. D. Ripley. 2002. Random and mixed effects. Pages 271–300 *in* Chambers, J., W. Eddy, W. Härdle, S. Sheaffer, and L. Tierney, editors. Modern applied statistics with S. Fourth edition. Springer-Verlag, New York, New York, USA.

Verhulst, S., and K. Oosterbeek. 2001. Experimental evidence for effects of human disturbance on foraging and parental care in oystercatchers. Biological Conservation 101(3):375–380.

Walter, S. E., and D. H. Rusch. 1997. Visibility bias on counts of nesting Canada geese. Journal of Wildlife Management 61(3):768–772.

Watts, A. C., J. H. Perry, S. E. Smith, M. A. Burgess, B. E. Wilkinson, Z. Szantoi, P. G. Ifju, and H. F. Percival. 2010. Small unmanned aircraft for low-altitude aerial surveys. Journal of Wildlife Management 74(7):1614–1619.

Weissensteiner, M. H., J. W. Poelstra, and J. B. W. Wolf. 2015. Low-budget read-to-fly unmanned aerial vehicles: an effective tool for evaluating the nesting status of canopy breeding bird species. Journal of Avian Biology 46(4):425–430.

Wiens, J. A., J. F. addicott, T. J. Case, and J. Diamond. 1986. Overview: the importance of spatial and temporal scale in ecological investigations. Community Ecology: 145–153. Harper and Row, New York, New York, USA.

Williams, B., D. F. Brinker, and B. D. Watts. 2007. The status of colonial nesting wading bird populations within the Chesapeake Bay and Atlantic barrier island-lagoon system. Waterbirds 30(1):82–92.

Williams, K. A., P. C. Frederick, P. S. Kubilis, and J. C. Simon. 2008. Bias in aerial estimates of the number of nests in white ibis and great egret colonies. Journal of Field Ornithology 79(4):438–447.

Wilson, T. E., J. Wheeler, M. C. Green, and E. Palacios, editors. 2012. Reddish egret conservation action plan. Reddish egret conservation planning workshop, October 2012. Corpus Christi, Texas, USA.