Use of visible spectrum sUAS photography for land cover classification at nest sites of a declining bird species (*Falco sparverius*)

Matthew Kamm & J. Michael Reed

Department of Biology, Tufts University, 200 College Avenue, Medford, Massachusetts,

Keywords
Citizen science, ENVI, habitat selection, supervised classification, UAV, Unmanned aerial vehicles

Abstract
Photography with small unmanned aircraft systems (sUAS) offers opportunities for researchers to better understand habitat selection in wildlife, especially for species that select habitat from an aerial perspective (e.g., many bird species). The growing number of commercial sUAS being flown by recreational users represents a potentially valuable source of data for documenting and studying wildlife habitat. We used a commercially available quadcopter sUAS with a visible spectrum camera to classify habitat for American Kestrels (*Falco sparverius*; Aves), as well as to evaluate aspects of image processing and postprocessing relevant to a simple habitat analysis using citizen science photography. We investigated inter–observer repeatability of habitat classification, effectiveness of cross-image classification and Gaussian filtering, and sensitivity to classification resolution. We photographed vegetation around nests from both 25 m and 50 m above takeoff elevation, and analyzed images via maximum likelihood supervised classification. Our results indicate that commercial off-the-shelf sUAS photography can distinguish between grass, herbaceous, woody, bare ground, and human-modified cover classes with good (kappa > 0.6) or strong (kappa > 0.8) accuracy using a 0.25 m$^2$ minimum patch size for aggregation. There was inter-subject variability in designating training samples, but high repeatability of supervised classification accuracy. Gaussian filtering reduced classification accuracy, while coarser classification resolution out-performed finer resolution due to “speckling noise.” Image self-classification significantly outperformed cross-image classification. Mean classification accuracy metrics (kappa values) across different photo heights differed little, but, importantly, the rank order of images differed noticeably.

Introduction
Small unmanned aerial systems (sUAS, “drones”), are gaining popularity as tools in ecological research. These platforms with onboard sensors allow research that previously required manned flights, such as surveys of seabirds (Hodgson et al. 2016) and large mammals (Chrétien et al. 2016), sampling air quality (Rossi and Brunelli 2016), and surveying tree canopies (van Andel et al. 2015). In particular, sUAS seem like excellent tools for studying habitat selection in volant birds, since they allow researchers to get a “bird’s eye view,” analogous to what birds see when flying over a site. The potential for citizen science engagement in capturing sUAS photographs is large; according to the U.S. Federal Aviation Administration, in 2016 > 670 000 drones were registered in the USA and 1.9 million were sold to hobbyists (Vanian 2016). Drone hobbyists might be engaged to aid research and conservation efforts by taking aerial photos of wildlife habitat. This could make aerial imagery with a broad spatial and temporal scope available to researchers through citizen science projects. However, use of remotely sensed data from consumer-grade cameras is fraught with potential image analysis pitfalls.

Our goal is to demonstrate use of sUAS photography for assessing wildlife habitat at the nest-site level, with a focus on making best use of off-the-shelf sUAS photographs to answer conservation-relevant questions. We do this using the American Kestrel (*Falco sparverius*), a
Visible-Spectrum sUAS Land Cover Classification

M. Kamm & J. M. Reed

small, widespread raptor that forages in grasslands and other open habitats (Smallwood and Bird 2002). American Kestrels are declining across much of North America; while there are likely multiple causes for this decline, loss of nesting habitat is widely believed to be a significant factor (Smallwood and Collopy 2009; Smallwood et al. 2009). This has prompted conservation in the form of nest box programs (Fig. 1) with a long-term interest in identifying the most appropriate sites for nest box placement (Hamerstrom et al. 1973; Toland and Elder 1987; Katzner et al. 2005; Liebana et al. 2013). Recent work suggests that artificial nest boxes may not always be beneficial to kestrel populations, especially if nest site availability is not limiting (McClure et al. 2017). Wildlife managers should place artificial nesting sites with the best possible information about the needs of the species. However, UAS photography offers an opportunity to evaluate cover types at potential nest sites and identify patterns among sites kestrels select.

Our objectives are to (1) demonstrate feasibility of visible spectrum UAS photography to classify habitat around kestrel nest box locations using supervised classification (Chabot et al. 2014), and (2) evaluate potential pitfalls and best practices in the classification process for the benefit of prospective users of the technology. These analyses illustrate some basic principles of remote sensing for the benefit of readers who have not worked with remotely sensed data. We evaluated the effectiveness of using a Gaussian low-pass filter to reduce high-frequency noise, the viability of using training data from a single site to classify other sites, and the sensitivity of classification accuracy to different image resolutions. Ultimately, we offer observations for those wishing to perform similar analyses. We focus on sUAS with consumer-grade cameras, rather than on near-IR, IR, or UV photography despite potential advantages of the latter, such as better differentiation of plant species using near-infrared data (Mitchell et al. 2016). Visible spectrum camera sUAS are less expensive and are flown by hundreds of thousands of hobbyists. If images from these cameras can provide useful land cover data, the potential contributions to habitat assessment by recreational pilots could be tremendous (Johnson et al. 2017).

Background

Many wildlife species, especially highly mobile organisms like birds, are believed to select habitat at multiple spatial scales (Wiens 1989; Mayor et al. 2009; Drever et al. 2015). For species with larger territories (e.g., >20 ha), it might not be practicable to survey vegetation features on foot, yet the organism may select habitat based on fine-scale variation in features like grassland structure (Cunningham and Johnson 2006).

Aerial photography can be used to assess fine-scale features across a broad area (e.g., Dufour et al. 2013; Kuzmin 2016), but in seasonal environments, available aerial and satellite photography is often taken when leaves are not present on trees, to maximize ground visibility. This makes much available aerial imagery less suitable for assessing vegetation structure during a bird’s breeding season, when rapid seasonal plant growth can change habitat structure markedly within weeks (Rotenberry and Wiens 1980). Additionally, traditional aerial photography cannot always provide images at a resolution fine enough to identify features of grassland habitat, whereas sUAS photographs can (Rango et al. 2009). To date, sUAS use to quantify avian habitat has been limited to classifying areas of broad, uniform cover types (Rodriguez et al. 2012), or high-contrast wetland habitats (Chabot et al. 2014).

Because kestrels are visual hunters, the physical structure of their foraging grounds plays a major role in their ability to find and capture prey (Bildstein and Collopy 1987), making the species an excellent candidate for habitat assessment via sUAS photography. They also have territories too large to survey practicably on foot (often >20 ha), yet the birds may cue on fine-scale features when selecting nest-sites (Smallwood and Bird 2002). This
study focuses exclusively on the area immediately surrounding the nest site, since the nest-site scale is the most sensitive to decisions about nest-box placement.

Materials and Methods

Field methods

Field sites were selected from known kestrel breeding areas in eastern Massachusetts, U.S.A. (42°15′–42°50′ N, 70°49′–2°38′ W). Aerial photographs were taken May–August, 2015, where permission to fly could be obtained. From this set, we randomly selected twelve kestrel boxes (six occupied, six unoccupied) for analysis in this study. Field sites had minimal terrain relief. Only two sites of 12 had elevation change >4 m across the photographed area, as assessed using Google Earth data. Both sites had gentle slopes with maximum grades <20%. We performed ground surveys around each nest site to assess vegetative cover, identifying major cover plants (those accounting for >10% cover by visual assessment) to species level where possible, and recorded their approximate distribution by sketching maps for use in ground truthing. These data were later used to define classifier training samples as particular cover types (e.g., grass, herbaceous) for classification in software ENVI (Environment for Visualizing Images, Harris Geospatial Solutions).

Flights were conducted by two researchers: one pilot responsible for flying the sUAS and image acquisition, and one spotter to monitor reactions and proximity of kestrels and other birds (as in Junda et al. 2015). At study sites where kestrels were known to be present, UAV flights elicited either no reaction or an alert reaction (sensu Mulero-Pázmany et al. 2017), consistent with findings by other researchers that flights by electric UAV 20–45 m from focal species are unlikely to provoke strong behavioral responses (Mulero-Pázmany et al. 2017). American Kestrels specifically are resilient to disturbances at the nest site, including capture and handling (Smallwood 2016).

Photogrammetry methods

We used a commercially available DJI Phantom Vision 2 Plus quadcopter drone with standard 14-megapixel camera, 6.17 × 4.55 mm sensor, and 4384 × 3288 pixel field for all photography. Photographs at each site were taken

Figure 2. Flow chart of the steps used to process and quantify drone photographs of habitat cover.
at 25 m and 50 m (± 0.1 m, as reported by the drone’s onboard altitude sensor) heights above takeoff level to assess habitat from altitudes at which prospecting kestrels might fly. Three photographs were taken at each site at each height, centered above the nest box facing nadir. The best-quality photograph from each set was used in subsequent analyses (Fig. 2). Photographs were taken within two hours of solar noon to reduce shadows.

**Photo processing and supervised classification**

Because the DJI Phantom 2 Vision Plus camera imposes a fisheye distortion, we corrected each.jpg photo using GNU Image Manipulation Program (GIMP) version 2.6 (https://www.gimp.org/). GIMP uses spherical correction to remove fisheye distortion via its “Lens Distortion” tool. This tool magnifies and crops the outer edges of the photograph in a radially symmetric way, and similar distortions have been shown to be effective for sUAS GoPro-style cameras in past studies (Nel 2015). Final ground areas covered by each corrected photo are reported below. Photos were not orthorectified due to a lack of ground control points. While this likely introduced minor error into area calculations, previous studies found that for images covering a small geographic extent in areas with minimal relief, orthorectification is not necessary (Tobias 2008). Corrected photos covered approximately 38 × 20 m at ground level from a height of 25 m, and 80 × 42 m from 50 m. This gave approximate ground size per pixel of 67.6 mm² and 326 mm², respectively. Corrected photos were imported to ENVI 5.3. ENVI is a widely used software platform for processing remote sensing data, including aerial photos (Goetz 2009). ENVI has been used in aerial photo analysis to investigate diverse vegetation research questions (Underwood 2003; Zhang et al. 2009). The steps we describe in processing and assessing images are depicted in Figure 2.

In ENVI, we applied a Gaussian low-pass filter with the default dimensions [3,5] to each image. This smooths spectral variation to improve feature identification by blurring high-frequency noise (e.g., Montabone 2010; Kim et al. 2015). This is a commonly recommended step in image processing (Burger and Burge 2016) that has been used in sUAS photography studies (Kim et al. 2015). This can facilitate discrimination between cover types by reducing spectral variation caused by, for example, multiple leaves in the same tree that reflect light at different angles. We assumed that filtering would improve cover type differentiation, but filtering could also interfere with classification by removing characteristic differences between cover types. To test these predictions, we compared kappa coefficients between filtered and unfiltered photos from all 12 sites at both heights.

We defined training samples for each cover type by drawing polygons on the raster with ENVI’s “Create ROIs from Geometry” digitization tool. Relevant cover types included Grasses (Poaceae), Herbaceous growth (non-woody plants not in the grass family), Dead/Bare Ground (sparse dead vegetation, bare earth), Human-modified (e.g., buildings, asphalt), and Woody (e.g., tree crowns, shrubs) (Fig. 3). Cover types were chosen because of their potential biological relevance to kestrels, which are visual hunters that typically capture prey on or near the ground.

Figure 3. Aerial photograph of Site 4, with Regions of Interest defined by colored outlines: light green = Grass, dark green = Herbaceous, brown = Bare Ground, Magenta = Woody, red = Human-modified (in this case, a path covered with wood chips). This photo has also been modified with a Gaussian low-pass filter to reduce noise and increase contrast between cover types. The gray arrow identifies location of the nest box. The arrow is 10 m long for scale.
after diving from above (Smallwood and Bird 2002). Bare
ground offers no cover for prey, which facilitates hunting,
but large areas of bare ground are unlikely to support
much prey. Grass and herbaceous growth can conceal
prey, but kestrels can pass through them to hunt. Woody
growth is difficult for kestrels to see through and to pass
through when pursuing prey, preventing birds from hunt-
ing effectively. Human-modified features, e.g., impervious
surfaces, have a wide range of effects on kestrels, includ-
ing potential exposure to noise and pollutants (Varland
and Loughin 1993), although kestrels can breed success-
fully in urban areas (Smallwood and Bird 2002). We did
not classify plant growth to finer detail, because kestrel
ecology did not justify it.

Using areas of known cover from ground surveys, we
drew polygons on each aerial photo within each cover type
to create training samples (Fig. 3). Similar habitat assess-
ment studies using ENVI showed that defining fewer cover
types (≤five) is most effective for achieving good differenti-
atation (Chabot et al. 2017). We selected representative sam-
ples of each cover type, ensuring that only one cover type
was in each training sample to avoid mistraining the classi-
fier. After we defined training samples, we analyzed each
photo through supervised classification using MAXLI-
KE_RUN, a wrapper for ENVI’s Maximum Likelihood
classifier (Canty 2010). This function randomly assigns 2/3
of the pixels in each ground truth polygon as training data,
and then tests the classifier on the remaining pixels
(Fig. 3). Many contemporary image analyses use object-
oriented approaches that group pixels with similar charac-
teristics into discrete objects, rather than the pixel-by-pixel
approach we used here. Object-oriented analysis is a
powerful approach, but maximum likelihood pixel-based
approaches are equal or superior to object-oriented
approaches when training sample sizes are small, e.g., when
there is only one image to work from (Yu et al. 2006).

Inter–observer repeatability of supervised
classification

Because there are subjective elements to creating training
samples, we were interested in determining inter–observer
classification accuracy. This is particularly important for
citizen-science projects. We gave five subjects (graduate
students in biology with limited or no image-processing
experience) the same image with instructions to select
training samples and perform maximum likelihood classifi-
cation. All subjects independently analyzed the same photo
taken from 50 m and filtered with a Gaussian low-pass fil-
ter. We compared their results with one-another and with
classification of the same image by the author (MK).

Classification postprocessing

Once image classification was completed, there often were
small patches of one habitat type imbedded within a lar-
gar habitat type, creating a “salt-and-pepper” effect (see
Fig. 4). This effect can be caused by misclassification
(e.g., identifying a tree leaf as a forb) and is well-docu-
mented when classifying high-resolution remotely sensed
imagery (Hsieh et al. 2001). We used ENVI’s Classification
Aggregation tool to assign these small clusters to sur-
rounding cover types. This tool sets a minimum patch
size for a cluster of pixels, and smaller clusters are

![Figure 4](image_url)

Figure 4. The same photograph of Site 4 as in Figure 4 that has been completely classified via Supervised Classification into grass (light green),
forbs (dark green), woody (magenta), bare ground (brown) and human-modified areas (red). Note the speckling effect (small pixels of one habitat
type embedded in a much larger different habitat type), which will be removed in the next processing step. The white arrow identifies location of
the nest box. The arrow is 10 m long for scale.
subsumed into the surrounding cover type. We were interested in determining what minimum patch size gave the most accurate and biologically informative classification. Too large a minimum could lose important small-scale information about habitat structure, while too small a minimum would not resolve “salt-and-pepper” errors. Chabot et al. (2014) used a minimum patch size of 0.25 m² when analyzing wetland habitat for a similarly sized bird, the Least Bittern (*Ixobrychus exilis*). Although American Kestrels are behaviorally different, this seemed to be a reasonable size from the perspective of evaluating territory features. Consequently, we used minimum patches of 0.25 m² at both 25 m and 50 m scales (for our images this was 3696 pixels and 768 pixels, respectively); to examine patch size sensitivity, we also evaluated 0.2 m² and 0.3 m² (± 20%). Also, since American Kestrels take small prey such as meadow voles (*Microtus pennsylvanicus*) (Smallwood and Bird 2002), we assumed that they might respond to finer-scaled habitat structure at ground level. Therefore, we evaluated a trio of smaller minimum patch sizes. A typical meadow vole measures about 170 mm long by 35 mm wide (approximately 5950 mm²); any ground feature of this size or larger might obscure prey from a hunting kestrel. Therefore, we processed minimum patches at this value (17 pixels for 25 m photographs), as well as 7 and 27 pixels, for both 25 m and 50 m photos.

**Confusion matrices**

After postprocessing, we created a confusion matrix for each image, which compares the computer’s pixel classifications within the user-defined test samples to the ground truth of those samples (Congalton 1983; Underwood 2003; Govender et al. 2008). The confusion matrix returns information for each image on how many pixels within test samples were correctly identified, as well as how many pixels were misidentified as other cover types (i.e., the degree to which the classifier was ‘confused’). The primary statistic of interest in the confusion matrix is the kappa coefficient (Foody 2002). This is a measure of how well a classifier discriminates among different types while taking into account the prevalence of each cover type within the image, and the chance of classifying each pixel correctly by chance. Its value can range from -1 (total disagreement) to 1 (total agreement); it is a more unbiased measure of a classifier’s utility than is overall accuracy (Titus et al. 1984). Kappa coefficients of 0.6–0.8 are considered to show good agreement with ground truth, while values of 0.8–1.0 show strong agreement (Landis and Koch 1977; Chust et al. 2008). We also recorded Producer’s and User’s Accuracies for each image (Supplemental Material).

**Table 1. Kappa coefficients of supervised classification for photos of each kestrel nest site.**

| Site | 25 m kappa | Rank | 50 m kappa | Rank | Cropped 50 m kappa | Rank |
|------|------------|------|------------|------|--------------------|------|
| 6    | 0.9879     | 1    | 0.9241     | 2    | 0.9608             | 1    |
| 8    | 0.8966     | 5    | 0.9973     | 1    | 0.8944             | 2    |
| 3    | 0.8998     | 4    | 0.8758     | 3    | 0.8299             | 5    |
| 7    | 0.9142     | 2    | 0.7502     | 8    | 0.8422             | 4    |
| 2    | 0.7721     | 9    | 0.8100     | 4    | 0.8630             | 3    |
| 9    | 0.8585     | 7    | 0.8360     | 6    | 0.7453             | 8    |
| 11   | 0.8949     | 6    | 0.8305     | 7    | 0.5877             | 10   |
| 4    | 0.8140     | 8    | 0.7940     | 5    | 0.7958             | 6    |
| 10   | 0.9102     | 3    | 0.4003     | 12   | 0.7351             | 7    |
| 5    | 0.7703     | 10   | 0.5118     | 9    | 0.6401             | 9    |
| 12   | 0.5757     | 12   | 0.7037     | 10   | 0.5642             | 11   |
| 1    | 0.5965     | 11   | 0.6173     | 11   | 0.4338             | 12   |
| Mean | 0.824      |      | 0.754      |      | 0.743              |      |
| SD   | 0.127      |      | 0.172      |      | 0.155              |      |

150 m photos cropped to the field of view of the 25 m photos.

2Higher rank indicates more accurate classification.

**Crossed comparisons**

Image analysis can take considerable time; in our study, processing each image took ~45 min, with 20-30 min of that time dedicated to creating training samples. We investigated whether training data from one image could be used to accurately classify habitat in other photos. If so, time could be saved, using a single master training image to classify others. We selected two sites that each had a diversity of cover types and used the training data from these sites to classify all other sites. We used the same minimum patch sizes in postprocessing, and the
test samples from the original image, to evaluate success of crossed classification. Therefore, each site was classified three times, each using a different training sample: the site itself (referred to as “self”), the data from Site 4, and the data from Site 9 (Sites are listed in Table 1). We then compared kappa coefficients across the three sets.

50 m cropped photos

By comparing photos taken from 25 m with photos taken from 50 m, we hoped to determine whether an increased field of view at a higher altitude had an associated loss of precision identifying cover types. To investigate this, we cropped each 50 m photo to include the same field of view as each 25 m photo. We analyzed these photos according to the methods above, including defining their own training samples to account for small differences in feature alignment between the 25 m and 50 m-cropped photos, which prevented us from using the training samples for the original 25 m photos.

Discriminating Kestrel Habitat

Once the best parameters for accurate classification were identified, we used the classified images to examine whether we could detect systematic differences in cover types around occupied and unoccupied kestrel nest boxes. We used a binomial GLM to test for effects of different cover types on likelihood of occupancy.

Results

Supervised Self-Classification

Self-classification performed better than crossed classifications at all flying heights. Self-classification of sites using a minimum patch of 0.25 m² had similar mean kappa values at 25 m, 50 m, and 50 m-cropped scales (Table 1). The 25 m images had the highest mean kappa and lowest standard deviation by a small margin. However, these means obscure changes in rank order of sites between image scales. For example, Site 10 shifts between having the third, seventh, and twelfth highest classification accuracy depending on which height is examined. This variation is important to consider when evaluating classification accuracy. Rank orders of kappa values by site changed a fair amount between height treatments.
(correlation of ranks, $r = 0.76$) (Table 1). Ranks for some sites changed substantially, and ranks of the 50 m Cropped images were significantly different from those of both 25 m ($P = 0.01$, $r = 0.71$) and 50 m ($P = 0.008$, $r = 0.72$) images (Table 1).

**Repeatability of supervised classification**

Six subjects independently created training samples for supervised classification (Table 2). Between-subject training sample patches differed widely in sizes (mean total area = 1 639 816 pixels ± 506 468; range 753 533–2 189 336). Despite this variation, supervised classification of all participants’ training samples achieved good accuracy (kappa $0.6$–$0.8$), and there was high inter-subject consistency in kappa values. Linear regression revealed no meaningful correlation between the size of training samples and accuracy of the classification ($r^2 = 0.38$).

**Effects of Gaussian low-pass filtering**

We found that Gaussian low-pass filtering significantly reduced kappa coefficients for our sites at 25 m (paired $t$-test, $P = 0.002$; Filtered mean = 0.697 ± 0.170 sd, Unfiltered mean = 0.824 ± 0.127) and at 50 m ($P = 0.0003$; Filtered mean = 0.699 ± 0.177, Unfiltered mean = 0.754 ± 0.172). Mean differences in kappa coefficients between unfiltered and filtered images was 0.127 at 25 m and 0.055 at 50 m, which suggests the unfiltered images were significantly more accurate at both scales than the filtered images.

**Effects of minimum patch size**

We evaluated six minimum patch sizes at each height class. At both spatial scales the larger minimum patch sizes gave statistically significantly higher kappa coefficients; an approximately 10% increase at 25 m (Small patch mean = 0.711 ± 0.164, Large patch mean = 0.824 ± 0.124, paired $t$-test, $P < 0.001$, df = 35), and 4% increase at 50 m (Small patch mean = 0.711 ± 0.170, Large patch mean = 0.753 ± 0.169, paired $t$-test, $P < 0.001$, df = 35) (Fig. 5). Comparisons of kappa coefficient sensitivity to minimum patch size showed some small differences. At the 25 m scale, mean kappa coefficients increased ~1% (0.01) at each increasing step size from 7 to 17 pixels ($P < 0.001$) and from 17 to 27 pixels ($P < 0.001$). For the larger minimum patch sizes at 25 m, differences between sizes did not produce significantly different mean kappa coefficients (all $P > 0.05$) (Fig. 5). At 50 m, differences between patch sizes followed a similar pattern: significant but small increases of ~1% at small patch sizes ($P < 0.01$ between 7, 17, and 27 pixel patches) and no significant differences between larger patch sizes (all $P > 0.05$, Fig. 5).

**Effectiveness of crossed classifications**

After previous analyses revealed that larger minimum patch sizes resulted in more accurate classification, the larger (0.25 m$^2$) minimum patch size was used in processing images for all subsequent analyses. Crossed classifications showed an overall much poorer fit compared to self-classification, as well as high variance within groups.

**Figure 6.** Kappa coefficients after supervised classification at 12 sites; 1.0 indicates perfect classification. Red: self-classification (red) Green: classifications using Region of Interest (ROI) data from Site 4 (green), and Site 9 (blue). Negative kappa coefficients indicate a classification accuracy that was worse than classification by random placement. Dashed vertical lines indicate mean values; solid black lines indicate medians.
At 25 m, crossed classification with Site 4 had a mean kappa coefficient of 0.340 ± 0.40. Crossed classifications with Site 9 had a mean kappa coefficient of 0.280 ± 0.31. These values are considerably lower than those from self-classification, which were 0.82 ± 0.13 at the 25 m scale and 0.75 ± 0.17 at 50 m (Fig. 6). The 50 m crossed comparisons showed similarly lower accuracies and high variances (Fig. 7).

As with self-classified data, mean values for cross classification also obscured inter-site changes in rank orders between treatments (Fig. 6). Some crossed classifications at the 50 m scale had negative kappa coefficients, indicating performance worse than random. These sites were classified as entirely Woody or entirely Herbaceous, presumably because those cover types in the reference images covered a wider spectral range than in the target image, resulting in all cover types being aggregated. Both crossed classifiers performed poorly, but they performed poorly in different ways. Classification by Site 4, for example, performed well at a small number of sites at both 25 m and 50 m, presumably due to similarities in cover species and light conditions with the reference site, but was outperformed by the classifier from Site 9 at most other sites (Fig. 7).

Finally, we examined which cover types were the most difficult to classify by calculating the average proportion of each cover type that was either correctly classified, or incorrectly classified as another cover type (Fig. 8). Bare Ground and Human-modified cover were confused with other cover types least often, while the three vegetative cover types (Grass, Herbaceous, Woody) were confused with each other more often. Woody was the cover type most often misclassified.

Use of Images to identify Kestrel Habitat

Analysis of the proportion of each cover type around occupied and unoccupied boxes showed no significant differences (P > 0.1 at all scales). This suggests that occupancy of nest cavities may not be driven by microhabitat immediately around the cavity.

Discussion

We achieved good or strong classification rates of different cover types at every site using visible spectrum sUAS photography. Past studies showed that visible spectrum sUAS photographs can be used to classify habitat more accurately than satellite images that incorporate near-infrared (NIR) spectral data (Dunford et al. 2009). That said, most studies of remotely sensed vegetation benefit from the inclusion of NIR spectral data. For image acquisition at low cost by volunteers and citizen scientists, however, our work suggests that off-the-shelf consumer camera photos can be useful in differentiating cover types.

This approach shows promise for fine-scale evaluation of wildlife habitat, but the steps taken in postprocessing can affect classifier performance. In our study, a Gaussian low-pass filter actually decreased classification accuracy. Some applications of a Gaussian low-pass filter to reduce noise focused on identifying features that are distinct from neighboring pixels, such as cracks in buildings (Kim et al. 2015). However, because the Gaussian low-pass filter blurs high-contrast edges, it could make differentiation more difficult between similar features such as two types of green vegetation. Consequently, we recommend that studies of wildlife habitat assessment using off-the-shelf.

Figure 7. Kappa coefficients for (A) 25 m photos and (B) 50 m photos using self-classification (circles), crossed classification with Regions of Interest (ROIs) from Site 4 (triangles), and crossed classification with ROIs from Site 9 (squares).
sUAS cameras evaluate their classifications with and without Gaussian filtering.

Results from repeatability of classification success for drawing training samples show low inter-subject variability. This suggests that firsthand site knowledge is not necessary to achieve good classification results. This is particularly encouraging with regards to researchers using photographs collected by citizen scientists. However, we caution readers to carefully evaluate crowd-sourced photos, and to offer guidelines, training, and feedback as well as subjecting data to secondary validation (See et al. 2013). We also note that there is likely some effect of where experimenters choose to draw training samples for supervised classification, as the range of kappa values in our study across all six trials was 0.68–0.79, indicating variability in training the classifier. We recommend that other researchers also examine the size of this effect so it can be better understood and controlled for in future studies.

Using different minimum patch sizes during aggregation of classified cover types also influenced classification accuracy, as might be expected in heterogeneous landscapes (Goodchild and Gopal 1989). At both 25 and 50 m heights, larger minimum patch sizes yielded more accurate results compared to smaller (vole-sized) minima. This suggests that, at least for our study, classification errors caused by speckling noise retained in the image by smaller patches outweighs the potential loss of habitat structure from larger patches. In many images, the largest minimum patch size tested gave the highest kappa coefficient, suggesting that the ideal minimum patch size for our images might be larger than any we evaluated. However, the precision of classifications from this method is limited by the precision of the hand-drawn ground truth sketches. Because original ground truth information was recorded in the field by hand, this necessarily imposed a lower limit on the size of the training polygons created in ENVI. Nonetheless, we feel that the results from the large patch classification gave good approximations of salient habitat features for kestrels.

Using training samples generated on one “master” site (image) to classify other sites resulted in much poorer classification results than creating unique training samples for each image. This is likely not surprising to remote sensing professionals, but we feel that the example can serve as a useful cautionary tale against attempting to shortcut a time-consuming step in the postprocessing workflow. Individual site differences in vegetation, angle of light, and weather can all undermine this approach. Fewer than 25% of sites showed good or better accuracy when using training samples from a different site to classify them. Although generating training samples is

---

Figure 8. Average proportions of accurately classified and misclassified cover types. For example, the top-left chart at 25 m shows that the majority of Grass was correctly classified as Grass, with occasional misclassification as Herbaceous or Woody and rare misclassification as Bare Ground.
time-consuming, we recommend creating unique ones for each image to be classified.

Site effects were apparent in which categories were most frequently misclassified. In our study system, Bare Ground and Human-modified areas were consistently well-separated from other cover types, but the three vegetative cover types (Grass, Herbaceous, Woody) were less consistently identified. At some sites, (e.g., 3,4), vegetation types were readily separated, while at other sites (e.g., 1) all vegetative types showed significant misclassification. Sites with noticeable shadows cast by trees tended to have pixels over-assigned to Woody cover, which often has darker foliage. In some sites, differences between classification error rates at different heights were striking. Having a wider view and more areas to sample from sometimes reduced confusion across vegetative cover types. At Site 8, however, where the 50 m image included Woody cover while the 25 m and 50 m cropped images did not, the addition of a third category of vegetative cover introduced significant confusion. An increased error rate associated with a larger numbers of categories has been reported in other remote-sensing studies (Dronova et al. 2012; Chabot and Bird 2013). This problem might be resolved with data supplementation, for example, digital surface models (DSMs) of vegetation created by LIDAR added as an additional layer. Data about vegetation height would help distinguish Woody and Herbaceous plants from Grasses (Su and Bork 2007).

So, can off-the-shelf consumer sUAS photography aid researchers in classifying habitat features for birds? We believe our results show that it can, as long as precautions are taken during data collection and processing. Consumer-grade cameras used on commercial drones may differ by pixel size, sensor size, focal length, and other characteristics. These characteristics should be reported to help control variation introduced in this manner. Obviously, this approach will not be ideal for all study systems; discrimination between different plant species or groups of green vegetation in a forest environment remains a vexing problem in remote sensing (Dunford et al. 2009). Sites with pronounced relief, such as mountainous areas, are notoriously susceptible to geometric distortion, requiring image orthorectification when assessing proportions of different cover types (Rocchini and Di Rita 2005) or integrating images into spatially explicit datasets. That said, in grassland habitats where interspersion of patches of bare ground or particular shrubs at small spatial scales affects plant community composition and physiognomy (e.g., Fisher and Davis 2010), citizen-collected sUAS images without ground control points can still be useful to researchers. The potential utility of drone photography in wetland habitats, where water is interspersed with emergent vegetation, has already been demonstrated (Chabot et al. 2014). It would be particularly useful for research questions that investigate changes in vegetation across short time scales (weeks), and for habitat types that are difficult to access, such as mangrove (Rhizophoraceae) forest.

Acknowledgments
We thank Tufts Institute of the Environment, a Tufts Faculty Research Award, and the NSF REU for grants supporting this research. This manuscript benefited from the comments of four anonymous reviewers. Shana Hodosh and Jordyn Fisk assisted in transporting and operating the UAV in the field. Jonathan Gale, Patrick Florance. And Jar-lath O’Neil-Dunne were valuable resources for questions related to drones, remote sensing, and image analysis in ENVI. We also thank Zarin Machanda and Julia Pilowsky for assistance in data analysis, and Rory Fuller, Brenna Gormally, Natalie Kerr, Jessica Rozek, and Charles van Rees for participating in our repeatability analysis.

References
van Andel, A. C., S. A. Wich, C. Boesch, L. P. Koh, M. M. Robbins, J. Kelly, et al. 2015. Locating chimpanzee nests and identifying fruiting trees with an unmanned aerial vehicle: locating chimpanzee nests with UAVs. Am. J. Primatol. 77, 1122–1134.
Bildstein, K. L., and M. W. Collopy. 1987. Hunting behavior of Eurasian (Falco tinnunculus) and American Kestrels (F. sparverius): a review. Pp. 66–82 in D.M. Bird, R. Bowman, eds. The ancestral kestrel. Raptor Res. Rep. No. 9.
Burger, W., and M. J. Burge. 2016. Digital Image Processing: an Algorithmic Introduction Using Java. Springer, London, UK.
Canty, M. J. 2010. Image Analysis, Classification, and Change Detection in Remote Sensing with Algorithms for ENVI/IDL. Second Edition. CRC Press, Boca Raton, Florida.
Chabot, D., and D. M. Bird. 2013. Small unmanned aircraft: precise and convenient new tools for surveying wetlands. J. Unmanned Veh. Syst. 01, 15–24.
Chabot, D., V. Carignan, and D. M. Bird. 2014. Measuring habitat quality for least bitterns in a created wetland with use of a small unmanned aircraft. Wetl. 34, 527–533.
Chabot, D., C. Dillon, O. Ahmed, and A. Shemrock. 2017. Object-based analysis of UAS imagery to map emergent and submerged invasive aquatic vegetation: a case study. J. Unmanned Veh. Syst. 5, 27–33.
Chretien, L.-P., J. Théau, and P. Ménard. 2016. Visible and thermal infrared remote sensing for the detection of white-tailed deer using an unmanned aerial system: detection of White-tailed Deer Using an UAS. Wildl. Soc. Bull. 40, 181–191.
Chust, G., I. Galparsoro, Á. Borja, J. Franco, and A. Uriarte. 2008. Coastal and estuarine habitat mapping, using LIDAR.
height and intensity and multi-spectral imagery. *Estuar. Coast. Shelf Sci.* 78, 633–643.

Congalton, R. G. 1983. Assessing landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogramm. Eng.* 49, 1671–1678.

Cunningham, M. A., and D. H. Johnson. 2006. Proximate and landscape factors influence grassland bird distributions. *Ecol. Appl.* 16, 1062–1075.

Drever, M. C., L. W. Gyg, J. Nielsen, A. K. Stuart-Smith, I. P. Ohanjanian, and K. Martin. 2015. Using field data and GIS-derived variables to model occurrence of Williamson’s Sapsucker nesting habitat at multiple spatial scales. *PLoS ONE* 10, e0130849.

Dronova, I., P. Gong, N. E. Clinton, L. Wang, W. Fu, S. Qi, et al. 2012. Landscape analysis of wetland plant functional types: the effects of image segmentation scale, vegetation classes and classification methods. *Remote Sens. Environ.* 127, 357–369.

Dufour, S., I. Bernex, J. Betheder, S. Corgne, L. Hubert-Moy, J. Nabucet, et al. 2013. Monitoring restored riparian vegetation: how can recent developments in remote sensing sciences help? *Knowl. Manag. Aquat. Ecosyst.* 410, 10.

Dundon, R., K. Michel, M. Gagnage, H. Piégay, and M.-L. Trémeolo. 2009. Potential and constraints of unmanned aerial vehicle technology for the characterization of mediterranean riparian forest. *Int. J. Remote Sens.* 30, 4915–4935.

Fisher, R. J., and S. K. Davis. 2010. From wiens to robel: a review of grassland-bird habitat selection. *J Wildl Manage* 74:265–273.

Foody, G. M. 2002. Status of land cover classification accuracy assessment. *Remote Sens. Environ.* 80, 185–201. Available from: http://www2.geog.ucl.ac.uk/~mdisney/teaching/teachingNEW/GOOGG141/papers/foody.pdf [accessed 1 January 2017]

Goetz, A. F. H. 2009. Three decades of hyperspectral remote sensing of the earth: a personal view. *Remote Sens. Environ.* 113, S5–S16.

Goodchild, M. F., and S. Gopal. 1989. *The Accuracy of Spatial Databases*. CRC Press, London, UK.

Govender, M., K. Chetty, V. Naiken, and H. Bulcock. 2008. A comparison of satellite hyperspectral and multispectral remote sensing imagery for improved classification and mapping of vegetation. *Water SA* 34, 147–154.

Hamerstrom, F., F. N. Hamerstrom, and J. Hart. 1973. Nest boxes: an effective management tool for kestrels. *J. Wildl. Manag.* 37, 400.

Hofstetter, A., C. M. Baylis, R. Mott, A. Herrod, and R. H. Clarke. 2016. Precision wildlife monitoring using unmanned aerial vehicles. *Sci. Rep.* 6, 22574.

Hsieh, P.-F., L. C. Lee, and N.-Y. Chen. 2001. Effect of spatial resolution on classification errors of pure and mixed pixels in remote sensing. *IEEE Trans. Geosci. Remote Sens.* 39, 2657–2663.

Johnson, P., B. Ricker, and S. Harrison. 2017. Volunteered Drone Imagery: Challenges and constraints to the development of an open shared image repository.

Proceedings of the 50th Hawaii International Conference on System Sciences. Available from: http://scholarspace.manoa.hawaii.edu/handle/10125/41396 [accessed 23 February 2017]

Junda, J., E. Greene, and D. M. Bird. 2015. Proper flight technique for using a small rotary- winged drone aircraft to safely, quickly, and accurately survey raptor nests. *J. Unmanned Veh. Syst.* 3, 222–236.

Katzner, T., S. Robertson, B. Robertson, J. Klucarsits, K. McCarty, and K. L. Bildstein. 2005. Results from a long-term nest-box program for American Kestrels: implications for improved population monitoring and conservation. *J. Field Ornithol.* 76, 217–226.

Kim, J.-W., S.-B. Kim, J.-C. Park, and J.-W Nam. 2015. Development of Crack Detection System with Unmanned Aerial Vehicles and Digital Image Processing. Advances in Structural Engineering and Mechanics (ASEM15). Available from: http://www.i-asem.org/publication_conf/asem15/5.JCSSS15/1w3F2_2MS852_2545F1.pdf [accessed 8 February 2017]

Kuzmin, A. 2016. Automatic segment-level tree species recognition using high resolution aerial winter imagery. *Eur. J. Remote Sens.* 49, 239–259.

Landis, J. R., and G. G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33, 159.

Liébana, M. S., J. H. Sarasola, and M. Á. Santillán. 2013. Nest-box occupancy by neotropical raptors in a native forest of Central Argentina. *J. Raptor Res.* 47, 208–213.

Mayor, S. J., D. C. Schneider, J. A. Schaefer, and S. P. Mahoney. 2009. Habitat selection at multiple scales. *Ecocience* 16, 238–247.

McCulre, C. J. W., B. P. Pauli, and J. A. Heath. 2017. Simulations reveal the power and peril of artificial breeding sites for monitoring and managing animals. *Ecol. Appl.* 27:1155–1166.

Mitchell, J., N. Glenn, M. Anderson, and R. Hruska. 2016. Flight Considerations and hyperspectral image classifications for dryland vegetation management from a fixed-wing UAS. *Environ. Manag. Sustain. Dev.* 5, 41.

Montabone, S. 2010. Pp. 106–109 Beginning digital image processing: using free tools for photographers, 1st ed. Apress, New York, USA.

Mulero-Pazmány, M., S. Jenni-Eiermann, N. Strebel, T. Sattler, J. J. Negro, and Z. Tablado. 2017. Unmanned aircraft systems as a new source of disturbance for wildlife: a systematic review. *PLoS ONE* 12, e0178448.

Nel, J. F. 2015. Post-processing of UAV-captured images for enhanced mapping by image stitching. In 2015 IEEE 5th International Conference on Consumer Electronics - Berlin (ICCE-Berlin) (pp. 500–504). Berlin, Germany: IEEE. https://doi.org/10.1109/icce- Berlin.2015.7391322

Rango, A., A. Laliberte, J. E. Herrick, C. Winters, K. Havstad, C. Steele, et al. 2009. Unmanned aerial vehicle-based remote sensing for rangeland assessment, monitoring, and management. *J. Appl. Remote Sens.* 3, 033542–033542.
Rocchini, D., and A. Di Rita. 2005. Relief effects on aerial photos geometric correction. *Appl. Geogr.* **25**, 159–168.

Rodríguez, A., J. J. Negro, M. Mulero, C. Rodríguez, J. Hernández-Pliego, and J. Bustamante. 2012. The eye in the sky: combined use of unmanned aerial systems and GPS data loggers for ecological research and conservation of small birds. *PLoS ONE* **7**, e50336.

Rossi, M., and D. Brunelli. 2016. Autonomous gas detection and mapping with unmanned aerial vehicles. *IEEE Trans. Instrum. Meas.* **65**, 765–775.

Rotenberry, J. T., and J. A. Wiens. 1980. Temporal variation in habitat structure and shrubsteppe bird dynamics. *Oecologia* **47**, 1–9.

See, L., A. Comber, C. Salk, S. Fritz, M. van der Velde, C. Perger, et al. 2013. Comparing the quality of crowdsourced data contributed by expert and non-experts. *PLoS ONE* **8**, e69958.

Smallwood, J. A. 2016. Effects of researcher-induced disturbance on American Kestrels breeding in nest boxes in northwestern New Jersey. *Journal of Raptor Research* **50**, 54–59.

Smallwood, J. A., and D. M. Bird. 2002. American Kestrel (*Falco sparverius*). The Birds of North America Online. https://doi.org/10.2173/bna.602

Smallwood, J. A., and M. W. Collopy. 2009. Southeastern American kestrels respond to an increase in the availability of nest cavities in North-Central Florida. *Journal of Raptor Research* **43**, 291–300.

Smallwood, J. A., M. F. Causey, D. H. Mossop, J. R. Klucarits, B. Robertson, S. Robertson, et al. 2009. Why are American Kestrel (*Falco sparverius*) populations declining in North America? evidence from nest-box programs. *Journal of Raptor Research* **43**, 274–282.

Su, J. G., and E. W. Bork. 2007. Characterization of diverse plant communities in Aspen Parkland rangeland using LiDAR data. *Appl. Veg. Sci.* **10**, 407–416.

Titus, K., J. A. Mosher, and B. K. Williams. 1984. Chance-corrected classification for use in discriminant analysis: ecological applications. *Am. Midl. Nat.* **111**, 1.

Tobias, M. M. 2008. Monitoring shorebird habitat using photogrammetry: the case of western snowy plover at coal oil point reserve, Santa Barbara, California. *Phys. Geogr.* **29**, 275–288.

Toland, B. R., and W. H. Elder. 1987. Influence of nest-box placement and density on abundance and productivity of American Kestrels in central Missouri. *Wilson Bull.* **99**, 712–717. Available from http://www.jstor.org/stable/4162483 [accessed 27 January 2016]

Underwood, E. 2003. Mapping nonnative plants using hyperspectral imagery. *Remote Sens. Environ.* **86**, 150–161.

Vanian, J. 2016. *Drone Sales Are About To Go Crazy*. Available from: http://fortune.com/2016/03/25/federal-governmen-drone-sales-soar/

Varland, D. E., and T. M. Loughin. 1993. Reproductive success of American Kestrels nesting along an interstate highway in central Iowa. *Wilson Bull.* **105**, 465–474.

Wiens, J. A. 1989. Spatial scaling in ecology. *Funct. Ecol.* **3**, 385.

Yu, Q., P. Gong, N. Clinton, G. Biging, M. Kelly, and D. Schirokauer. 2006. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogramm. Eng. Remote Sensing* **72**, 799–811.

Zhang, H., Y. Lan, R. Lacey, W. C. Hoffmann, and Y. Huang. 2009. Analysis of vegetation indices derived from aerial multispectral and ground hyperspectral data. *Int. J. Agric. Biol. Engin.* **2**, 33–40.

**Supporting Information**

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

Summary of proportions of land cover types across all sites at all resolutions.