



Deep learning analysis of nest camera video recordings reveals temperature-sensitive incubation behavior in the purple martin (*Progne subis*)

Heather M. Williams¹ · Robert L. DeLeon²

Received: 28 June 2019 / Revised: 4 December 2019 / Accepted: 13 December 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Incubation is a key life history stage for birds, and incubation attentiveness can have significant fitness consequences for both parents and offspring. Incubation is, however, a challenging phenomenon to observe and studies generally either measure some proxy of the target behavior, or risk disturbing birds through direct observation. More recently, nest cameras have provided a non-intrusive way to directly observe incubation, but analysis of these data is time-consuming. Here, we use the results of the first deep learning model which automated analysis of nest camera video recordings from eight purple martin (*Progne subis*) nests over the entire incubation period at a 1-s resolution. We mathematically define the initiation of incubation, characterize the change in nest attentiveness during incubation, and analyze the factors determining nest attentiveness and on- and off-bout duration during the incubation process. A random forest regression model identified the most important predictors of nest attentiveness. Attentiveness decreased with increasing temperature, but the strength of this response increased above the presumed physiological zero egg temperature, below which egg development ceases. This implies that the purple martins are able to adjust their incubation behavior in a complex, multiple-state manner to an extrinsic stimulus. Our study highlights the value of high-resolution datasets created using artificial intelligence for the analysis of nest camera video recordings of animal behavior.

Significance statement

The use of artificial intelligence for image classification tasks is becoming commonplace in society. This technology is beginning to be used to automate the analysis of video recordings of wildlife behavior. Here, we use the results of the first such classification from nest camera video recordings of the purple martin (*Progne subis*) to determine the factors affecting incubation attentiveness (the proportion of time that the adults spend in contact with eggs). Incubation attentiveness is important because it can affect hatch rate and have carry-over effects both for the condition of the incubating adults and the quality of the resulting offspring. Our analysis found that attentiveness was mainly affected by ambient temperature, with incubating adults reducing their efforts as ambient temperature reaches the minimum threshold for egg development.

Keywords Incubation · Purple martins · Nest cameras · Deep learning · Physiological zero · Ambient temperature

Communicated by J. A. Graves

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00265-019-2789-2>) contains supplementary material, which is available to authorized users.

✉ Heather M. Williams
hw49@buffalo.edu

¹ Department of Environment and Sustainability, State University of New York at Buffalo, Buffalo, New York, USA

² Center for Computational Research, State University of New York at Buffalo, Buffalo, New York, USA

Introduction

How to collect high-quality behavioral data from wild animals without disturbing them and compromising the integrity of the data is a perennial problem in behavioral ecology. It is a particularly relevant problem when studying breeding behaviors in birds due to the cryptic nature of many birds' nests and the fear of nest abandonment and reduced hatching success due to experimental disturbance (Blackmer et al. 2004; Ibanez-Alamo et al. 2012). Direct observation (when it is even possible) is extremely time-consuming and demanding for the researcher, meaning that most studies cover just a small

proportion of the breeding period for a small part of each day (Smith et al. 2015). In recent years, real-time nest video cameras have become common place in behavioral studies of breeding birds (Cox et al. 2012). Video cameras have the advantage of decreased disturbance to birds compared with direct observation. Viewing real-time recordings however is extremely time-consuming and data are usually subsampled (Davis and Holmes 2012), meaning that despite the technical capacity to study the full breeding period, the vast majority of studies are still forced to use some subset of the data. Amininasab et al. (2016) provided a major advance to this situation with their “BirdBox” software which partially automates viewing of nest box footage. The program provides the most likely times of adult entry and exit points based on frame by frame luminescence. However, the researcher must still manually verify these likely events.

The use of deep learning convolutional neural networks as tools to automate image recognition and classification is becoming commonplace in computing. These same techniques are starting to be successfully applied to wildlife camera footage and have proven capable of accurately identifying and counting species given an adequately labeled training dataset (Serrano et al. 2018; Tabak et al. 2018; Willi et al. 2018). Automating the classification of incubation behavior removes (or reduces) the need to subsample data and makes analysis of the entire incubation period both feasible and scalable (Williams et al. 2019). Deep learning has recently been used for the first time with nest camera video recordings to automate the classification of incubation behavior in the purple martin (*Progne subis*) (Williams et al. 2019). The eastern subspecies of purple martin (*Progne subis subis*) is an ideal species with which to test the utility of deep learning for incubation analysis. The species nests colonially, providing a relatively large potential sample size all in the same geographic location (Allen and Nice 1952). It also nests exclusively in man-made nest boxes (Tarof and Brown 2013), making it easy to standardize factors associated with the nest box and to install nest cameras. Purple martins are single-brooded female-only incubators, with an incubation period of 15–17 days and a typical clutch size of 3–6 eggs (Allen and Nice 1952).

Incubation behavior in birds is a key life history trait with potentially high fitness consequences for offspring. Patterns of parental attentiveness during incubation can determine hatchling success (Deeming 2002), affect embryo developmental rate (Martin and Schwabl 2008) and offspring phenotype (Hepp et al. 2006), and they can alter the risk of nest predation (Martin et al. 2000) and can shape the egg microbial environment (Cook et al. 2005). Incubation typically represents a costly energetic investment for the incubating parent(s) (Williams 1996) with attentiveness necessarily trading off against other essential behaviors such as feeding (Deeming 2002). As such, incubation behavior has been studied relatively intensely in birds (Deeming and Reynolds 2015) with

researchers using it as a model to test life history evolutionary hypotheses (Martin 2004).

Incubation behavior is often defined by attentiveness and on- and off-bout durations. Nest attentiveness (or constancy) is defined as the percentage of time that eggs are in contact with an adult (Skutch 1962), while on- and off-bouts are the average durations of time that the incubating bird spends in and out of contact with the eggs respectively (Deeming 2002). Researchers have often strived to determine intrinsic and extrinsic factors determining incubation behavior. Attentiveness is known to vary by taxonomic group, with Passeriformes averaging 75% (Deeming 2002). Within the passerines, there is further variation between species where both parents incubate and only females incubate, with shared incubators having significantly higher overall nest attentiveness (Deeming 2002). Within a species, attentiveness has sometimes been found to vary with clutch size, with larger clutches demanding greater attentiveness (Blagosklonov 1977). Cooper and Voss (2013) found that egg cooling rate increases with egg age, leading to the expectation of increased attentiveness and on-bout durations and/or decreased off-bout durations throughout the incubation period, but empirical evidence in support of this is so far sparse in the literature (Deeming 2002). Finally, there is an expected relationship between incubation behavior and ambient temperature, with birds increasing attentiveness in response to ambient temperatures below egg physiological zero (PZT—the temperature below which egg development is suspended) and decreasing attentiveness above PZT (Conway and Martin 2000). Theoretically, this increase in attentiveness at low temperatures is characterized by shorter off-bouts as incubators avoid exposing eggs to cold temperatures for long durations (Conway and Martin 2000).

Despite decades of interest in incubation behavior, however, the underlying datasets are often limited by the difficulties in collecting data for the entire incubation period and directly observed datasets are generally limited in scope both in terms of the number of individuals viewed and the resolution and coverage of the data from each nest (Davis and Holmes 2012; Smith et al. 2015). In response to this limitation, researchers have turned to the use of in-nest devices which capture some proxy of incubation attentiveness. Temperature sensors have been used, for example, to infer incubation in prairie chickens (*Tympanuchus cupido*) (Dallmann et al. 2016) and blue tits (*Cyanistes caeruleus*) (Amininasab et al. 2016). Pressure-sensitive devices have been tested with the common eider (*Somateria mollissima*) (Bottitta et al. 2002), and PIT tags have been used to measure nest arrivals and departures in Leach’s storm petrels (*Oceanodroma leucorhoa*) (Zangmeister et al. 2009). While these proxy devices are able to collect data more consistently throughout the incubation period, they can never definitively

demonstrate that incubation is actually taking place (Smith et al. 2015).

Here, we use a deep learning model (Williams et al. 2019) to estimate nest attentiveness and its corollaries, on- and off-bout duration, from purple martin nest camera recordings from the entire period starting with egg laying to incubation and post incubation periods. We use change point analysis to mathematically identify the initiation of incubation for each nest. We then use a random forest modeling approach to test whether purple martin incubation behavior is constant throughout this period, and whether it is sensitive to extrinsic factors such as ambient temperature and rainfall and/or intrinsic factors such as clutch size and the age of the incubating female.

Methods

Field data collection

Purple martin artificial gourds (Troyer horizontal gourd Conley II with tunnel entrance, Part THGC) were fitted with in-nest cameras with a wide-angle lens (CCTV HD-TVI Mini Security Camera 1080P 2.0MP 15 mm lens) at our field site at Iroquois National Wildlife refuge in Western New York State (43.111°N, 78.409°W). We modified the gourd cap by removing the opaque cover and replacing it with translucent plastic to allow more light into the nest. Cameras were installed prior to nest building. Cameras were set to record to a security DVR (Alibi 4000 Series 16-Channel HD-TVI 3.0 Hybrid+) from 6 a.m. until 9 p.m. from nest initiation until all eggs hatched. Data were not recorded overnight as there was insufficient light in the nests to discern incubation behavior in the video. Data were collected for eight purple martin nests (1 in 2017 and 7 in 2018) for an approximate total of 2280 h of footage. Our sample size of eight nests was determined by a combination of nest box occupancy and nests for which we obtained footage of the full incubation sequence. Video streams were automatically split and stored as individual 33-min files. Critical dates and information about each nest are available in Table 1.

Our field site hosts a weather station approximately 100 m from our nest boxes which recorded ambient temperature, precipitation, and wind-speed every hour throughout our study period. Temperature sensors (iButton DS1923 Humidichrons) were placed in the bottom of each nest box such that they were situated underneath the purple martins' nests. We did not observe any evidence of the purple martins attempting to remove or relocate sensors. Temperature was recorded every hour, and the difference between ambient temperature and nest temperature was calculated.

Deep learning neural network

We trained a neural network to automatically classify each video frame as “incubating” or “not incubating.” Full details of this process are available in Williams et al. (2019), but briefly, a tensor-flow model was trained on 12,144 manually classified images from our dataset. For this classification, “incubation” was defined as when one or more adult birds were in the nest and at least one of them was sitting on the eggs. Images were classified as “not incubation” when there was either no adult bird in the nest, or if an adult was present, but eggs were still exposed (Fig. 1). This definition of incubation does, therefore, include some instances of male birds sitting on nestlings although this cannot be technically classed as incubation due to the absence of a brood patch in male purple martins (Allen and Nice 1952). Each 33-min video clip took approximately 4 min to analyze when 1 frame per second was considered using only a single Nvidia Tesla P100 or V100 GPU node.

The resulting model produced a binary classification of incubation behavior at a 1-s resolution. Internal model testing showed the model classified 99.5% of still frames of validation images correctly. Modeled nest attentiveness was then expressed as a percentage for each 33-min video clip. We aggregated the 1-s data of our output to both negate the effect of any single frame misclassifications and to reach a resolution that is likely biologically meaningful for the hypotheses we wished to test. We further tested model accuracy by comparing modeled attentiveness percentages in 43 video clips with attentiveness estimated from manual viewing. Modeled and manually viewed attentiveness estimates were within 1% of each other (Williams et al. 2019).

To determine on- and off-bout durations, raw data of model output showing binary attentiveness values (incubating or not) at one point per second were aggregated for each channel to give the modal value per minute. This allowed us to estimate bout durations to the nearest whole minute and to disregard state changes of less than 30 s. These very short state changes were sometimes produced by single frame model classification errors and sometimes by behaviors such as when an incubating female would briefly stand up from the eggs, turn around, and sit back down—neither case would likely have a significant effect on egg temperature. The same set of 43 validation videos was manually classified for on- and off-bout durations, using the same approach of taking modal values per minute. Comparisons with this and the deep learning outputs never differed by more than 2 min.

Defining the start of incubation

Purple martins are generally thought to start incubation after laying their penultimate egg (Allen and Nice 1952; Tarof and Brown 2013) and are thus predicted to have 1 day of hatch

Table 1 Timing, clutch size, and adult age classes for each nest. Start of incubation is as defined by changepoint analysis (see “defining the start of incubation”). SY denotes a second-year bird (i.e., in its first potential breeding season) and ASY denotes an after-second-year bird (i.e., it has

potentially already bred at least once). Mean nest attentiveness refers to the mean attentiveness between 6 a.m. and 9 p.m. each day from the start of incubation until the first egg hatches. Bout duration is given in minutes

Nest ID	Date of first egg	Start of incubation	Date of first hatch	Days of hatch asynchrony	Clutch size	Age class of male/ female	Mean nest attentiveness	Mean on (off) bout duration
Ch01_2018	Jun 9 2018	Jun 11, 2018	Jun 27, 2018	2	5	SY/ASY	69.11%	16.04 (8.06)
Ch02_2018	May 20, 2018	May 22, 2018	Jun 8, 2018	1	6	ASY/ASY	72.74%	15.98 (7.06)
Ch04_2018	Jun 13/18	Jun 16, 2018	Jul 2, 2018	0	4	SY/ASY	68.23%	13.57 (6.68)
Ch05_2018	May 30, 2018	May 31, 2018	Jun 19, 2018	0	4	ASY/ASY	70.57%	14.57 (6.57)
Ch07_2018	May 22, 2018	May 25, 2018	Jun 9, 2018	1	5	ASY/ASY	76.44%	14.30 (5.70)
Ch09_2018	May 23, 2018	May 27, 2018	Jun 12, 2018	1	6	ASY/ASY	73.61%	14.69 (6.31)
Ch12_2018	May 26, 2018	May 20, 2018	Jun 14, 2018	2	6	ASY/SY	78.77%	20.98 (5.43)
Ch07_2017	Jun 17, 2017	Jun 18, 2017	Jul 6, 2017	1	4	SY/SY	69.67%	24.56 (10.31)

asynchrony where one egg hatches 1 day later than the rest of the clutch. As our nests had more variation in hatch asynchrony (0–2 days, Table 1) than expected by this rule, we decided to define the start of incubation empirically by using change point analysis to find the time at which attentiveness increases from a pre-incubation baseline to true-incubation for each nest. We limited the video data for each channel to the day the penultimate egg was laid, plus the 3 days prior and the 3 days after. Using the R package “change point” (Killick et al. 2016), we found change points in the mean and variance of nest attentiveness using the “At-most-one-change” (AMOC) method for each nest (Hinkley 1970). For one nest (ch04_2018), this method predicted the start of incubation 3.63 days after the penultimate egg was laid and would have resulted in an implausibly short incubation period of 13 days. For this nest, we therefore relaxed our criteria to allow more than one change point to be predicted by using the “Pruned-exact-linear-time” (PELT) method (Killick et al. 2012). To increase our power to illustrate the characteristics of the start of incubation, we made one final change point analysis (AMOC method) with the mean attentiveness values from

all nests included in a single analysis, with the data for each channel aligned on the penultimate egg laying day. As a check on the video data, we also did a similar AMOC change point analysis on the temperature difference data (i.e., nest temperature as determined by sensors minus the ambient temperature) from seven nests. Similar to the video data, we aligned the temperature data from the seven nests on the day of the penultimate egg.

Random forest modeling of factors determining incubation patterns

We used a random forest regression modeling approach to determine factors affecting incubation attentiveness using the randomForest R package (Liaw and Wiener 2002). The full dataset was subsetting to include only the period of “true” incubation (defined here as the start of incubation found in the change point analysis above, until the day that the first egg hatched) and was randomly split into training and testing sets (80:20). Candidate predictors included the day of incubation (where day 0 is the first day of incubation and incubation



Fig. 1 Typical images from the deep learning training set, illustrating incubation (a) and the two main cases of “non-incubation” (b, c)

ends when the first egg hatched); the time of day (between 6 a.m. and 9 p.m.); ambient temperature ($^{\circ}\text{C}$), rainfall (mm), wind speed ($\text{km}^{-\text{h}}$), ordinal date, clutch size, adult age classes (After-second-year (ASY) or Second-year (SY)), nest orientation (expressed as decimal degrees from North), and nest identity. We used AIC to select the most parsimonious set of predictors by starting with a full model and dropping the lowest-performing variable while calculating AIC at each step. The model with the lowest AIC was then tuned by finding the value of “mtry” (the number of variables selected as candidates at each split) and the number of trees which minimized the mean squared error on the training set. Errors on training and test data were examined; out-of-bag and validation performance was compared; predictor variables were ranked by importance; biological relationships were examined using partial plots and the variance explained was examined.

A correlation matrix of continuous predictors identified a strong positive correlation between ordinal date and day of incubation ($r = 0.44$, $N = 3282$, $P < 0.001$) (Online Resource 1: Fig. 1a). We therefore removed ordinal date from our set of predictors and re-ran the analysis.

Unsurprisingly, a repeated-measures (to account for pseudoreplication) correlation test showed there was also a strong positive correlation between each attentiveness measurement and its predecessor ($r = 0.64$, $N = 3271$, $P < 0.001$). To examine how this temporal autocorrelation affected the model, we calculated attentiveness at two additional temporal resolutions (mean daily attentiveness and three attentiveness measures per day (6 a.m.–9:30 a.m., 9:30 a.m. –5 p.m., 5 p.m.–9 p.m.)). We used repeated-measures correlation tests to check whether the positive correlation between each attentiveness measurement and its predecessor is similar at these temporal scales. We also repeated the main random forest analysis of these different temporal resolutions of attentiveness to check whether predicted relationships were similar for the original 33-min sampling frame.

Finally, we used on-bout duration and off-bout duration as the dependent variables in two random forest models with the same full set of predictors that was used for attentiveness. We used AIC for model selection using the same procedure as detailed for the attentiveness model.

Comparison with general linear mixed model

To provide a comparison with the results of the random forest model, we conducted a general linear mixed model (GLMM) of attentiveness using the same dataset and candidate predictors. Analysis was implemented in the R package lme4 (Bates et al. 2015) using nest identity as the random factor. Variables were selected for inclusion in a final model by stepwise backward elimination of non-significant predictors from the global model. The marginal and adjusted R^2 value was determined for the final model using the R package MuMIn (Bartoń 2018)

and fitted relationships visualized using the visreg package (Breheny and Burchett 2017).

Nest attentiveness and temperature

Ambient temperature was found to be an influential variable in all of the random forest models and in the GLMM. The relationship between ambient temperature and attentiveness was therefore examined further for adherence to a theoretically predicted threshold or hinge function. The R package “chngpt” (Fong and Permar 2017) was used to determine a statistically significant changepoint in the data. The splines package in base R was used to fit a linear spline function with a single knot at the pre-determined temperature changepoint. A linear regression was used to determine the statistical significance and R^2 between this function and nest attentiveness. Finally, we examined the relationship between ambient temperature and on- and off-bout durations to compare it with the attentiveness and ambient temperature fit.

All data analysis was conducted using R Studio version 1.1.383 (RStudio Team 2015) running R version 3.5.2 (R Development Core Team 2015). The use of artificial intelligence to classify incubation behavior (as opposed to manual classification) minimized the opportunity for observer bias and, in this sense, it can replace traditional blinded methods.

Results

All eight nests remained active throughout the incubation period and each nest achieved 100% hatching success. Nest attentiveness was found to be dependent on a number of variables but overall, during the incubation period (from the start of incubation until the first egg hatched), mean incubation attentiveness was 72%. Mean incubation attentiveness of individual nests varied from a minimum of 68.2% in ch04_2018 to a maximum of 78.8% in ch12_2018. On- and off-bout durations are also complicated functions of several variables but overall mean on-bout duration over the same period was 16 min and mean off-bout duration was 7 min (Table 1).

Determining the onset of incubation

Our combined analysis of mean values for all video channels indicated a mean incubation onset at 6 p.m. on the evening that the penultimate egg was laid (Fig. 2). This was largely supported by the temperature difference (in-nest temperature-ambient temperature) changepoint data from the seven nests that were instrumented with temperature sensors, which show an incubation onset of ~6 a.m. on the penultimate egg day (Online Resource 1: Fig. 3a).

At the individual nest level, for seven of eight nests, AMOC changepoint analysis placed the start of incubation

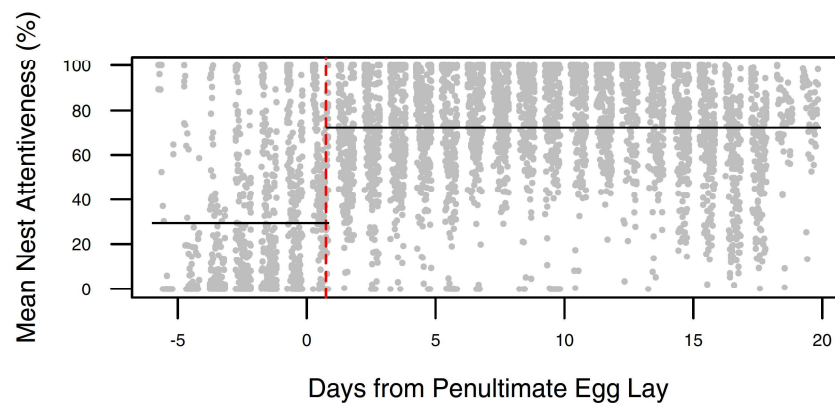


Fig. 2 Defining the onset of incubation. Each point represents an individual measurement of nest attentiveness at a 33-min temporal resolution. With data from all channels aligned on the day on which the penultimate egg was laid (the day on which the penultimate egg was laid is day zero on the x-axis), a changepoint in mean attentiveness was

detected at 0.75 days after the start of the day (shown with the red dashed line). This changepoint indicates the start of incubation. Black lines show mean attentiveness values before (29.42%) and after (72.36%) the start of true incubation

within 2 days of the penultimate egg lay (Table 1). Examination of the data shows that the start of incubation (i.e., the changepoint) is associated with very low attentiveness (0–0.2) becoming rare occurrences (Fig. 2, Online Resource 1: Fig. 2Aa–g). For the one channel where the AMOC changepoint method did not provide a plausible start date to incubation (ch04_2018), the relaxed criteria of the PELT method indicated an incubation initiation within 1 day of the penultimate egg being laid and conformed to a similar visual pattern as the other channels and was thus adopted as the start date (Online Resource 1: Fig. 2Ah).

Random forest modeling of factors determining incubation behavior

In our main random forest model (33-min resolution attentiveness with ordinal date included as a predictor), AIC model selection indicated the most parsimonious model included ambient temperature, nest identity, time of day, day of incubation, and ordinal date as predictors (Online Resource 1: Table 1A). This model explained 57% of the variance on the training dataset with low error rates (Table 2) and had a mean squared error of 0.025 on the testing dataset (or a mean absolute error of 12% attentiveness).

Ranking of variable importance showed that the model was mainly informed by ambient temperature, time of day and day of incubation; with minor effects of day of year and nest identity (Online Resource 1: Fig. 4A). Attentiveness decreased with increasing temperature above around 24 °C, increased at dawn and dusk; it slightly increased from incubation initiation until around day 6 of incubation; it slightly increased with ordinal date and showed some variation between individual nests (Fig. 3).

Removing the correlated predictor, ordinal date, from the model reduced the variation explained slightly to 56%, but as

variable rankings and error rates remained similar, we maintained ordinal date in the model for the main analysis as dictated by AIC (Online Resource 1: Tables 1A, 2).

Correlations between attentiveness measures and their previous measurement remained strong and positive at when measured daily and at a rate of three measures per day (3 daily: $r = 0.26$, $N = 355$, $P < 0.001$; daily: $r = 0.53$, $N = 114$, $P < 0.001$). Repeating the random forest analysis with the different temporal resolutions of attentiveness led to small changes in model fit, but predicted relationships and predictor rankings remained similar (Table 2). All subsequent analysis and discussion of attentiveness is therefore based primarily on the model which included ordinal date as a predictor and described attentiveness at a 33-min resolution (Table 2).

On-bout duration was most parsimoniously explained by ambient temperature, time of day, wind speed, nest identity, day of year, nest orientation, rainfall, and clutch size according to the AIC (Online Resource 1: Table 1A). The resulting random forest model of on-bout duration did not explain as much variation as seen in the attentiveness model (23% vs. 57%), but predictor rankings and shape of relationships were similar. On-bout duration decreased above 20 °C (with a noisy relationship at lower temperatures), increased at dawn and dusk, and peaked mid-way through the incubation period (Online Resource 1: Fig. 5A). Random forest modeling of off-bout duration was not informative, with none of the predictor variables having a clear effect and little variation in durations (Table 2). Off-bout duration did not vary substantially with temperature, time of day, or day of incubation (Online Resource 1: Fig. 5A).

Comparison with general linear mixed model

We compared the performance of a random forest model with that of a GLMM for the main attentiveness model. After

Table 2 Showing the results of random forest models. *MSE* mean squared error, *MAE* mean actual error (% attentiveness). Top three predictor variables are shown in ranked order according to their

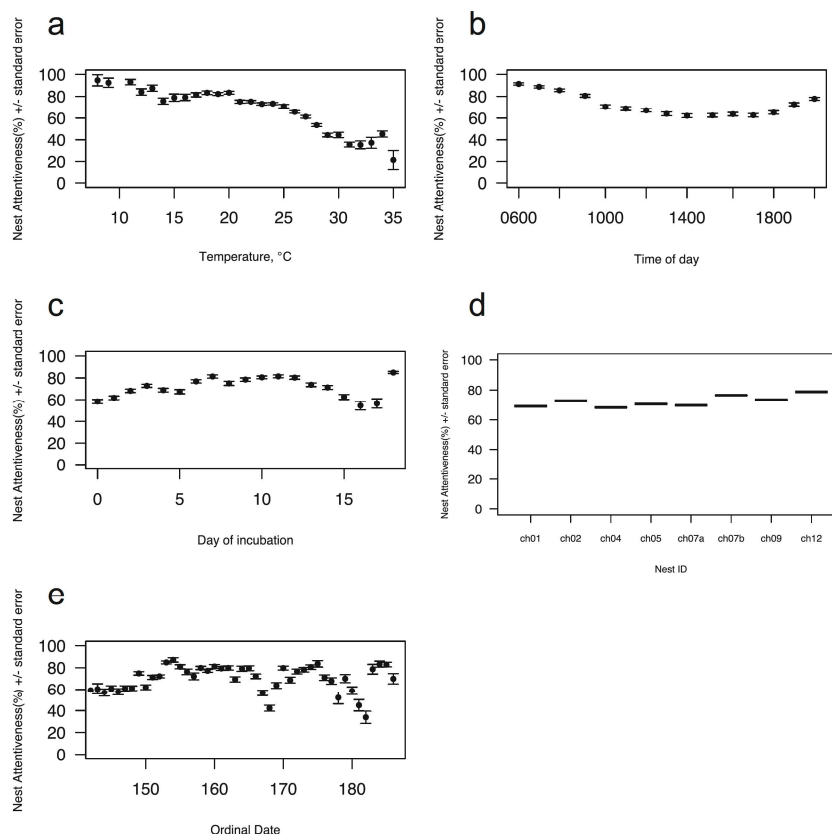
contribution to model fit. The attentiveness model shown in italics is the main model used in subsequent analysis

Incubation metric	Variance explained (%)	MSE	MAE	Top 3 predictor variables
<i>Attentiveness (33-min resolution)</i>	57.03	0.026	11.59%	<i>Temperature</i> <i>Time</i> <i>Day of incubation</i>
Attentiveness (33-min resolution) excluding ordinal date	55.96	0.025	11.54%	Temperature Day of incubation Time
Attentiveness (3/day resolution)	63.39	0.012	11%	Temperature Time Day of incubation
Attentiveness (daily resolution)	40.07	0.012	11%	Temperature Day of incubation Ordinal date
On-bout duration	23.22	497.01	10.68 min	Temperature Time Day of incubation
Off-bout duration	5.24	197.92	4.86 min	Ordinal date Clutch size Day of incubation

removing non-significant predictors from the model (clutch size, male age and female age), the GLMM had a marginal R^2 of 0.229 and a conditional R^2 of 0.235. The final model

predicted that nest attentiveness decreased with day of year ($Estimate = -1.84^{-3}$, $Df = 8$, $P = 0.046$), ambient temperature ($Estimate = -1.81^{-2}$, $Df = 3217$, $P < 0.001$), hour of day

Fig. 3 The relationship between incubation attentiveness and ambient temperature (a), time of day (b), day of incubation (c), nest ID (d), and ordinal date (e). Black circles show mean attentiveness values for each 1 °C increase in temperature (a), hour (b) or day (c, e). Error bars show standard error



($Estimate = -6.49^{-3}$, $Df = 3290$, $P < 0.001$), wind speed ($Estimate = -4.86^{-3}$, $Df = 3220$, $P < 0.001$), and rainfall ($Estimate = -2.59^{-2}$, $Df = 3288$, $P < 0.001$) but increased with day of incubation ($Estimate = 5.33^{-3}$, $Df = 3482$, $P < 0.001$) with only a very small effect of nest identity ($Nest\ variance = 0.0003407$, $Residual\ variance = 0.0450431$).

Nest attentiveness and temperature

Change point analysis found a statistically significant change point in the relationship between nest attentiveness and temperature at 24 °C ($ML\ ratio\ maximal\ statistic = 173.91$, $P < 0.001$). Fitting this spline to the data with a single linear hinge at 24 °C gave a significant fit ($F = 512.2$, $P < 0.001$, $multiple\ R^2 = 0.238$) (Fig. 4).

Discussion

Use of a deep learning model to automate estimations of nest attentiveness from nest camera data provided us with a dataset at 1-s temporal resolution covering the entire incubation period of purple martins with minimal disruption to the birds and was only limited by inadequate in-nest lighting preventing the collection of data overnight. Validation of this dataset showed that the deep learning model approached human accuracy in estimating nest attentiveness (Williams et al. 2019).

Two studies have previously attempted to measure incubation nest attentiveness in the purple martin. Kendeigh (1952) used a thermocouple in a single Purple Martin nest for 6 days of the incubation period and found an average attentiveness of 77%, and Allen and Nice (1952) directly observed three nests for a 4-h period in the middle of the day and estimated attentiveness at 70%. Our mean estimation of attentiveness (72%) sits comfortably between these estimates and, more broadly, in line with estimated passerine attentiveness of 75% (Deeming 2002). Our estimate of daytime mean on-bout duration of 16 min is also in-line estimates for other passerine cavity nesters such as estimated on-bouts of 15 min in tree swallows

(*Tachycineta bicolor*) (Ardia et al. 2010) and 17.5 min in European starlings (*Sturnus vulgaris*) (Reid et al. 1999).

Attentiveness should be relatively robust to small “single frame” errors in deep learning classifications as the actual result is averaged over many frames. Conversely, although we found strong adherence between modeled and manually validated videos, modeled on- and off-bout durations are likely to be more error prone as a single frame error can substantially change the measurement. Here, we mitigated this error by disregarding on- and off-bouts of less than 30 s. Although we found similar relationships between on-bout durations and attentiveness, we focus our analysis on attentiveness as the better-determined metric.

Our attentiveness metric could be considered an overestimate as it includes instances of males sitting on eggs in addition to “true” incubation by females. Although male purple martins are not generally considered to assist with incubation and do not have a brood patch (Allen and Nice 1952), we found that ~15% of the total attentiveness in our set of manually viewed videos was achieved by males. Furthermore, the males tended to fully sit on the eggs in the absence of the females in a manner consistent with incubation, as opposed to only egg guarding. While thermal transfer is likely to be reduced in male “incubation” compared with female incubation with a brood patch, as there will certainly still be some incubating effect, we do not see its inclusion as a shortcoming in our analysis.

Nest attentiveness in any given period was highly positively correlated with nest attentiveness in the previous period in three different temporal resolutions. Furthermore, attentiveness in all three temporal resolutions was principally described by the same covariates—namely temperature, time of day, and some measure of date (ordinal or day of incubation). This implies that attentiveness is driven by external conditions (such as those included in our modeling) as opposed to internal states, i.e., birds do not reduce attentiveness in a period simply because they already invested in high attentiveness in the previous period, but rather adjust attentiveness to environmental need.

Although it is generally thought that purple martins commence incubation after laying their penultimate egg (Allen and Nice 1952; Tarof and Brown 2013), our data showed some individual variation in this trait, with incubation starting up to 1 day before or after this event. This variation, coupled with the gradual increase in pre-incubation attentiveness observed in some nests, highlights the importance of taking an individual nest-level approach to defining the start of incubation.

We found strong evidence that purple martins adjust their incubation behavior in response to ambient temperature, with a small decline in attentiveness from 10 to 24 °C, and a much sharper decline above 24 °C. This appears to be achieved by decreasing on-bout duration while maintaining a relatively

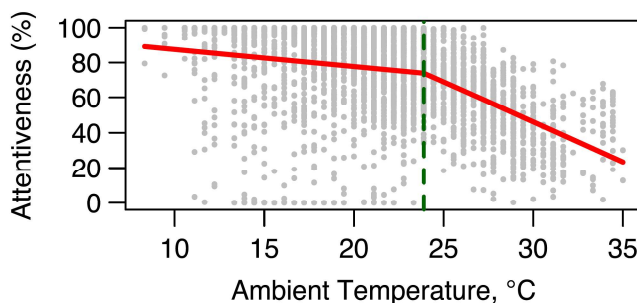


Fig. 4 The fitted relationship between nest attentiveness and temperature. Gray points show the raw data, the red line shows the fitted relationship, and the dashed green line shows the location of the change point and where the spline was fitted

constant off-bout duration. This result supports the observations of Allen and Nice (1952) who reported an increase in purple martin attentiveness during a “cold snap” of weather, but is contra to Finlay (1976) who did not find any significant effect of ambient temperature on adult entrances and departures from martin nests as quantified by a photoelectric device. Many studies have found a negative relationship between ambient temperature and incubation attentiveness in temperate regions (Conway and Martin 2000; Deeming 2002; Smith et al. 2012), and it is explained by the need to spend more effort to maintain egg temperature during colder weather. We found support for a threshold-type response, where the effect of increasing temperature becomes much more pronounced above 24 °C. This is very similar to Kendeigh’s (1952) description of house wren (*Troglodytes aedon*) incubation attentiveness declining sharply beyond 26 °C and the findings of Conway and Martin (2000) with orange crowned warblers (*Vermivora celata*). Once incubation is initiated, exposure of eggs to temperatures below physiological zero (PZT) can result in decreased viability (Viega 1992; Arnold 1993). The exact temperature at which PZT occurs is debated and most likely varies by species (Deeming 2002) but is widely estimated around 26 °C (Haftorn 1988), in broad agreement with the estimated changepoint in our data. It appears that the purple martins are not only adjusting their behavior to ambient temperature in a linear fashion but are capable of a relatively complex two-step response to temperature, where they are detecting when PZT is achieved and reducing their incubation effort. It is unclear whether they do this by direct assessment of either ambient or egg temperature or through some signaling from the embryos which may themselves be capable of sensing their own temperature (Reed and Clark 2011) and communicating with their parents via emission of volatile chemicals (Saino et al. 2015).

We detected higher attentiveness before ~9:30 a.m. than in the middle of the day and a parallel slight increase in attentiveness after ~6 p.m. This is in perfect agreement with Kendeigh’s (1952) thermocouple study of a single martin nest which found attentiveness to be highest at dawn and dusk and lower during the middle of the day. The same phenomenon was reported in uniparental incubating shorebirds (Smith et al. 2012) and Sprague’s pipit (*Anthus spragueii*) (Davis and Holmes 2012), but no trends in attentiveness were found through the day in western meadowlark (*Sturnella neglecta*) (Powell et al. 2012). It is likely that this trend is, at least partly, driven by the correlation with ambient temperature, with warmer temperatures in the middle of the day simultaneously decreasing the need for incubation and providing better hunting opportunities (purple martins are aerial insectivores) away from the nest. However, the correlation in our data between time of day and temperature was relatively weak (Online Resource 1: Fig. 1A). While we were not able to record data overnight due to insufficient light levels in the nest, it is

thought that passerine nest attentiveness approaches 100% overnight during incubation (Deeming 2002). It could be that the increased dawn and dusk attentiveness captured here is the onset of that behavior. Increased nocturnal attentiveness could be a positive evolutionary response to increased egg predation pressure overnight or could simply be a natural consequence of purple martins not hunting at night and hence no longer dividing their time between feeding and incubating.

Given that egg cooling rate increases with egg age (Cooper and Voss 2013), we might expect attentiveness to increase with day of incubation. This was indeed observed in black-capped chickadees (*Poecile atricapillus*) (Cooper and Voss 2013) and prairie nesting ducks (Loos and Rohwer 2004). However, other studies of female-only incubators have found a decrease in attentiveness over the course of incubation (Brown and Fredrickson 1987; Yerkes 1998). In common with Kendeigh’s (1952) study of house wrens (*Troglodytes aedon*), our model showed a slight increase in attentiveness over the first few days of incubation followed by a relatively constant pattern over the rest of the incubation period, although, in our data, the day of incubation was a relatively minor predictor of attentiveness. It is possible that discrepancies between studies with respect to constancy of incubation over the incubation period are driven by differences in how incubation initiation is defined between studies, or underlying correlations with ambient temperature.

Interestingly, our data showed very little variation in incubation attentiveness by intrinsic factors such as adult age class or clutch size. That there was little difference in nest attentiveness by adult age class (SY vs. ASY) was surprising given that multiple studies have indicated that first breeding seasons in purple martins are not as productive as for experienced birds. SY female martins lay smaller clutches and have lower nesting success (Davidar and Morton 1993), and SY males have been found to provide less parental care (Morton and Derrickson 1990) and provision less food to nestlings (Wagner et al. 1997) than ASY males. Beyond purple martins, inexperienced breeders have been shown to be less attentive incubators in Canada geese (*Branta canadensis*) (Aldrich and Raveling 1983) and to have decreased survival of their offspring in herring gulls (*Larus argentatus*) (Bogdanova et al. 2007), but to have no difference in attentiveness in blue tits (*Cyanistes caeruleus*) (Amininasab et al. 2016). We also expected that nest attentiveness would increase with increasing clutch size as has been observed in pied flycatchers (*Ficedula hypoleuca*) (Blagosklonov 1977) and black-capped chickadees (*Poecile atricapillus*) (Cooper and Voss 2013) and, more generally, with increasing “clutch mass” across 354 avian species (Deeming

2002). The lack of a relationship in our study may be due to the relatively small variation in clutch size in our dataset (4–6 eggs) compared with 4–12 eggs in the flycatcher study and 4–9 eggs in the chickadee study.

We compared the results of a random forest and GLMM for their ability to model nest attentiveness in the purple martin. We found that the random forest was able to explain a much higher proportion of the variation in our dataset, presumably because it allows for complex non-linear relationships in the data (Conway and Martin 2000). For example, while the random forest model allowed for a graded response to ambient temperature, with different degrees of response in different temperature ranges, the GLMM fits a single slope to the entire temperature range and thus misses the step-change the attentiveness-temperature relationship after PZT is reached. While random forest models are sometimes criticized for overfitting the data (Segal 2004), the close relationship between our training and validation data error rates indicates excellent model generalization; therefore, this is not likely the case in this instance.

Changes in nest attentiveness with temperature, time of day, and day of incubation were apparently driven by changes in on-bout duration. Mirroring overall attentiveness, on-bout duration was principally affected by temperature, with decreased bout duration with ambient temperature above ~ 20 °C. This likely reflects a re-allocation of time away from prolonged incubation which is likely not necessary at in-nest temperatures above PZT in favor of other activities such as feeding. There is a noisy relationship with temperature and on-bout duration below ~ 20 °C. This may be due to a slight negative relationship being masked by an increase in bout duration with day of incubation which itself co-varies with temperature.

Off-bout duration showed very little variation in our dataset overall and did not clearly co-vary with any of the explanatory variables tested. If purple martins are principally varying on-bout duration in response to incubation needs, it is possible that a relatively consistent off-bout duration is driven by the consistent need to guard eggs from predators (Komdeur and Kats 1999) or from intraspecific and interspecific nest usurpers. So, as all of our nests were in the same location and in similar colony sizes across years, it is possible that the stable off-bout duration is due to stable nest-loss risks.

In this study, we have used a novel method of automating classification of nest camera videos using deep learning to provide data for a detailed analysis of incubation behavior in the purple martin. This has given us the ability to analyze continuous data for the whole incubation period, rather than having to use just a subset due to the laborious nature of manually classifying video footage. The results indicate that purple martins are adjusting their incubation behavior in a

complex manner according to ambient temperature conditions. We have demonstrated the value of using deep learning analysis of nest camera video recordings for behavioral science. The same techniques could be used to estimate different metrics of incubation such as “total incubation” time (with the inclusion of overnight video footage) which might relate to more sensitive evolutionary metrics such as egg volume. The same approach could also be used to create high-resolution data of other in-nest behaviors such as provisioning or grooming.

Acknowledgments The authors are grateful to Carl Zenger, Celeste Morien, Lee Brun, Craig Brun, Joseph Toth, Logan Fahrenkopf, Samantha Wilcox, and Ramya Sridhar for field and nest camera assistance; to the staff at Iroquois NWR, especially Tom Roster and Paul Hess; and to Shawn Matott and Yashas Shamaraju for their assistance with deep learning models. We are grateful to two anonymous reviewers for constructive comments which improved our final manuscript.

Author contributions HMW conceived the idea, developed the methods, and conducted the experimental measurements. Both authors conducted the analysis. HMW wrote the paper. Both authors gave final approval for publication.

Funding information This study was made possible through funding from the Mark Diamond Research Fund, the Purple Martin Conservation Association, the North American and New York State Bluebird Societies, the Friends of Iroquois National Wildlife Refuge and NSF grant number 1556577.

Data availability The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Compliance with ethical standards

Ethical approval This study adhered to the ethical procedures in Guidelines to the Use of Wild Birds in Research. The birds were not subjected to additional disturbance as a result of this study.

Conflict of interest The authors declare that they have no conflict of interest.

References

- Aldrich TW, Raveling DG (1983) Effects of experience and body-weight on incubation behavior of Canada geese. *Auk* 100:670–679
- Allen RW, Nice MM (1952) A study of the breeding biology of the purple martin (*Progne subis*). *Am Midl Nat* 47:606–665
- Amininasab SM, Kingma SA, Birker M, Hildenbrandt H, Komdeur J (2016) The effect of ambient temperature, habitat quality and individual age on incubation behaviour and incubation feeding in a socially monogamous songbird. *Behav Ecol Sociobiol* 70:1591–1600
- Ardia DR, Pérez JH, Clotfelter ED (2010) Experimental cooling during incubation leads to reduced innate immunity and body condition in nestling tree swallows. *Proc R Soc Lond B* 277:1881–1888
- Arnold TW (1993) Factors affecting egg viability and incubation time in prairie dabbling ducks. *Can J Zool* 71:1146–1152

- Bartoń K (2018) MuMIn: multi-model inference. R package version 1.42.1, <https://cran.r-project.org/package=MuMIn>
- Bates D, Maechler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67:1–48
- Blackmer AL, Ackerman JT, Nevitt GA (2004) Effects of investigator disturbance on hatching success and nest-site fidelity in a long-lived seabird, Leach's storm petrel. *Biol Conserv* 116:141–148
- Blagosklonov KN (1977) Experimental-analysis of rhythm of incubation in pied flycatcher (*Ficedula hypoleuca*). *Sov J Ecol* 8:340–344
- Bogdanova MI, Nager RG, Monaghan P (2007) Age of the incubating parents affects nestling survival: an experimental study of the herring gull *Larus argentatus*. *J Avian Biol* 38:83–93
- Bottitta GE, Gilchrist HG, Kift A, Meredith MG (2002) A pressure-sensitive wireless device for continuously monitoring avian nest attendance. *Wildlife Soc B* 30:1033–1038
- Breheny P, Burchett W (2017) Visualization of regression models using visreg. *R J* 9:56–71
- Brown PW, Fredrickson LH (1987) Time budget and incubation behavior of breeding white winged scoters. *Wilson Bull* 99:50–55
- Conway CJ, Martin TE (2000) Effects of ambient temperature on avian incubation behavior. *Behav Ecol* 11:178–188
- Cook MI, Beissinger SR, Toranzos GA, Arendt WJ (2005) Incubation reduces microbial growth on eggshells and the opportunity for trans-shell infection. *Ecol Lett* 8:532–537
- Cooper CB, Voss MA (2013) Avian incubation patterns reflect temporal changes in developing clutches. *PLoS One* 8:e65521
- Cox WA, Pruett MS, Benson TJ, Chiavacci SJ, Thompson FR III (2012) Development of camera technology for monitoring nests. In: Ribic CA, Thompson FR III, Pietz PJ (eds) Video surveillance of nesting birds. University of California Press, Berkeley, pp 185–209
- Dallmann JD, Anderson LC, Raynor EJ, Powell LA, Schacht WH (2016) iButton® temperature loggers effectively determine prairie grouse nest absences. *Great Plains Res* 26:117–123
- Davidar P, Morton ES (1993) Living with parasites - prevalence of a blood parasite and its effect on survivorship in the purple martin. *Auk* 110:109–116
- Davis SK, Holmes TG (2012) Sprague's pipit incubation behavior. In: Ribic CA, Thompson FR III, Pietz PJ (eds) Video surveillance of nesting birds. University of California Press, Berkeley, pp 67–76
- Deeming DC (2002) Avian incubation. Behavior, environment and evolution. Oxford University Press, Oxford
- Deeming DC, Reynolds SJ (2015) Nests, eggs and incubation. Oxford University Press, Oxford
- Finlay JC (1976) Some effects of weather on purple martin activity. *Auk* 93:231–244
- Fong H, Permar G (2017) chngpt: threshold regression model estimation and inference. *BMC Bioinformatics* 18:454
- Haftorn S (1988) Incubating female passerines do not let the egg temperature fall below the 'physiological zero temperature' during their absences from the nest. *Ornis Scand* 19:97–110
- Hepp GR, Kenamer RA, Johnson MH (2006) Maternal effects in wood ducks: incubation temperature influences incubation period and neonate phenotype. *Funct Ecol* 20:308–314
- Hinkley DV (1970) Inference about the change-point in a sequence of random variables. *Biometrika* 57:1–17
- Ibanez-Alamo JD, Sanllorente O, Soler M (2012) The impact of researcher disturbance on nest predation rates: a meta-analysis. *Ibis* 154:5–14
- Kendeigh SC (1952) Parental care and its evolution in birds, vol XXII. University of Illinois Press, Urbana
- Killick R, Fearnhead P, Eckley IA (2012) Optimal detection of changepoints with a linear computational cost. *J Am Stat Assoc* 107:1590–1598
- Killick R, Haynes K, Eckley IA (2016) changepoint: an R package for changepoint analysis. R package version 2.2.2, <https://cran.r-project.org/package=changepoint>
- Komdeur J, Kats RKH (1999) Predation risk affects trade-off between nest guarding and foraging in Seychelles warblers. *Behav Ecol* 10: 648–658
- Liaw A, Wiener M (2002) Classification and regression by randomForest. *R News* 2:18–22
- Loos ER, Rohwer FC (2004) Laying-stage nest attendance and onset of incubation in prairie nesting ducks. *Auk* 121:587–599
- Martin TE (2004) Avian life-history evolution has an eminent past: does it have a bright future? *Auk* 121:289–301
- Martin TE, Schwabl H (2008) Variation in maternal effects and embryonic development rates among passerine species. *Phil Trans R Soc B* 363:1663–1674
- Martin TE, Scott J, Menge C (2000) Nest predation increases with parental activity: separating nest site and parental activity effects. *Proc R Soc Lond B* 267:2287–2293
- Morton ES, Derrickson KC (1990) The biological significance of age-specific return schedules in breeding purple martins. *Condor* 92: 1040–1050
- Powell LA, Giovanni MD, Groepper S, Reineke ML, Schacht WH (2012) Attendance patterns and survival of western meadowlark nests. In: Ribic CA, Thompson FR III, Pietz PJ (eds) Video surveillance of nesting birds. University of California Press, Berkeley, pp 61–66
- R Development Core Team (2015) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna <http://www.r-project.org/>
- Reed WL, Clark ME (2011) Beyond maternal effects in birds: responses of the embryo to the environment. *Integr Comp Biol* 51:73–80
- Reid JM, Monaghan P, Ruxton GD (1999) The effect of clutch cooling rate on starling, *Sturnus vulgaris*, incubation strategy. *Anim Behav* 58:1161–1167
- RStudio Team (2015) RStudio: integrated development for R. RStudio, Inc., Boston <http://rstudio.com>
- Saino N, Webster B, Hayes W, Pike TW (2015) Avian egg odour encodes information on embryo sex, fertility and development. *PLoS One* 10:e0116345
- Segal MR (2004) Machine learning benchmarks and random forest regression. In: eScholarship repository. University of California, <https://escholarship.org/uc/item/35x3v9t4>. Accessed 11/11/19
- Serrano SA, Benítez-Jimenez R, Nuñez-Rosas L, del Coro AM, Greeney H, Reyes-Meza V, Morales E, Escalante HJ (2018) Automated detection of hummingbirds in images: a deep learning approach. In: Martínez-Trinidad J, Carrasco-Ochoa J, Olvera-Lopez J, Sarkar S (eds) Pattern recognition proceedings of the 10th Mexican conference, Puebla, Mexico. Springer Nature, Basel, pp 155–166
- Skutch AF (1962) The constancy of incubation. *Wilson Bull* 74:115–152
- Smith PA, Dauncey SA, Gilchrist HG, Forbes MR (2012) The influence of weather on shorebird incubation. In: Ribic CA, Thompson FR III, Pietz PJ (eds) Video surveillance of nesting birds. University of California Press, Berkeley, pp 89–103
- Smith JA, Cooper CB, Reynolds SJ (2015) Advances in techniques to study incubation. In: Deeming DC, Reynolds SJ (eds) Nests, eggs and incubation. Oxford University Press, Oxford, pp 179–195
- Tabak MA, Norouzzadeh MS, Wolfson DW et al (2018) Machine learning to classify animal species in camera trap images: applications in ecology. *Methods Ecol Evo* 10:585–590
- Tarof S, Brown C (2013) Purple martin (*Progne subis*). In: Poole A. The Birds of North America. Online. In: Poole A (ed) The birds of North America. Cornell Laboratory of Ornithology, Ithaca. <https://doi.org/10.2173/bna.287>
- Viega JP (1992) Hatching asynchrony in the house sparrow: a test of the egg-viability hypothesis. *Am Nat* 139:669–675
- Wagner RH, Davidar P, Schug MD, Morton ES (1997) Do blood parasites affect paternity, provisioning and mate-guarding in purple martin? *Condor* 99:520–523
- Willi M, Pitman RT, Cardoso AW, Locke C, Swanson A, Boyer A, Veldthuis M, Fortson L, Gaggiotti O (2018) Identifying animal

- species in camera trap images using deep learning and citizen science. *Methods Ecol Evol* 10:80–91
- Williams JB (1996) Energetics of avian incubation. In: Carey C (ed) *Avian energetics and nutritional ecology*. Springer, Boston, pp 375–415
- Williams HM, Matott LS, DeLeon RL (2019) Automated deep learning analysis of purple martin videos depicting incubation and provisioning. In: *Proceedings of Practice and Experience in Advanced Research Computing (PEARC '19)*, Chicago, IL, p 64
- Yerkes T (1998) The influence of female age, body mass, and ambient conditions on redhead incubation constancy. *Condor* 100:62–68
- Zangmeister JL, Hausmann MF, Cerchiara J, Mauck RA (2009) Incubation failure and nest abandonment by Leach's storm-petrels detected using PIT tags and temperature loggers. *J Field Ornithol* 80: 373–379

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.