# Written Report for “Analysis of bird survey data to refine monitoring designs and survey protocols” (EC Contract No. 3000704376)

By Lionel Leston

# Introduction

In recent decades, many landbird species have experienced continued long-term population declines across Canada and North America (Rosenberg et al. 2019), based on data from broad-scale, all-bird monitoring programs like the North American Breeding Bird Survey (hereafter, “BBS”) (Sauer et al. 1994). As a result, several species are listed on Schedule 1 of Canada’s Species-at-Risk Act (SARA). Aerial insectivores like the Common Nighthawk (*Chordeiles minor*) (hereafter, “CONI”) are one of the most strongly declining groups (Canaday 1997, Sekercioglu et al. 2002, Nebel et al. 2010, Hallman et al. 2014, Paquette et al. 2014, Smith et al. 2015). For several species including crepuscular aerial insectivores like the CONI, their habitat preferences or activity periods are likely mismatched with current protocols such as those of the BBS. The methods of the BBS are designed to maximize the detection of diurnal songbird species, many of which can be detected within the 3-minute count periods used by the BBS. In contrast, the crepuscular CONI is likely to be more active and detectable at times when songbirds are less active, so its occupancy and abundance may be poorly represented within existing all-bird surveys like the BBS. Inaccurate estimates of occupancy and abundance could result in unreliable population distribution, trend estimates, and recommendations for this species’ recovery. The Canadian Wildlife Service (Ontario Region) (CWS-ON) has conducted several pilot surveys and research projects in recent years with the goal of evaluating and improving survey protocols for species with low data precision.

# Objectives

* Run occupancy models with detection covariates to assess which variables most strongly affect probability of detection of CONI across Canada.
* Determine optimal recording length and number of recordings to reliably detect CONI that are present.
* Assess accuracy of model predictions of CONI presence/absence (e.g. receiver-operating characteristic [ROC] curves, rates of false positive detections).
* Run generalized additive mixed models to determine how CONI call rates (booms and peent calls) vary with ordinal day and time since sunrise.
* After determining optimal survey design for CONI, compare and contrast this survey design to BBS protocols to show how protocols might be adjusted to increase CONI detection.

# Methods

## Recordings

Nocturnal bird surveys focusing on CONI were conducted from June 1 to August 28, 2014 at 23 sites (1-3 sites at each of 12 general locations spanning 3000 km from Ottawa, ON (45°21 N, -76°0 W) to Yellowknife, NT (62°42 N, -116°6 W)) using autonomous recording units (ARUs). Sites within general locations were ≥ 1.8 km apart to minimize the probability of double-counting individual CONI at different sites. Sites were selected opportunistically, coinciding with other projects conducted by colleagues and volunteers. Sites were also selected based on their *a priori* suitability as CONI habitats (e.g. extensive openings with sand, gravel or rock; recently burned or harvested areas (Brigham et al. 2011)), as determined from mapped imagery or by biologists in the field deploying ARUs in the areas.

CONI sounds include low-frequency (0.4-1.0 kHz) “booms” made by flexing the wings at the bottom of an aerial dive. These booms are made presumably by males at dawn and dusk by males to maintain and defend discreet aerial territories (Weller 1958, Brigham et al. 2011), often over habitat features used for feeding or breeding (Caccamise 1974). Another more common vocalization is a mid-frequency contact call or "peent" (2-4 kHz), used for territorial displays, defensive threats, and courtship (Brigham et al. 2011). Patterns in both sound types were analyzed separately from each other, since booms and peents appear to be associated with different behaviours and functions and since detection of one sound type might be differently affected by environmental variables from the other sound type.

Nighthawk activity was measured using Song Meter ARUs (Model SM2+, Wildlife Acoustics Inc., Maynard, MA; firmware version 3.2.5) ARUs. ARUs were deployed in late May or early-June at all locations and retrieved from mid-July to mid-August, and were programmed to record continuously from 1-hr before sunset to 1-hr after sunrise local time at each site, starting 1 June 2014; the same recording program was repeated on a 4-day interval throughout the breeding season or until ARUs were retrieved. Nighthawk sounds are typically low frequency and given the large number of recordings required, ARUs were programmed to record at a sampling rate of 16 kHz and a bit depth of 16 bits to conserve battery power. Files were saved in the uncompressed waveform audio file (.wav) format. A single 32 GB and a single 16 GB SD memory card (48 GB total) were used together in each ARU to store all acoustic files within ARUs, with standard alkaline batteries installed in units prior to deployment.

Two partially-automated approaches were used to obtain data from acoustic recordings. First, peent vocalizations were obtained using an automated time-frequency band-limited energy detector (BLED; see Mills 2000), in Raven Pro (version 1.4; Charif et al. 2010). Frequency range and time durations used to parameterize the detector were defined based on a random sampling of 100 peents from our recordings. For the remaining detector parameters, settings were adjusted as necessary from default settings to find the best configuration for isolating "peent" vocalizations. After running the detector, the resulting list of candidate peent detections was verified manually for accuracy. False positives were scored as 0 and true positives were scored as 1 in the resulting output file.

Second, a visual scanning approach was used to count non-vocal "booms" by viewing spectrograms in Raven Pro. Using the timed auto page-advance function, 1-min recordings were viewed and rapidly assessed, enabling processing of an hour’s recordings in approximately 1 min, depending on the frequency of nighthawk booming. The analyst initially audibly confirmed each visually detected candidate "boom" until candidate signals could be confidently confirmed using only visual detection. When a "boom" was confirmed on a recording, the timed auto-page advance function was paused and the signal was selected by drawing a box around it, capturing the start time (in seconds since the beginning of the recording), duration, and frequency range.

## Converting Continuous Recording Data to Interval Data

Continuous recorded data for each site included the date and start and end times of each recording, and the time (seconds after the start of the recording) when either peents or booms were detected. Each recording began 1 hour before local sunset and ended 1 hour after local sunrise. The **suncalc** package in R (Agafonkin and Thieurmel 2017) was used to estimate the start and end of civil, nautical, and astronomical twilight periods 1 and 2 as well as the “night” twilight period between astronomical twilight 1 and astronomical twilight 2.

I used the **lubridate** package in R (Grolemund and Wickham 2011) to break down each night’s single recording at each site into intervals of specific duration. I used durations ranging from 1 minute to 20 minutes in 1-minute increments to create a range of interval times including those periods typically used in bird surveys across studies, along with 1-hour intervals. Each time I broke a continuous recording down into specific intervals, I saved the results in a separate CSV file (“*0\_data/ processed/1\_IntervalUsed*”).

*Assigning Specific Detection Events and Other Data to Specific Time Intervals*

I assigned peents and booms to specific time intervals separately. After using the **lubridate** package to calculate the date-time of a specific detection event based on the event’s time in a recording, I used the **intrval** package in R (Sólymos 2017) to map individual detection events from a given site and date to a specific time interval within that site’s and date’s recording. Specific time intervals that lacked any detection events were assigned “NA” values. I looped through the 21 specific interval files, mapping detection events to 21 mapped interval files for each kind of CONI vocalization (“*0\_data/processed/2\_BoomDetectionsMapped*”, “*0\_data/processed/2\_PeentDetectionsMapped*”).

Once specific intervals were assigned detections or no detections, I used the **lubridate** package to determine time since sunset (hereafter, “TSSS”) for the start time of each interval. I used the **suncalc** package to determine the moon fraction (ranging from “new moon” = 0 to “full moon” = 1) as a measure of moon illumination that might influence CONI activity through the amount of moonlight when foraging. I assigned specific intervals to “Twilight Period” based on the intervals’ times within the recording relative to the twilight period times at a given site on a given day, using the categories “Before”, “Civil”, “Nautical”, “Astronomical”, “Night”, and “After”. Finally, I assigned mean nightly temperature to intervals for those dates when temperatures had been taken at each site (“*0\_data/processed/3\_BoomsMapped\_SunAndMoon*”, “*0\_data/processed/3\_PeentsMapped\_SunAndMoon*”).

## Sampling from Intervals

To create data sets for testing the effects of sample size and sample duration on CONI detection probability, I randomly drew all observations from different numbers of intervals at each site across the whole season, from each of the interval files with mapped detections, TSSS, twilight period and temperature. For a given interval duration, I randomly sampled 1, 2, and 3 intervals from each site from any time in the season. For larger sample sizes (4, 8, 12, 16, 20, 24, 28, and 32 intervals), I stratified the intervals files by creating half-month-long sampling periods, then randomly sampled 1, 2, 3, 4, 5, 6, 7, and 8 samples per half-month period between June 1 and July 31. I excluded all recordings from August because CONI activity was determined to be negligible by that time. I also excluded all sites missing nightly mean temperature data. Finally, I excluded sampling from any intervals that were less than the specified interval duration within a given file. Such intervals occurred because each continuous nightly recording could not be perfectly divided into intervals of the same duration. Each sample was stored in a separate file (“*0\_data/processed/4\_NSampleBooms*”, “*0\_data/processed/4\_NSamplePeents*”).

## Assessing Activity Rates with Generalized Additive Mixed Models

To examine how activity rates (number of peents or booms counted per interval) varied with ordinal day and TSSS for sites occurring at different latitudes, I used the **gamm4** package in R (Wood et al. 2017) to run a generalized additive mixed model (GAMM) to select the likeliest nonlinear functions of ordinal day and TSSS that fit the data. Based on exploration of raw count data, I split sites into two groups by latitude (LatgroupN: “North” ≥ 55 degrees, “South” < 55 degrees) since at that latitude the length of both nightly recordings and the survey season was significantly shorter at the northern sites.

Each GAMM was modelled using a negative binomial error distribution in which I specified the amount of overdispersion relative to a Poisson distribution (theta). The GAMM took the following form:

Count ~ LatgroupN + s(TSSS, k=4, bs=”cs”, by=LatgroupN) + s(Ordinal day, k=4, bs=”cs”, by=LatgroupN) + s(Mean nightly temperature, k=4, bs=”cs”, by=LatgroupN)

I used “site” as a random effect to account for correlations due to repeated visits from the same sites. I specified a nonlinear function with up to 4 knots or change points, to be connected by cubic splines.

Using the 10-minute interval data for both peent and boom detections, I first randomly withdrew without replacement detection and covariate data from 20 intervals per site as test data for validating the GAMMs. The remainder of the intervals were potential training data for generating GAMMs. Within 100 bootstrap iterations, I randomly drew 20 intervals with replacement from each site, ran the aforementioned GAMM, then stored the model coefficients from the GAMM, both for generating prediction plots with the training data and for validating the GAMM with the test data. From the 100 bootstrapped model coefficients for each term in the GAMM, I generated a bar plot (median + 5th and 95th quantile values) for each term’s coefficient values and matrix plots showing predicted numbers of peent and boom detections per 10-minute interval as a function of TSSS, ordinal day, and mean nightly temperature. For each bootstrapped set of model terms, I predicted numbers of peent and boom detections per 10-minute interval and compared predicted to actual numbers of detections using Spearman correlation coefficients. I generated median + 5th and 95th quantile values of the Spearman correlation coefficients for both the peent and boom GAMMs.

## Assessing Probability of Detection With Occupancy Models

Once I determined how peent and boom activity rate varied with TSSS, ordinal day, and mean nightly temperature, I used the results to select a set of TSSS and ordinal day intervals where activity rates were highest. I then randomly sampled a different number of intervals (“visits”) from interval data of different lengths in order to run occupancy models. For each sampled data set combination of sample number per site and sample duration, I ran single-season occupancy models with the **unmarked** package in R (Fiske and Chandler 2011) to measure how predicted probability of occupancy and probability of detecting CONI given occupancy varied with interval duration and number of visits. I ran separate sets of occupancy model analyses using 1) peent detections as the response variable, then 2) boom detections as the response variable.

Seventeen of 23 sites had peent detections while19 sites had boom detections in 2014. As assessing probability of detection given true occupancy was of interest to the study, I only used sites within an occupancy modelling analysis if peents or booms were detected at that site within 2014. In doing so, I removed the possibility that low detection probability for a given number of visits and interval duration was affected by true absence at a site. Modelling only sites where actual detections were known also gave me a way of validating predicted probability of occupancy. I ran a single “null” occupancy model in which occupancy and detection were assumed to be equal among all sites, after filtering the data to a narrow window of TSSS and ordinal day values where peent and boom activity rates (and presumably detection probability) were highest).

For each combination of interval duration (2, 4, 6, 8,10, 12, 14, 16, 18, 20 minutes) and number of visits sampled per site (3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 20), I generated 100 bootstrap samples in both the peent and the boom occupancy model analyses. I ran the null model and stored the predicted occupancy and detection probabilities from each bootstrap along with the interval duration and number of visits, then graphed occupancy and detection probabilities versus interval duration and number of visits.

## Assessing Activity Rates with Mixed-effects Models

To examine how activity rates (number of peents or booms counted per interval) varied with twilight period at different latitudes, I used the **glmer.nb** function in the lme4 package in R (Bates et al. 2015) to run two mixed-effects models with a negative binomial error distribution. The mixed-effects models took the following form:

Count ~ Twilight period (effect of twilight period does not vary with latitude) (***Non-interaction model***)

Count ~ Twilight period + Latitude + Twilight period\*Latitude (effect of twilight period varies with latitude) (***Interaction model***)

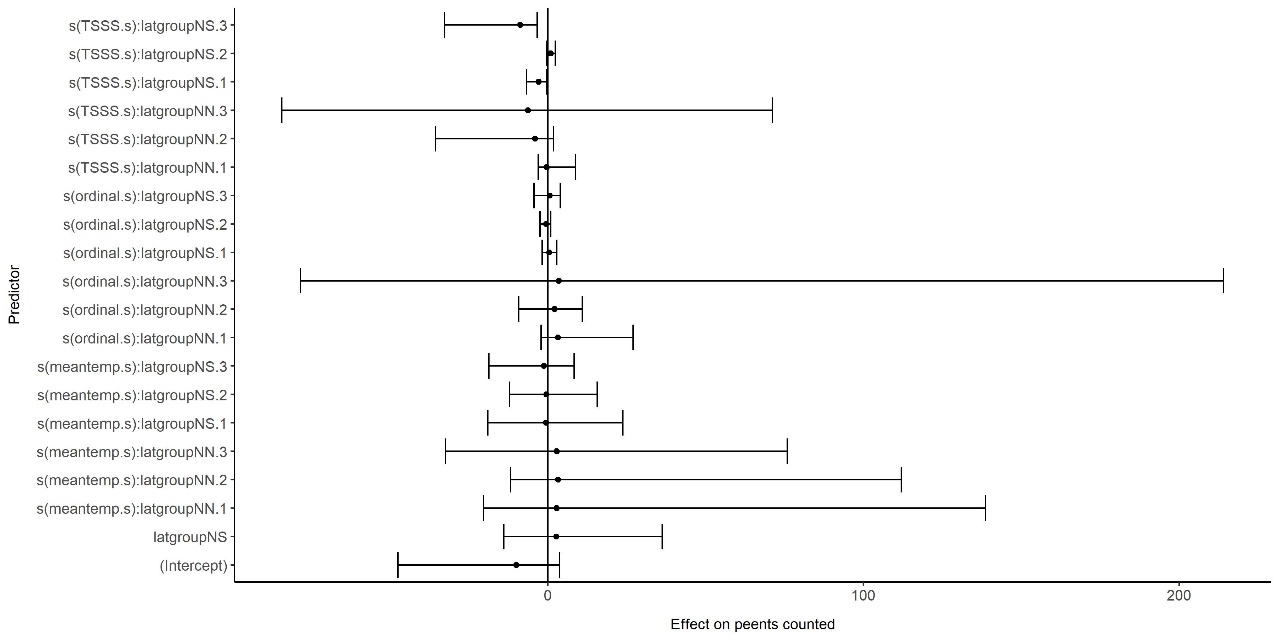
I used “site” as a random effect to account for correlations due to repeated visits from the same sites.

Using the 10-minute interval data for both peent and boom detections, I randomly drew equal numbers of observations per site and six twilight periods (“Before”, “Civil”, “Nautical”, “Astronomical”, “Night”, and “After). I excluded observations that occurred outside the period of highest activity rates according to the GAMMs (170 < ordinal day < 190). I then dropped the northernmost six of the 23 sites from analysis because they lacked certain twilight periods (“Night”, “Astronomical”) between ordinal days 170-190. I set aside 10 observations per site and twilight period as test data for validating mixed-effects models and generating predictions of activity rates. The remaining data served as model training data, from which 100 sample data sets were drawn (a maximum possible 12 observations per site and twilight period) via bootstrapping. I reclassified twilight period to distinguish between civil, nautical, and astronomical twilight periods that occurred before midnight (“Civil.1”, “Nautical.1”, “Astronomical.1”) and after midnight (“Civil.2”, “Nautical.2”, “Astronomical.2”). Including interaction terms with latitude, reclassification of twilight period resulted in a model with 18 fixed effect predictors. Within 100 bootstrap iterations, I randomly drew 12 intervals with replacement from each site, ran the aforementioned model, then stored the model coefficients to generate prediction plots and validate the model with the test data. From the 100 bootstrapped model coefficients for each term in the model, I generated a bar plot (median + 5th and 95th quantile values) for each term’s coefficient values and matrix plots showing predicted numbers of peent and boom detections per 10-minute interval as a function of twilight period, latitude, and the twilight period\*latitude interaction terms. For each bootstrapped set of model terms, I predicted numbers of detections per 10-minute interval and compared predicted to actual numbers of detections using Spearman correlation coefficients. I generated median + 5th and 95th quantile values of the Spearman correlation coefficients.

# Results

## Assessing Activity Rates with Generalized Additive Mixed Models

Bootstrapped coefficient terms from the GAMM for peent activity suggested that TSSS had a stronger impact than ordinal day or temperature (Figure 1). At southern sites, there were two peak times of peent activity around sunset and sunrise, while at northern sites there was one peak time of peent activity from 1 hour before to 2 hours after sunset (Figure 2). Peent activity rates were higher around ordinal day 170 at southern sites and ordinal day 180 at northern sites (Figure 3). Peent activity rate also increased slightly with increasing temperature at northern but not southern sites (Figure 4).



### Figure 1. Box plots showing the median, 5th and 95th quantile values for the model terms based on 100 bootstrapped sample data sets used in the generalized additive mixed model: Peent activity rate = latgroupN (“N”≥55, “S”<55) + (TSSS, knots = 4, basis = "cubic spline", by=latgroupN) + (Ordinal day, knots = 4, basis = "cubic spline", by=latgroupN) + (Mean nightly temperature, knots = 4, basis = "cubic spline", by=latgroupN).



### Figure 2. Matrix plots showing the predicted CONI peent activity rates (counts per 10-minute interval) at the southernmost site (“8246”) and northernmost site (“13588”) based on 100 bootstrapped sample data sets used in the generalized additive mixed model: latgroupN (“N”≥55, “S”<55) + (TSSS, knots = 4, basis = "cubic spline", by=latgroupN) + (Ordinal day, knots = 4, basis = "cubic spline", by=latgroupN) + (Mean nightly temperature, knots = 4, basis = "cubic spline", by=latgroupN). Left-side matrix plots show individual predicted peent activity rates from each bootstrap sample versus TSSS at the two sites. Right-side matrix plots show the median, 5th and 95th quantile values from the bootstrapped predicted peent activity rates for each value of TSSS.

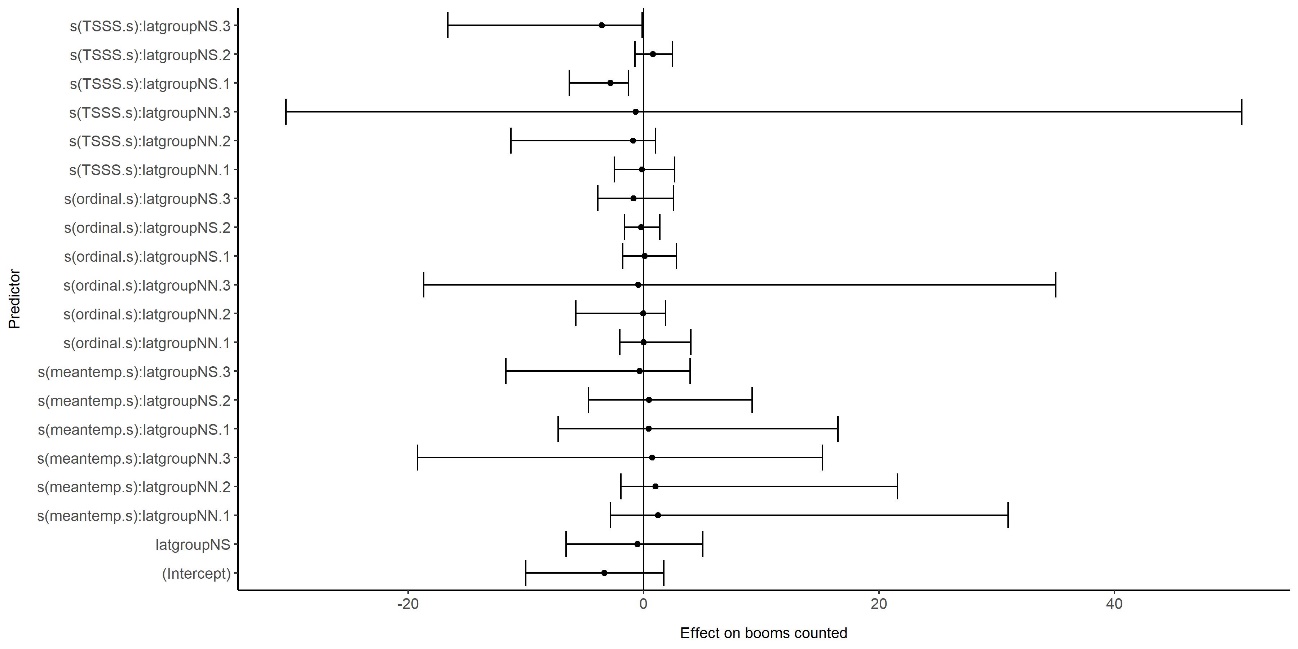


### Figure 3. Matrix plots showing the predicted CONI peent activity rates (counts per 10-minute interval) at the southernmost site (“8246”) and northernmost site (“13588”) based on 100 bootstrapped sample data sets used in the generalized additive mixed model: latgroupN (“N”≥55, “S”<55) + (TSSS, knots = 4, basis = "cubic spline", by=latgroupN) + (Ordinal day, knots = 4, basis = "cubic spline", by=latgroupN) + (Mean nightly temperature, knots = 4, basis = "cubic spline", by=latgroupN). Left-side matrix plots show individual predicted peent activity rates from each bootstrap sample versus ordinal day at the two sites. Right-side matrix plots show the median, 5th and 95th quantile values from the bootstrapped predicted peent activity rates for each value of ordinal day.



### Figure 4. Matrix plots showing the predicted CONI peent activity rates (counts per 10-minute interval) at the southernmost site (“8246”) and northernmost site (“13588”) based on 100 bootstrapped sample data sets used in the generalized additive mixed model: latgroupN (“N”≥55, “S”<55) + (TSSS, knots = 4, basis = "cubic spline", by=latgroupN) + (Ordinal day, knots = 4, basis = "cubic spline", by=latgroupN) + (Mean nightly temperature, knots = 4, basis = "cubic spline", by=latgroupN). Left-side matrix plots show individual predicted peent activity rates from each bootstrap sample versus temperature at the two sites. Right-side matrix plots show the median, 5th and 95th quantile values from the bootstrapped predicted peent activity rates for each value of temperature.

Bootstrapped coefficient terms from the GAMM for peent activity suggested that TSSS had a stronger impact than ordinal day or temperature (Figure 5). At southern sites, there were two peak times of peent activity around sunset and sunrise, while at northern sites there was one peak time of peent activity from 1 hour before to 2 hours after sunset (Figure 6). Peent activity rates were higher around ordinal day 165 at southern sites and ordinal day 185 at northern sites (Figure 7). Boom activity rate did not vary with increasing temperature at northern or southern sites (Figure 8).



### Figure 5. Box plots showing the median, 5th and 95th quantile values for the model terms based on 100 bootstrapped sample data sets used in the generalized additive mixed model: Boom activity rate = latgroupN (“N”≥55, “S”<55) + (TSSS, knots = 4, basis = "cubic spline", by=latgroupN) + (Ordinal day, knots = 4, basis = "cubic spline", by=latgroupN) + (Mean nightly temperature, knots = 4, basis = "cubic spline", by=latgroupN).



### Figure 6. Matrix plots showing the predicted CONI boom activity rates (counts per 10-minute interval) at the southernmost site (“8246”) and northernmost site (“13588”) based on 100 bootstrapped sample data sets used in the generalized additive mixed model: latgroupN (“N”≥55, “S”<55) + (TSSS, knots = 4, basis = "cubic spline", by=latgroupN) + (Ordinal day, knots = 4, basis = "cubic spline", by=latgroupN) + (Mean nightly temperature, knots = 4, basis = "cubic spline", by=latgroupN). Left-side matrix plots show individual predicted boom activity rates from each bootstrap sample versus TSSS at the two sites. Right-side matrix plots show the median, 5th and 95th quantile values from the bootstrapped predicted boom activity rates for each value of TSSS.



### Figure 7. Matrix plots showing the predicted CONI boom activity rates (counts per 10-minute interval) at the southernmost site (“8246”) and northernmost site (“13588”) based on 100 bootstrapped sample data sets used in the generalized additive mixed model: latgroupN (“N”≥55, “S”<55) + (TSSS, knots = 4, basis = "cubic spline", by=latgroupN) + (Ordinal day, knots = 4, basis = "cubic spline", by=latgroupN) + (Mean nightly temperature, knots = 4, basis = "cubic spline", by=latgroupN). Left-side matrix plots show individual predicted boom activity rates from each bootstrap sample versus ordinal day at the two sites. Right-side matrix plots show the median, 5th and 95th quantile values from the bootstrapped predicted boom activity rates for each value of ordinal day.



### Figure 8. Matrix plots showing the predicted CONI boom activity rates (counts per 10-minute interval) at the southernmost site (“8246”) and northernmost site (“13588”) based on 100 bootstrapped sample data sets used in the generalized additive mixed model: latgroupN (“N”≥55, “S”<55) + (TSSS, knots = 4, basis = "cubic spline", by=latgroupN) + (Ordinal day, knots = 4, basis = "cubic spline", by=latgroupN) + (Mean nightly temperature, knots = 4, basis = "cubic spline", by=latgroupN). Left-side matrix plots show individual predicted boom activity rates from each bootstrap sample versus temperature at the two sites. Right-side matrix plots show the median, 5th and 95th quantile values from the bootstrapped predicted boom activity rates for each value of temperature.

## Effect of Number of Visits and Visit Duration on Occupancy and Detection Probability

I drew occupancy model samples for peent detections from TSSS between -3600 and 3600 (1 hour before to 1 hour after sunset) and for ordinal days 175 to 185. Probability of predicting true occupancy based on peent detections increased with visit duration and with number of visits. To be likely to predict occupancy (ψ≥0.50), at least 6 visits were required if intervals were 20 minutes long. If intervals were shorter, more visits were necessary (Figure).

Given that a site is occupied, the probability of detecting peents doesn’t become likely (P≥0.50) until visits are at least 11 minutes long. There was a negative but nonsignificant relationship between detection probability and number of visits (Figure). The nonsignificance of this relationship may have been due to the greater uncertainty in estimated detection probability estimated from fewer visits, since uncertainty in detection probability estimates decreased as the number of visits increases (Figure). When intervals were 12 or more minutes long, detection probability was >0.50 with 3-4 visits.



### Figure 9. Probability of occupancy (represented as smoothed logistic curves) based on peent detections in the samples versus number of visits per site (3-20) and duration of visit (in minutes).



### Figure 10. Probability of occupancy (represented as prediction intervals) based on peent detections in the samples versus number of visits per site (3-20) and duration of visit (in minutes).



### Figure 11. Probability of detection given occupancy (represented as smoothed logistic curves) based on peent detections in the samples versus number of visits per site (3-20) and duration of visit (in minutes).



### Figure 12. Probability of detection given occupancy (represented as prediction intervals) based on peent detections in the samples versus number of visits per site (3-20) and duration of visit (in minutes).

I drew occupancy model samples for boom detections from TSSS between -3600 and 3600 (1 hour before to 1 hour after sunset) and for ordinal days 165 to 185. Probability of predicting true occupancy based on boom detections increased with visit duration and with number of visits. Predicted probability occupancy was high (ψ≥0.50) even with just 3 visits and 2-minute intervals (Figure).

Given that a site is occupied, the probability of detecting booms doesn’t become likely (P≥0.50) until visits are at least 16 minutes long. There was a negative but nonsignificant relationship between detection probability and number of visits (Figure). The nonsignificance of this relationship may have been due to the greater uncertainty in estimated detection probability estimated from fewer visits, since uncertainty in detection probability estimates decreased as the number of visits increases (Figure). In contrast, there was a significant positive relationship between detection probability and visit duration with visits ≥16 minutes long having greater detection probability than visits<10 minutes long (Figure).



### Figure 13. Probability of occupancy (represented as smoothed logistic curves) based on boom detections in the samples versus number of visits per site (3-20) and duration of visit (in minutes).



### Figure 14. Probability of occupancy (represented as prediction intervals) based on boom detections in the samples versus number of visits per site (3-20) and duration of visit (in minutes).



### Figure 15. Probability of detection given occupancy (represented as smoothed logistic curves) based on boom detections in the samples versus number of visits per site (3-20) and duration of visit (in minutes).

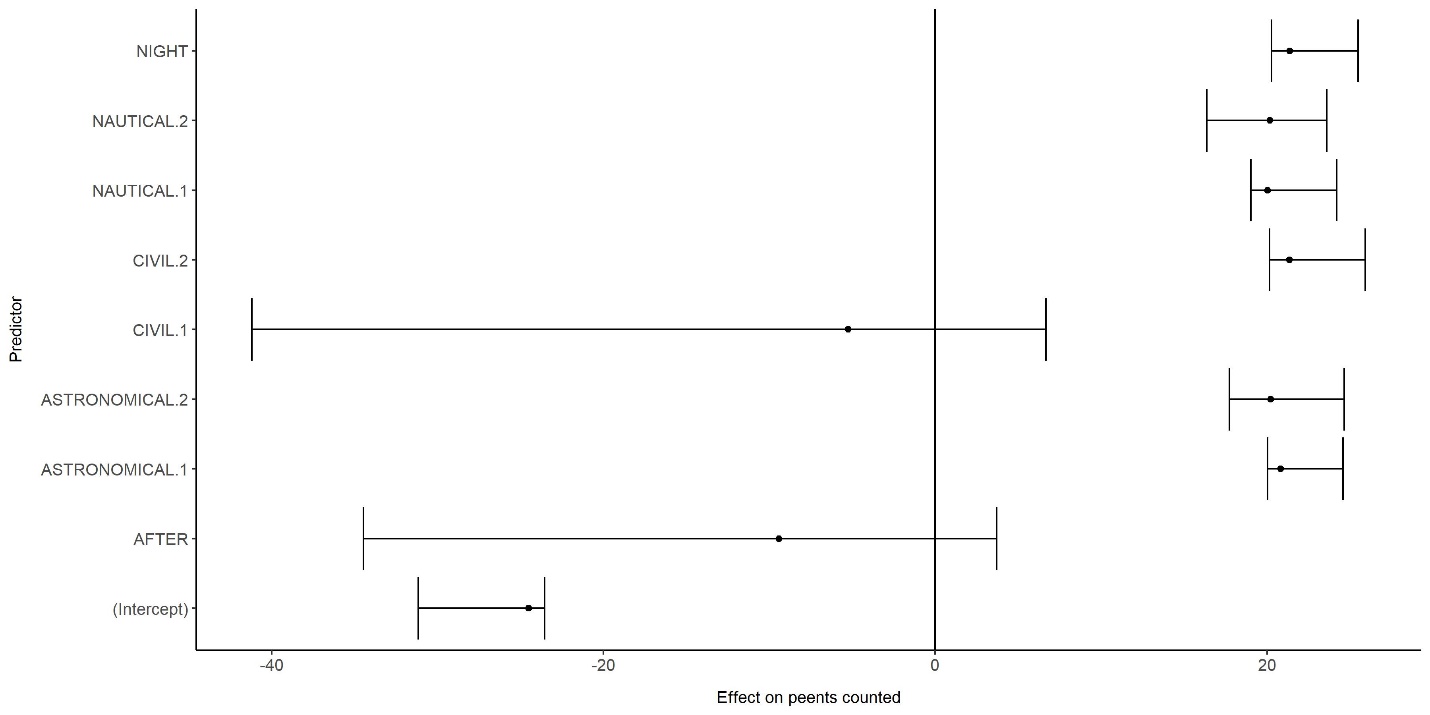


### Figure 16. Probability of detection given occupancy (represented as prediction intervals) based on boom detections in the samples versus number of visits per site (3-20) and duration of visit (in minutes).

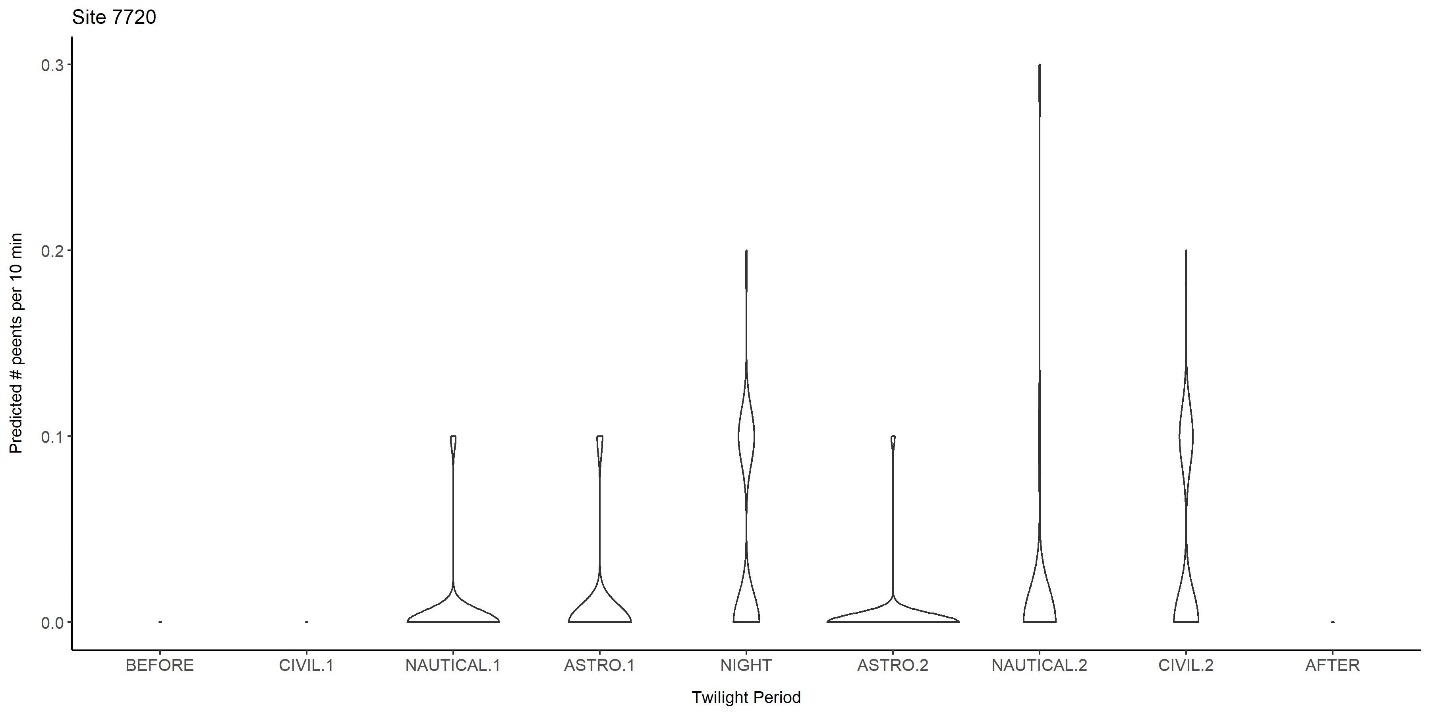
## Assessing Effects of Twilight Period and Latitude on Activity Rates

In the non-interaction model, the model converged for only 38 % of bootstrap samples. Peent activity was predicted to be higher for the twilight periods from nautical twilight 1 (before midnight) to civil twilight 2 (after midnight) than before sunset. Peent activity was not predicted to differ between before sunset, the civil twilight period immediately after sunset, or after sunrise (Figures 1-2). The goodness-of-fit of this model was poor: Spearman correlation coefficients between predicted and actual peent counts in the 38 sample data sets with predictions were weakly negative (median Spearman correlation = -0.07, 95 % BCI = -0.05 – -0.10).

In the interaction model, 98% of the peent activity models failed to converge during bootstrapping. As such, mixed model results for peent activity were questionable and not pursued further.

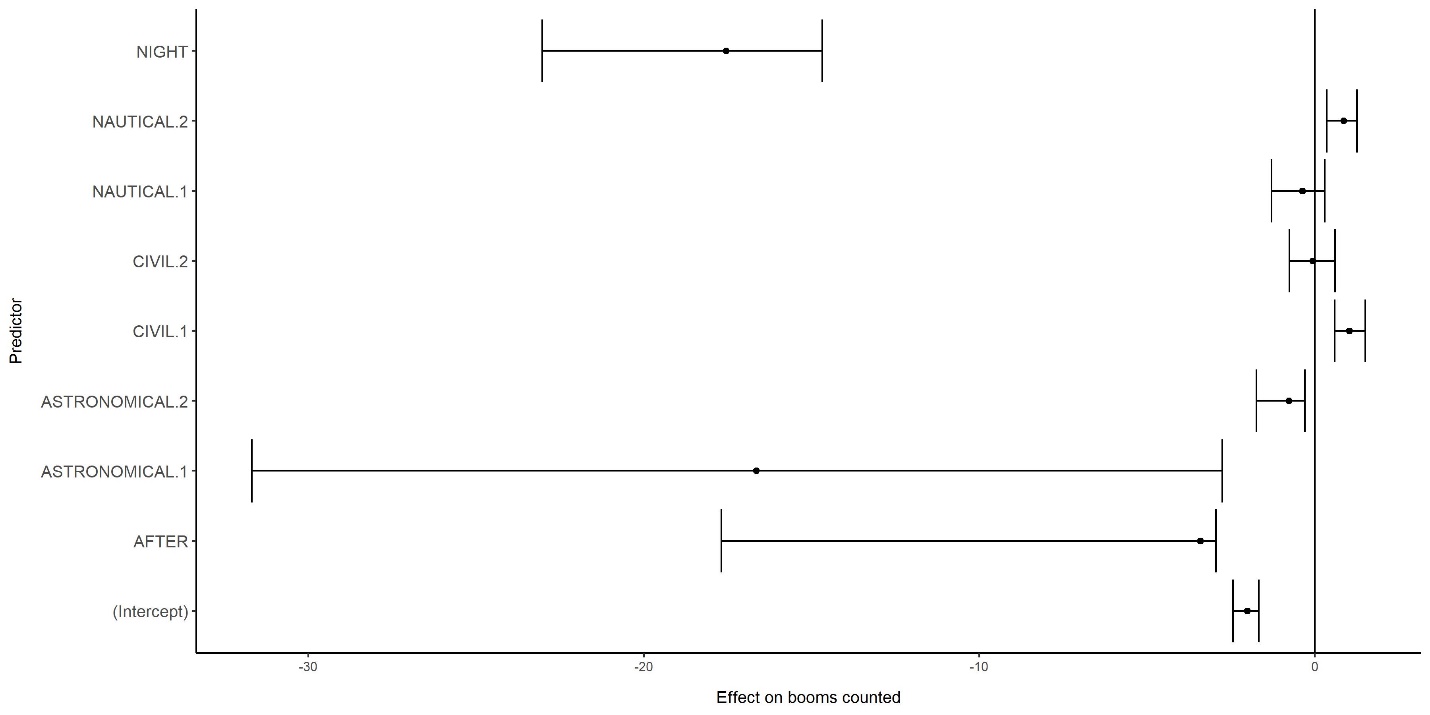


### Figure 17. Box plots showing the median, 5th and 95th quantile values for the model terms based on the 38 out of 100 bootstrapped sample data sets for which the following mixed model successfully converged and produced results: Peent activity rate = Twilight period. Note that in these results the “median” would be the mean of the two model coefficient values obtained for each fixed effect while the 5th and 95th quantiles would be respectively the lower and higher of the two model coefficient values. Effects of different twilight periods are relative to the 10-minute observation periods that took place before sunset (twilight period = “Before”).

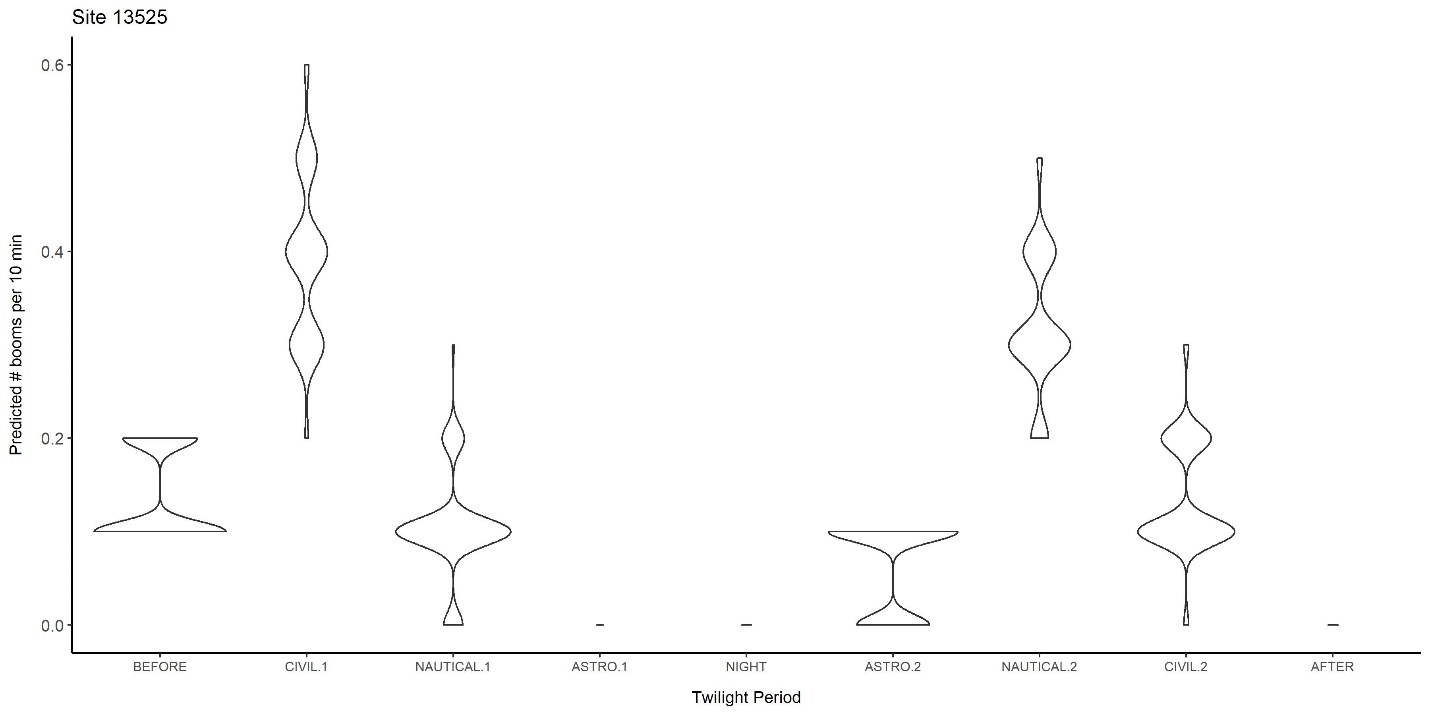


### Figure 18. Violin plots showing the distribution of predicted CONI boom activity rates (counts per 10-minute interval, rounded to 1 decimal place) in each of 9 twilight periods at each of 17 sites between ordinal days 170-190, based on the non-interaction mixed-effects model Peent Count ~ Twilight Period. Effects of different twilight periods are relative to the 10-minute observation periods that took place before sunset (twilight period = “Before”).

In the non-interaction model, the model converged for all 100 bootstrapped sample data sets. Boom activity was generally highest in the civil twilight period right after sunset and secondarily highest in the nautical twilight period after midnight. Boom activity was predicted to be negligible in the “Night” twilight period and after sunrise (Figure 3-4). When I compared predicted boom activity rates to actual boom detections per 10-minute interval, Spearman correlations were moderately positive (median Spearman correlation = 0.24, 95 % BCI = 0.25 – 0.26).

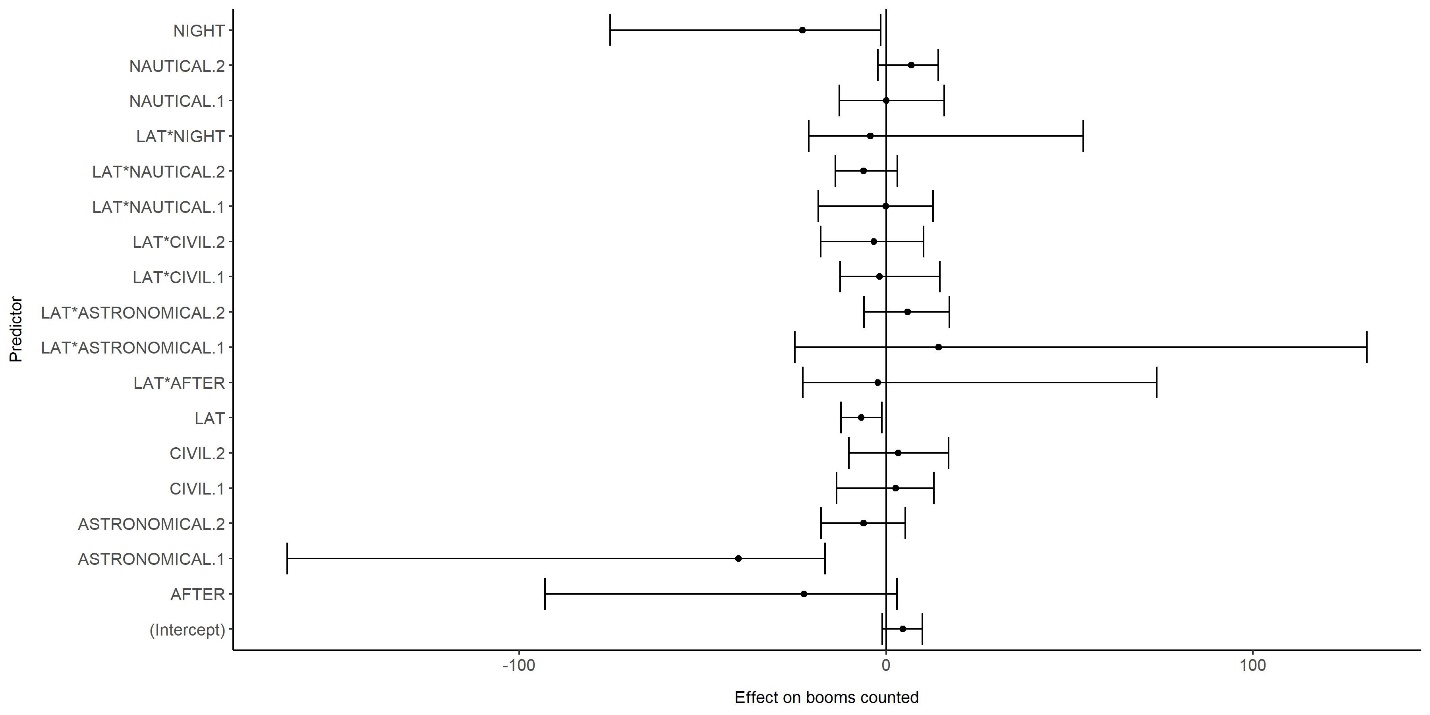


### Figure 19. Box plots showing the median, 5th and 95th quantile values for the model terms based on the 27 out of 100 bootstrapped sample data sets for which the following mixed model successfully converged and produced results: Boom activity rate = Twilight period. Effects of different twilight periods are relative to the 10-minute observation periods that took place before sunset (twilight period = “Before”).

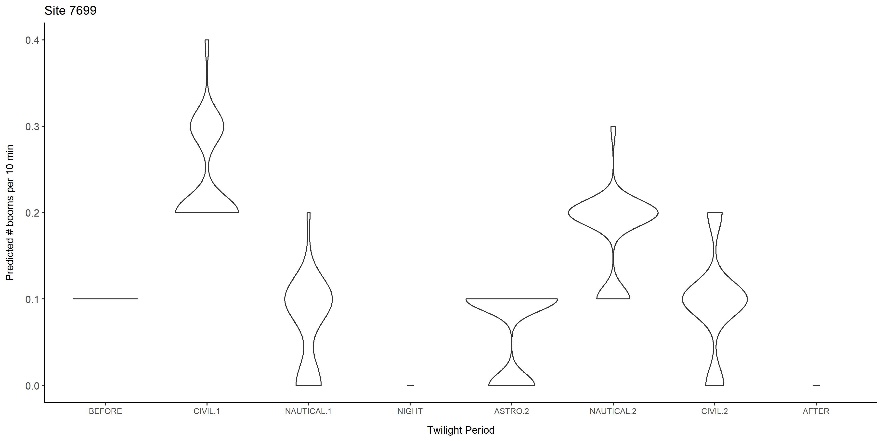
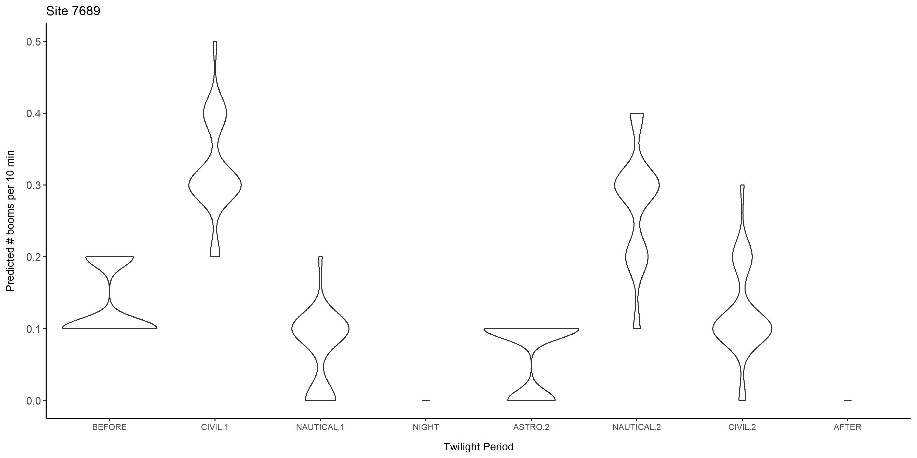
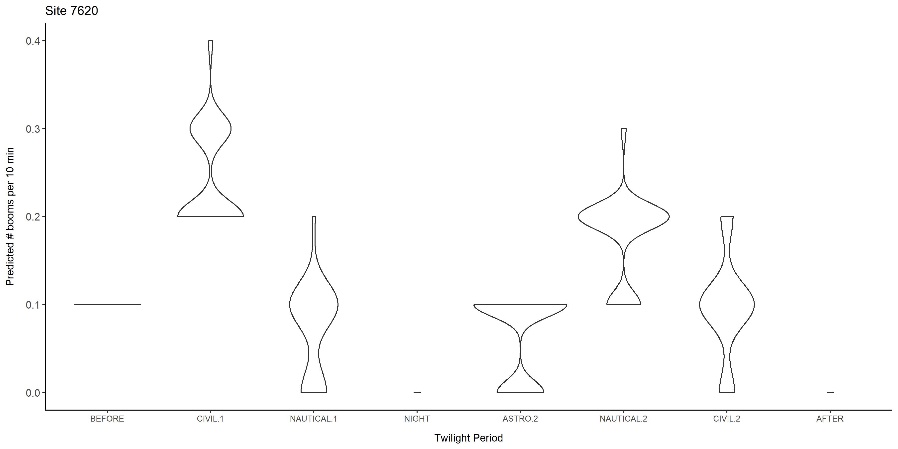
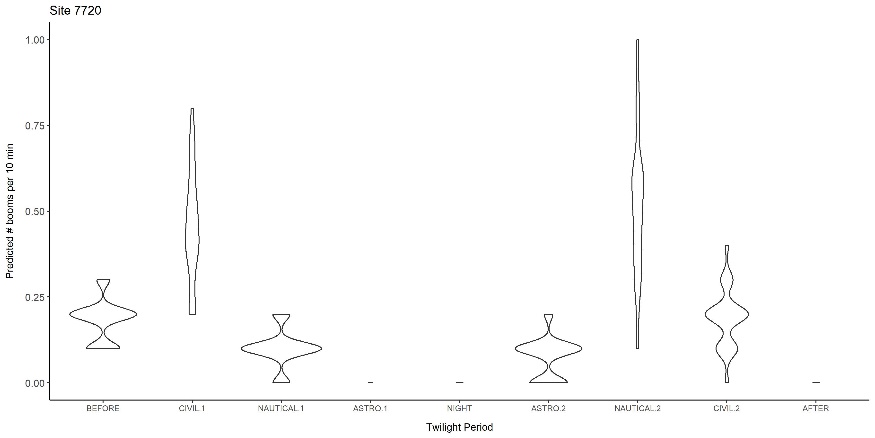


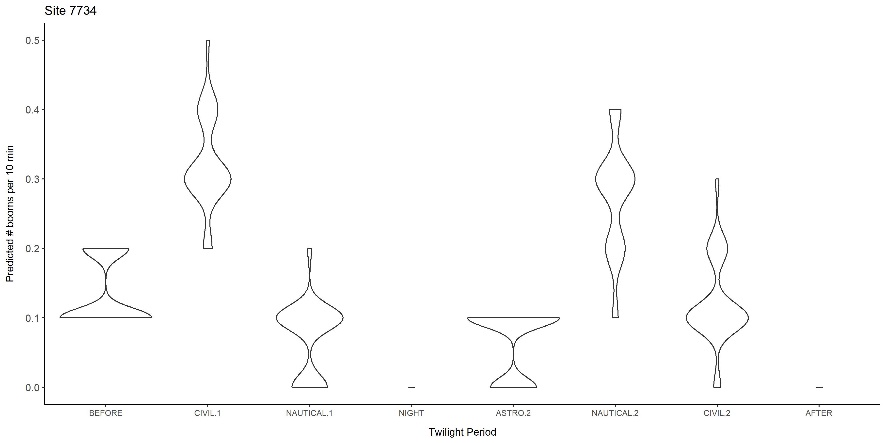
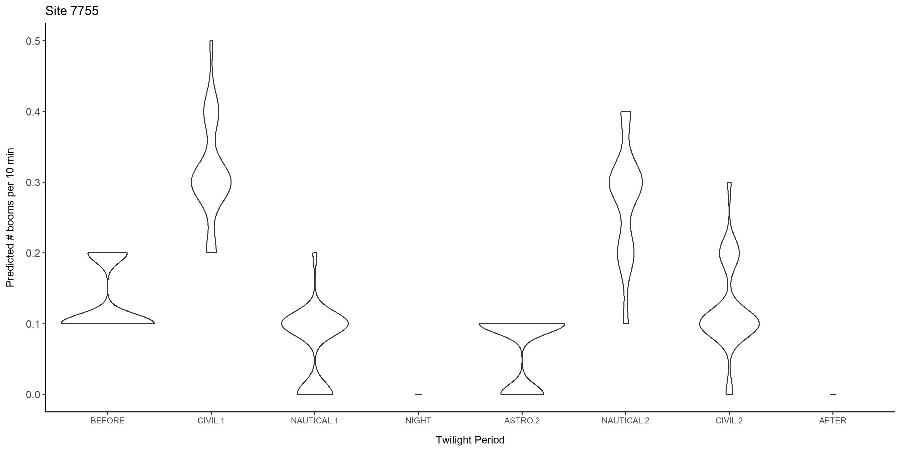
### Figure 20. Violin plots showing the distribution of predicted CONI boom activity rates (counts per 10-minute interval, rounded to 1 decimal place) in each of 9 twilight periods at each of 17 sites between ordinal days 170-190, based on the non-interaction mixed-effects model Boom Count ~ Twilight Period. Effects of different twilight periods are relative to the 10-minute observation periods that took place before sunset (twilight period = “Before”).

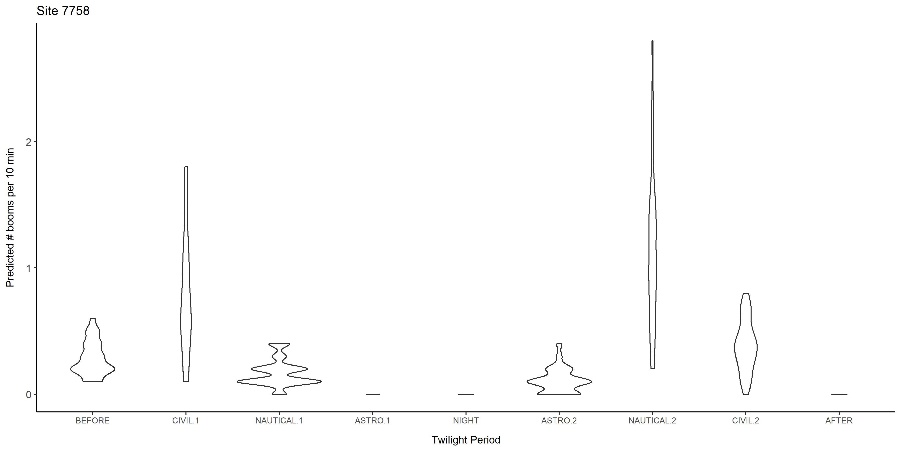
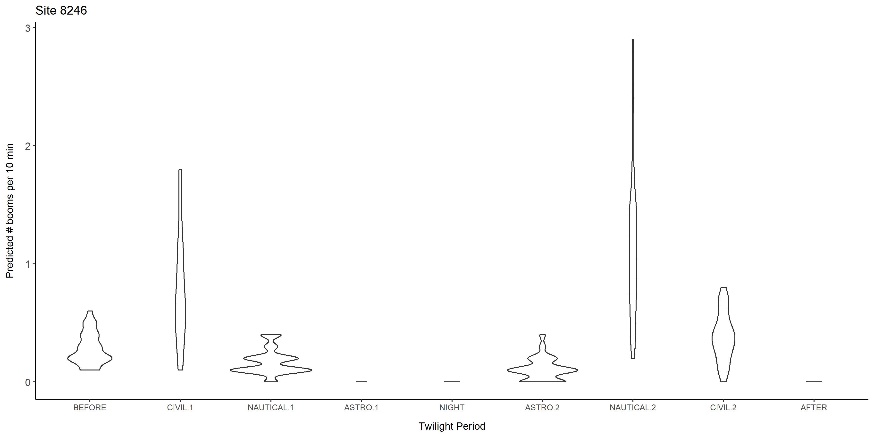
In the interaction model, only 27% of the boom activity models converged during bootstrapping. Based on these results, there was no strong evidence that twilight period effects on boom activity differed with latitude, although boom activity rates were lower at higher latitudes (Figure 5). When I compared predicted boom activity rates to actual boom detections per 10-minute interval, Spearman correlations were moderately positive (median Spearman correlation = 0.29, 95 % BCI = 0.27 – 0.31, higher than in the non-interaction model). As in the non-interaction model, boom activity was generally highest in the civil twilight period before midnight and secondarily highest in the nautical twilight period after midnight. Boom activity was predicted to be negligible in the “Night” twilight period and after sunrise (Figure 6).

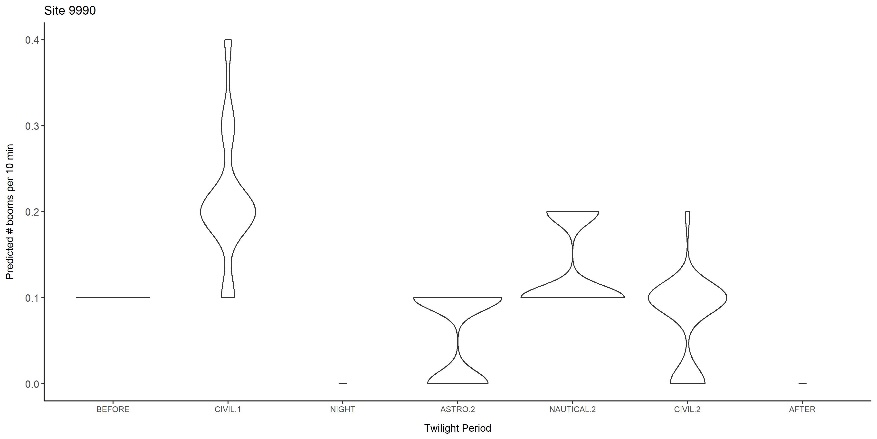
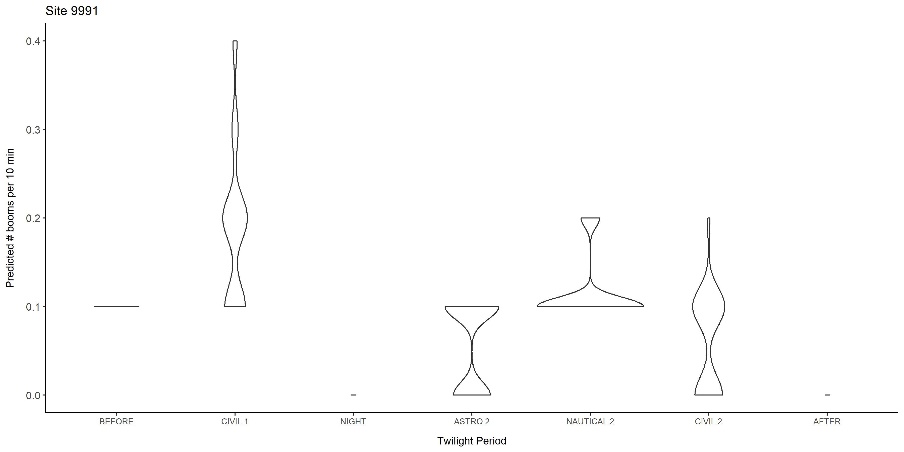


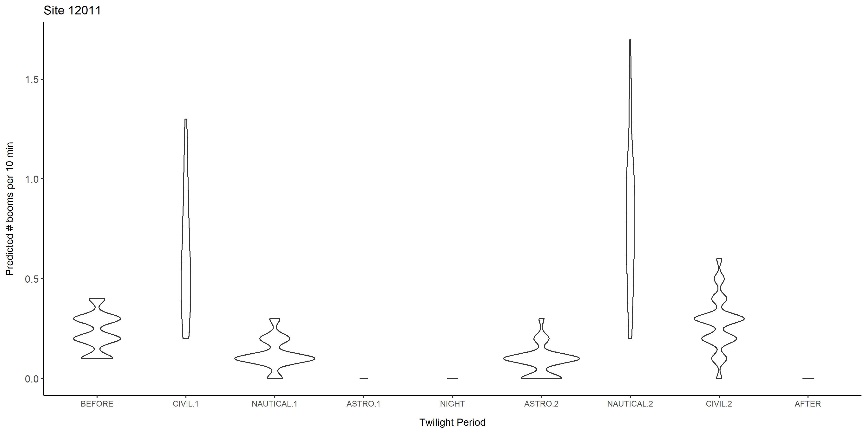
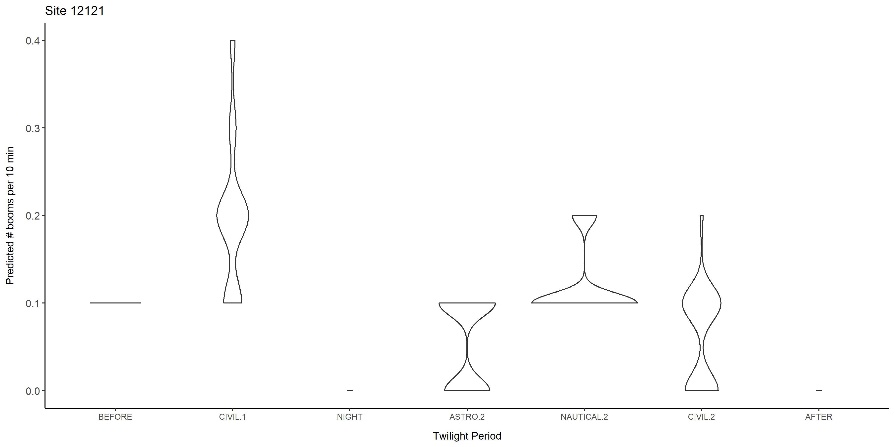
### Figure 21. Box plots showing the median, 5th and 95th quantile values for the model terms based on the 27 out of 100 bootstrapped sample data sets for which the following mixed model successfully converged and produced results: Boom activity rate = Twilight period + latitude + twilight period\*latitude. Effects of different twilight periods and twilight period\*latitude are relative to the 10-minute observation periods that took place before sunset (twilight period = “Before”).

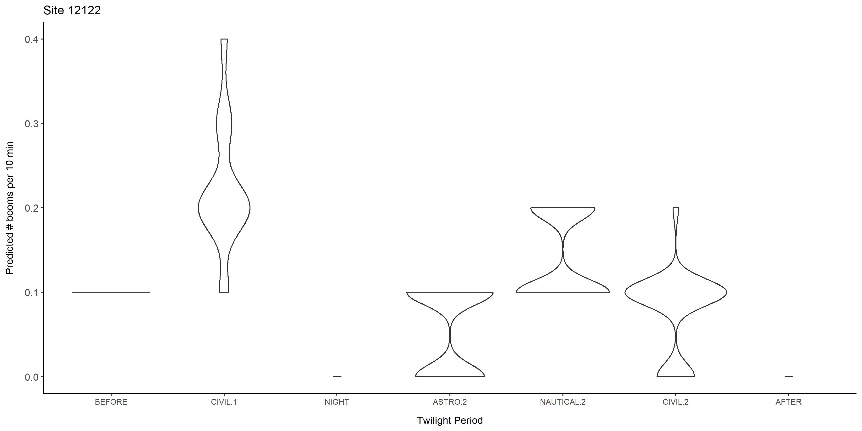
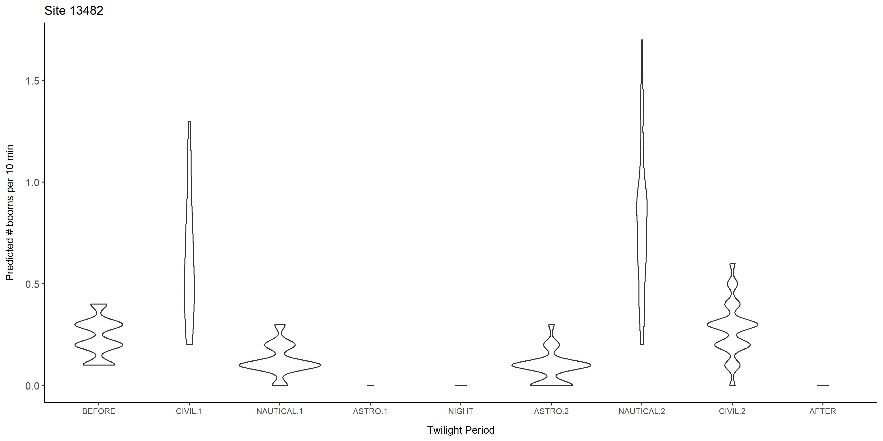
 

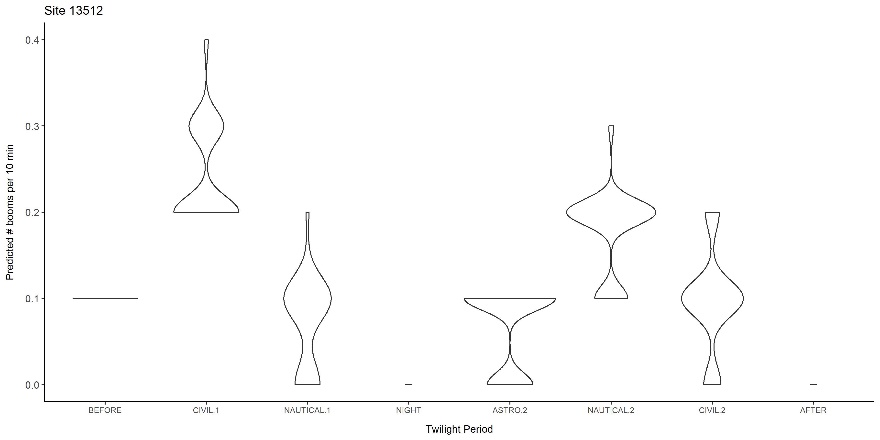
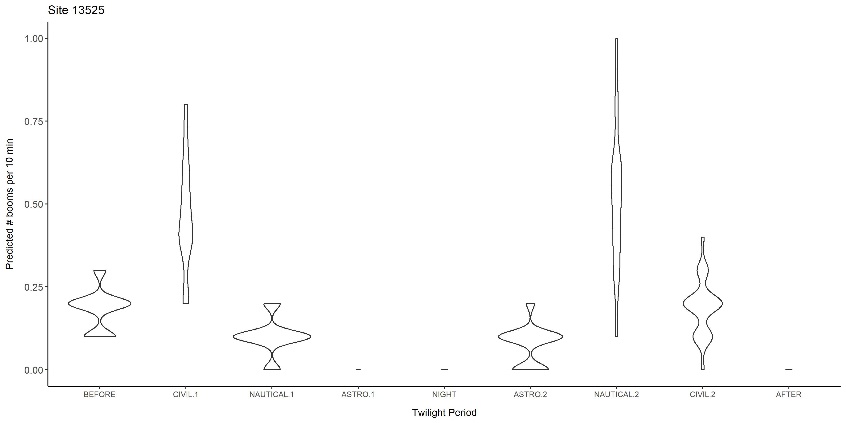
 

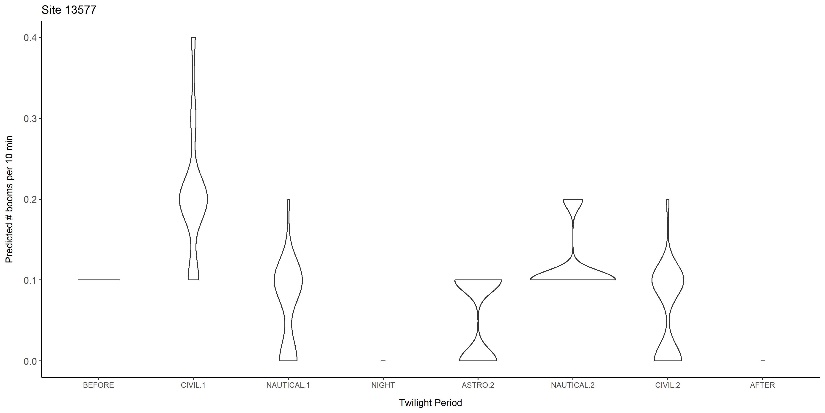
 



### Figure 22. Violin plots showing the distribution of predicted CONI boom activity rates (counts per 10-minute interval, rounded to 1 decimal place) in each of 9 twilight periods at each of 17 sites between ordinal days 170-190, based on the non-interaction mixed-effects model Boom Count ~ Twilight Period. Effects of different twilight periods are relative to the 10-minute observation periods that took place before sunset (twilight period = “Before”).

# Discussion

## Recommended Timing of CONI Surveys Within Season and Night

My analyses suggested that aside from increasing the number and/or duration of visits, both activity rates and probability of detection of CONI would be maximized by conducting surveys at or shortly after sunset, with point counts at sunrise generally being a second choice. CONI surveys should preferably be conducted from June 9 to June 29 when this species is more likely to be detected and is most frequently booming and calling.

Temperature and moonlight had weaker effects. When temperature occurred in top occupancy models as an effect, probability of peent detection increased with mean nightly temperature, possibly because warmer nights were associated with increased insect activity. Alternatively, CONI might have entered torpor on the coldest nights, though this possibility is questionable (Fletcher et al. 2004). Mean nightly temperature varied from 10.4 – 21.4 ° C. on nights that had temperature data taken. In contrast to temperature, peent and boom detection probabilities had contrasting relationships with respect to moonlight. As moon illumination increased, probability of detecting peents declined while probability of detecting booms increased slightly. If the intent of surveys is to establish evidence of breeding, nighthawk surveys should be scheduled when moonlight is waxing or full.

## Recommended Number of Survey Visits and Length of Surveys

If CONI data are collected so as to run occupancy models on the data, my analyses suggest that increasing both number of visits per site and the duration of visits will increase the probability that any present CONI are detected; however, model predictive power may not improve or may even worsen with an increased number of visits. My analyses suggested that accuracy of predicting joint probability of detection, as measured by area-under-the-curve, dropped sharply when there were more than 4 visits per site, whether I was analyzing detection of peents or booms. It should be noted that in these analyses I focused entirely on detection covariates without modelling site-specific occupancy, so it might have been possible to improve prediction accuracy by including site covariates. However, the study was not specifically designed to be analyzed by occupancy models, since CONI were detected at most of the sites in 2014.

Further, the use of continuously recorded ARU data to construct my simulated visits potentially gave me a much larger number of sample visits per site than is usually available to occupancy modelling studies. It is possible that some assumptions of occupancy modelling may break down with a very large number of visits. I ran occupancy models using visits over a two-month period. This period might be short enough to assume population closure, since I focused on detection covariates rather than occupancy covariates. However, CONI were only detected booming on a single day of visits at one of the sites. While it is possible that at that site, the CONI could have settled down to breed within 4 days of that date, it is also possible that CONI may have simply moved on from that site after a single boom-recording date, in which case the assumption of population closure was violated at that site. Multi-season occupancy models could have been used to relax the assumption of population closure. I used single-season occupancy models because I focused on detection rather than occupancy covariates.

Another related issue with how a large number of survey visits could cause occupancy model assumptions to fail is that if individual CONI are just passing through a site, a single visit with a detection out of a large number of visits will result in an extremely low estimated probability of detection given occupancy, and consequently an overestimated probability of occupancy. The percentage of sites where CONI are predicted to occur will then be overestimated, which could have serious implications for conservation if CONI are then mistakenly assessed as being at lower risk.

Reducing the number of visits per site to just 3-4 per season – the minimum recommended number visits for occupancy models – could reduce the amount of time spent processing data from each site, potentially enabling a larger number of sites to be visited using ARUs. In doing so, studies could potentially increase the inference of their results by covering a wider variety of conditions (e.g. nighthawk habitats) for the same amount of processing time outside of field work. Alternatively, instead of increasing the number of sites, researchers could use the processing time liberated by having fewer visits to increase the duration of recordings examined at each site. Given that visit duration was a stronger predictor than number of visits of both peent and boom detection probability, investing in longer recordings rather than more recordings is probably more likely to result in more detections of cryptic species.

Point counts within the BBS consist of 50 stops along a 39.4 km roadside route (an average of 788 m apart). At these stops, observers count all birds of all species seen and heard within a 3-minute period. Each point count is surveyed on a single visit per year (Sauer et al. 1994). Our occupancy model results indicate that the length of BBS visits is probably too short to detect many CONI that are present at a site, since detection probability in my analyses increased with visit duration and did not reach an asymptote with the durations I analyzed. Further, the probability of detecting CONI that are present will be low with a single visit, since I found that 12-20 visits were required before the proportion of sites with detections reached an asymptote.

Point counts within the BBS are tailored to detect as many species of birds as possible. Given that most species are songbirds, such point counts occur during the daytime when songbird species are most likely to be singing. My occupancy models and GAMMs suggest that CONI are more likely to be detected and peent-calling and/or booming more often around sunset, with a smaller secondary peak in activity around sunrise. While some CONI may be detected in early daytime point counts at sunrise, the BBS might increase the number of detections of this species for analyses by collecting point counts at sunset, especially in areas that are likely to be suitable CONI habitats (roadside areas with sand, gravel, and rocky outcrops, or roadsides passing through recently burned or logged areas). Since BBS point counts within a BBS route occur closer to each other than the distance between sites in the CWS-ON CONI study, BBS points selected for additional sunset surveys should be further apart than 788 m to reduce double-counting of CONI.

Apart from TSSS, the most important variable influencing detection probability of CONI was ordinal day. CONI were generally most likely to be detected in occupancy models from 160-180 days since January 1 (June 9-June 29). Peent and boom activity rates were also generally highest within this window of time. Additional point counts tailored towards the collection of CONI data at sunset should be concentrated within this 2-3 week period.

Although TSSS and ordinal day were the primary detection variables, the BBS is used to establish evidence of breeding in bird species; therefore, detection of breeding evidence in CONI in the BBS’s point counts might be increased slightly by conducting surveys on nights with greater moon illumination. However, the 3-minute point counts used in the BBS may be too short for moon illumination to be useful, since moonlight was only a factor in the top models for data sets with longer visit durations.

## Recommended Timing of CONI Surveys With Respect to Twilight Period

Similar to the GAMM results, mixed-effects models of twilight period suggested that boom activity rates were consistently higher at sites in the civil twilight period immediately after sunset. While a secondary peak in activity existed at some sites, during nautical twilight after midnight, this period will probably be less convenient for people to go out and monitor CONI activity. CONI boom activity was predicted to be negligible after sunrise, in contrast to the singing activity of most songbirds in the breeding season. Thus, if the purpose of surveys is detection of CONI breeding behaviour, dedicated surveys at civil twilight will detect higher rates of boom activity and more evidence of breeding than will multi-species, all-purpose surveys that are tailored to songbirds. The lack of variation in twilight period effects with increasing latitude suggests that CONI surveys done during civil twilight are appropriate for the latitudes that we tested in our twilight period models.

# Citations

Agafonkin, V., & Thieurmel, B. (2017). suncalc: Compute sun position, sunlight phases, moon position, and lunar phase. R package version 0.3.

Barton, K., & Barton, M. K. (2019). Package ‘MuMIn’. R package version, 1(6).

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48.

Brigham, R. M., Ng, J., Poulin, R. G., & Grindal, S. D. (2011). Common Nighthawk (*Chordeiles minor*). In A. Poole, editor. The Birds of North America Online. Cornell Lab of Ornithology, Ithaca, NY, USA.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. Sociological Methods & Research, 33(2), 261-304.

Buxton, R. T., & Jones, I. L. (2012). Measuring nocturnal seabird activity and status using acoustic recording devices: applications for island restoration. Journal of Field Ornithology, 83(1), 47-60.

Caccamise, D. F. (1974). Competitive relationships of the Common and Lesser Nighthawks. Condor, 76(1), 1-20.

Canaday, C. (1997) Loss of insectivorous birds along a gradient of human impact in Amazonia. Biological Conservation, 77, 63-77.

Charif, R. A., Waack, A. M., & Strickman, L. M. (2010). Raven Pro 1.4 User's Manual. Cornell Lab of Ornithology, Ithaca, NY. 367 pp.

Fiske, I., & Chandler, R. (2011). Unmarked: an R package for fitting hierarchical models of wildlife occurrence and abundance. Journal of statistical software, 43(10), 1-23.

Fletcher, Q. E., Fisher, R. J., Willis, C. K., & Brigham, R. M. (2004). Free-ranging common nighthawks use torpor. Journal of Thermal Biology, 29(1), 9-14.

Frommolt, K. H., & Tauchert, K. H. (2014). Applying bioacoustic methods for long-term monitoring of a nocturnal wetland bird. Ecological Informatics, 21, 4-12.

Frommolt, K. H. (2017). Information obtained from long-term acoustic recordings: applying bioacoustic techniques for monitoring wetland birds during breeding season. Journal of Ornithology, 158(3), 659-668.

Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate. Journal of Statistical Software, 40(3), 1-25.

Hallman, C. A., R. P. B. Foppen, C. A. M. van Turnhout, H. de Kroon, and E. Jongejans (2014) Declines in insectivorous birds are associated with high neonicotinoid concentrations. Nature, 511(7509), 341.

Nebel, S., A. Mills, J. McCracken, and P. Taylor (2010). Declines of aerial insectivores in North America follow a geographic gradient. Avian Conservation and Ecology, 5(2).

Oppel, S., Hervias, S., Oliveira, N., Pipa, T., Silva, C., Geraldes, P., Goh, M., Immler, E. & McKown, M. (2014). Estimating population size of a nocturnal burrow-nesting seabird using acoustic monitoring and habitat mapping. Nature Conservation, 7, 1.

Paquette, S. R., F. Pelletier, D. Garant, and M. Bélisle (2014). Severe recent decrease of adult body mass in a declining insectivorous bird population. Proceedings of the Royal Society B, 281, 20140649.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC bioinformatics, 12(1), 77.

Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., Stanton, J. C., Panjabi, A., Helft, L., Parr. M, & Marra, P. P. (2019). Decline of the North American avifauna. Science, 366(6461), 120-124.

Sauer, J. R., Peterjohn, B. G., & Link, W. A. (1994). Observer differences in the North American breeding bird survey. The Auk, 111(1), 50-62.

Sekercioglu, C. H., P. R. Ehrlich, G. C. Daily, D. Aygen, D. Goehring, and R. F. Sandi (2002) Disappearance of insectivorous birds from tropical forest fragments. Proceedings of the National Academy of Sciences, 99, 263-267.

Smith, A. C., M.-A. R. Hudson, C. M. Downes, and C. M. Francis (2015) Change points in the population trends of aerial-insectivorous birds in North America: synchronized in time across species and regions. PLOS One, 10(7): e0130768. doi:10.1371/journal.pone.0130768.

Sólymos, P. (2017). <https://cran.r-project.org/web/packages/intrval/intrval.pdf>.

Tegeler, A. K., Morrison, M. L., & Szewczak, J. M. (2012). Using extended‐duration audio recordings to survey avian species. Wildlife Society Bulletin, 36(1), 21-29.

Weller, M. W. (1958). Observations of the incubation behavior of a Common Nighthawk. Auk, 75(1), 48-59.Wood, S., Scheipl, F., & Wood, M. S. (2017). Package ‘gamm4’. Am. Stat., 45, 339.

# Abbreviations Used in This Report

ARU = Acoustic or Autonomous Recording Unit

BBS = North American Breeding Bird Survey

CONI = Common Nighthawk (*Chordeiles minor*)

TSSS = Time Since Sunset