National Park Service
U.S. Department of the Interior

Natural Resource Stewardship and Science

Development of a Robust Analytical Framework for Assessing Landbird Population Trends, Dynamics and Relationships with Environmental Covariates in the North Coast and Cascades Network

Natural Resource Report NPS/NCCN/NRR—2017/1483

ON THE COVER
Varied thrush
NPS photo

Development of a Robust Analytical Framework for Assessing Landbird Population Trends, Dynamics and Relationships with Environmental Covariates in the North Coast and Cascades Network

Natural Resource Report NPS/NCCN/NRR—2017/1483

Chris Ray,[1] James Saracco,[1] Kurt Jenkins,[2] Mark Huff,[3] Patricia Happe,[4] Jason Ransom[5]

[1]The Institute for Bird Populations
Point Reyes Station, CA

[2]U. S. Geological Survey
Forest and Rangeland Ecosystem Science Center
Olympic Field Station
Port Angeles, WA

[3]National Park Service
North Coast and Cascades Network
Ashford, WA

[4]National Park Service
Olympic National Park
Port Angeles, WA

[5]National Park Service
North Cascades National Park Service Complex
Sedro-Woolley, WA

July 2017

U.S. Department of the Interior
National Park Service
Natural Resource Stewardship and Science
Fort Collins, Colorado

The National Park Service, Natural Resource Stewardship and Science office in Fort Collins, Colorado, publishes a range of reports that address natural resource topics. These reports are of interest and applicability to a broad audience in the National Park Service and others in natural resource management, including scientists, conservation and environmental constituencies, and the public.

The Natural Resource Report Series is used to disseminate comprehensive information and analysis about natural resources and related topics concerning lands managed by the National Park Service. The series supports the advancement of science, informed decision-making, and the achievement of the National Park Service mission. The series also provides a forum for presenting more lengthy results that may not be accepted by publications with page limitations.

All manuscripts in the series receive the appropriate level of peer review to ensure that the information is scientifically credible, technically accurate, appropriately written for the intended audience, and designed and published in a professional manner.

Data in this report were collected and analyzed using methods based on established, peer-reviewed protocols and were analyzed and interpreted within the guidelines of the protocols. This report received formal peer review by subject-matter experts who were not directly involved in the collection, analysis, or reporting of the data, and whose background and expertise put them on par technically and scientifically with the authors of the information

Views, statements, findings, conclusions, recommendations, and data in this report do not necessarily reflect views and policies of the National Park Service, U.S. Department of the Interior. Mention of trade names or commercial products does not constitute endorsement or recommendation for use by the U.S. Government.

This report is available in digital format from the North Coast and Cascades Network Inventory and Monitoring website and the Natural Resource Publications Management website. To receive this report in a format that is optimized to be accessible using screen readers for the visually or cognitively impaired, please email irma@nps.gov.

Please cite this publication as:

Ray, C., J. Saracco, K. Jenkins, M. Huff, P. Happe, and J. Ransom. 2017. Development of a robust analytical framework for assessing landbird population trends, dynamics and relationships with environmental covariates in the North Coast and Cascades Network. Natural Resource Report NPS/NCCN/NRR—2017/1483. National Park Service, Fort Collins, Colorado.

NPS 963/139316, July 2017

Contents

Page

Figures

Appendices

Page

Executive Summary

During 2015-2016, we completed development of a new analytical framework for landbird population monitoring data from the National Park Service (NPS) North Coast and Cascades Inventory and Monitoring Network (NCCN). This new tool for analysis combines several recent advances in modeling population status and trends using point-count data and is designed to supersede the approach previously slated for analysis of trends in the NCCN and other networks, including the Sierra Nevada Network (SIEN). Advances supported by the new model-based approach include 1) the use of combined data on distance and time of detection to estimate detection probability without assuming perfect detection at zero distance, 2) seamless accommodation of variation in sampling effort and missing data, and 3) straightforward estimation of the effects of downscaled climate and other local habitat characteristics on spatial and temporal trends in landbird populations. No changes in the current field protocol are necessary to facilitate the new analyses. We applied several versions of the new model to data from each of 39 species recorded in the three mountain parks of the NCCN, estimating trends and climate relationships for each species during 2005-2014. Our methods and results are also reported in a manuscript in revision for the journal Ecosphere (hereafter, Ray et al.). Here, we summarize the methods and results outlined in depth by Ray et al., discuss benefits of the new analytical framework, and provide recommendations for its application to synthetic analyses of long-term data from the NCCN and SIEN. All code necessary for implementing the new analyses is provided within the Appendices to this report, in the form of fully annotated scripts written in the open-access programming languages R and JAGS.

Acknowledgments

This analysis was funded by the U.S. Geologcial Survey (USGS) National Parks Monitoring Program under Agreement G14AC00179, with additional in-kind support from the National Park Service North Coast and Cascades Inventory and Monitoring Network (NPS NCCN), NPS Olympic and North Cascades National Parks, The Institute for Bird Populations (IBP), and USGS Forest and Rangleand Ecosystem Sciences Center. We are grateful for comments provided in review by Joshua Schmidt and Thomas Rodhouse from the NPS Inventory and Monitoring Program. Data for the analyses was collected and managed by NPS NCCN under auspices of its Landbird Monitoring Project and Cooperative Agreement with IBP. We thank the many individuals from NCCN (especially John Boetsch), IBP (especially Mandy Holmgren, Robert Wilkerson and Rodney Siegel), and Olympic, Mount Rainier, and North Cascades National Parks for providing logistical and data management support that made this analysis possible. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Glossary

Imperfect detection -- Detection probability (detectability) is the probability of detecting an individual of a given species during a point count survey, given that individual is present. Detection is imperfect when detection probability is less than one. Many factors can affect detection probability for a given species and survey method, including the individual's distance from an observer, observer skill, weather, time of day or season, and habitat structure.

Modeling -- A model is a formal simplification of a more complex system that we are trying to understand. Modeling is the practice of abstracting relevant processes and describing them, often mathematically, in ways that allow useful and testable predictions. Estimating a parameter, like population size, is a form of modeling, in which key assumptions are made to simplify the relationship between what we know (e.g., a count) and what we want to know (the actual number of individuals present).

Hierarchical model -- Hierarchical models involve the decomposition of a complex relationship into a series of interdependent submodels to facilitate understanding and computation as well as to clarify all levels of uncertainty in the relationship. A simple example would involve a sequence of random variables, e.g., a count of y individuals, a population size of N, and a detection probability of p, with perhaps several factors affecting detection. The hierarchy in this example involves one level in which y is a function of N and p, and another level in which p is a function of factors like observer identity, habitat qualities and seasonal processes (e.g., initiation of breeding behavior).

Parameter -- A parameter is a feature of a population or process that we can estimate using a sample and sample statistics. For example, population density is a parameter that can be estimated from individuals detected per unit of area.

Bayesian model -- A Bayesian model is one that allows us to estimate parameter values by incorporating prior information about the distribution of values each parameter might take, using Bayes' rule. For example, if we were trying to estimate the abundance of a population and we had some samples from a current study and some samples from a prior study, we could use a Bayesian model to update our prior estimate of abundance using our current data. Our prior estimate (E) would have some distribution of probable values (e.g., a mean and standard error, or perhaps a range of potential values with no central tendency). This prior estimate can be written as p(E), the prior probability distribution or "prior density" of our estimate. An updated version of p(E), based on new data (D), would be p(E|D) or the probability density of E given D. Bayes' rule, which is often presented as $p(E|D) \sim p(D|E)p(E)$, provides a method for calculating p(E|D), which is our updated (better) estimate, aka the "posterior density" of E. Bayes' rule shows us that to find the posterior p(E|D), we adjust the prior p(E) by p(D|E), which is the probability density of observing D given some potential values of E, aka the likelihood of D given E. The best current estimate for E is the one that maximizes this likelihood, and the full posterior density for E is determined by multiplying the densities of the likelihood and the prior. Bayesian models are built on a rigorous theory and are quite useful in practice, but their use has lagged behind classical "frequentist" methods because calculating the joint probability density of all the parameters in even a moderately complex model was impractical before the advent of modern computers and simulation methods such as Markov chain Monte Carlo.

Posterior predictive distribution -- The posterior predictive distribution is the distribution that a set of predicted data would have, based on a predictive model and the distribution of observed data.

Markov chain Monte Carlo -- Markov chain Monte Carlo (MCMC) is a simulation technique used to draw samples from a posterior distribution, allowing for the analysis of even very complex Bayesian models. For example, plotting a raw or smoothed histogram of MCMC samples can reveal the shape of the posterior distribution, which defines the density of an estimated parameter or the joint density of multiple parameters. Here, the JAGS programmable platform (Plummer 2003) is used to perform MCMC simulation and to provide summaries of the resulting samples.

Credible interval -- A credible interval is the Bayesian counterpart to the traditional (frequentist) confidence interval, with an important distinction: Using the 95% confidence interval as an example, the frequentist version does not actually contain the target parameter value with a probability of 95%, but a Bayesian 95% *credible* interval *does* contain the target parameter value with a probability of 95%. This distinction means that Bayesian credible intervals can be used to interpret the probability of model outcomes and parameter values in a more straightforward manner. In this report, "95% CI" refers to a Bayesian credible interval.

Binomial random variable -- A binomial variable is a quantity that can vary according to a binomial distribution, its values being described well by the discrete probability distribution of the number of successes in a sequence of n independent yes/no experiments, each of which yields success with probability p, such as the outcome of attempts at detecting a bird.

Generalized linear models -- A linear model relates a response variable y to predictor variables under the assumption that the relationship is linear and that the error distribution for y is normal (Gaussian). The error distribution in this case describes the probability of any given difference between observed y and y predicted by the linear relationship. In generalized linear models, the assumption of a Gaussian error distribution is relaxed, allowing for response variables that have other error distributions, such as Poisson or binomial distributions.

Poisson model -- “Poisson model” is a general term indicating that we have assumed the variable of interest can be modeled as a random draw from a Poisson distribution, which differs from a Gaussian distribution in two key features: the variance of a Poisson distribution is equal to its mean, and its values are discrete and positive. A Poisson model expresses a count that arises from independent events that occur at an average rate in space and/or time. A common form of “Poisson model” is Poisson regression, a form of generalized linear regression used when the response variable has a Poisson distribution.

***N*-mixture models** -- An N-mixture model applied to abundance estimation typically pairs a Poisson model of true N (abundance) with a binomial model of observed N (counts). The binomial mixture model is named for the mixture of binomial models, each with a potentially distinct sample size N, that determine model likelihood. N-mixture models provide a hierarchical extension of generalized linear models, linking multiple GLMs (e.g., Poisson and binomial) to allow for structure in parameters at each hierarchical level (e.g., environmental covariates of abundance and effects of survey conditions on detection).

1.0 Challenges in the estimation of trends from point-count data

In 2007, the National Park Service (NPS) North Coast and Cascades Inventory and Monitoring Network (NCCN) began monitoring landbird populations in five National Parks, in a program that has served as a model for other resource monitoring efforts within the NCCN and in other networks. The peer-reviewed NCCN landbird monitoring protocol (Siegel et al. 2007) was tightly integrated with an analytical approach for assessing population trends in the larger parks, involving a multi-step process whereby 1) observations were corrected for imperfect detection using a distance-based detectability function, 2) mean transect-scale trends were estimated for each elevation stratum based on the corrected counts, and 3) trends were estimated at the park (or combined-park) scale based on the mean stratum-level trends and the relative size of the sampling frame in each elevation stratum. During production of the first multi-year synthesis report on landbird trends (Saracco et al. 2014), several limitations to the original analytical approach were identified, including 1) low confidence in some of the assumptions involved in estimating detectability from data on detection distance, 2) often prohibitively large sample sizes needed to fully account for all likely sources of heterogeneity in detection probability, 3) failure to incorporate uncertainty in slope estimation at the scale of the individual transect into variance estimation at larger scales, 4) difficulty in estimating park-level trends for species that were rare in one or more elevation strata, and 5) the lack of a clear framework for assessing effects of covariates on temporal or spatial variation in species densities. Addressing these limitations is of particular interest for monitoring landbirds not only in the NCCN but also in other NPS units, because the NCCN monitoring protocol has in large part been adopted also as the monitoring protocol for the Sierra Nevada Network, and has been published as part of the NPS Natural Resource Report series (Siegel et al. 2010).

1.1 Overview of recent advances in methods of point-count analysis

Recent methods applied to point-count data allow for modeling various components of the detection process and analyzing observations, population state and trends simultaneously (Royle 2004, Alldredge et al. 2007a, Kéry 2008, Nichols et al. 2009, Dail and Madsen 2011, Schmidt et al. 2013, Amundson et al. 2014, Kéry and Royle 2016). Given within-season replication of counts, such as the multiple count intervals adopted in the NCCN protocol, individual detection histories can be used within a closed-population framework to account for detection probability when modeling population size (Alldredge et al. 2007a) or occupancy (Saracco et al. 2011). Trends in occupancy or population size can be modeled using detection histories within and across seasons (Royle 2004, MacKenzie et al. 2006, Kéry et al. 2009). Time- and distance-of-detection data available from NCCN point-counts can be used to exploit a combined time-removal and distance-sampling method for modeling population status while accounting for both the availability and perceptibility of individual birds (Amundson et al. 2014), where availability and perceptibility are distinct components of detection probability (Alldredge 2004). The Amundson et al. (2014) approach is suited especially for modeling data from situations in which only a single visit to each point-count station is possible within the breeding season. Although single-visit studies cannot directly address variation in detection probability throughout the breeding season (Schmidt et al. 2013, Mizel et al. 2017), many studies lack the resources to support multiple visits. Instead of temporally independent counts, these studies employ multiple count intervals within a single visit. Methods also exist to allow for count intervals of varying length (Farnsworth et al. 2002), an important feature of the NCCN protocol. Most of these methods are based on "*N*-mixture" models (Royle 2004) that typically pair a Poisson model of true *N* (abundance) with a binomial model of observed *N* (counts). The binomial mixture model is named for the mixture of binomial models, each with a potentially distinct sample size *N*, that determine model likelihood. *N*-mixture models provide a hierarchical extension of generalized linear models (GLMs), linking multiple GLMs (e.g., the Poisson and binomial GLMs described here) to model structure in parameters at each level of the hierarchy as functions of environmental or survey conditions (Kéry and Royle 2016). All of these methods address limitations of the previous analytical approach, by accounting for uncertainty in the detection and state processes simultaneously and at multiple spatial scales (point, transect, stratum, park). Importantly, *N*-mixture models can account for random effects (e.g., observer effects on detection), allow for species presence at points with zero detections (the previous approach considered zero counts as true zeros uncorrected for detection probability), and provide a model-based method for assessing trends. Implemented in a Bayesian framework and using the Markov chain Monte Carlo (MCMC) method for parameter estimation, this approach also allows parameter summaries at multiple scales.

1.2 Products proposed for improved estimation of trends

To address the challenges inherent in estimating NCCN landbird trends, we (the principals and cooperators) proposed to develop a new framework for trend analyses combining recent advances in modeling detectability, population status and trend simultaneously as functions of environmental covariates. Ideally, this new, model-based framework should accept data from the original sampling design while providing solutions to the challenges identified, as well as providing greater flexibility for dealing with missed or reduced sampling effort. In addition, we sought clear applicability to bird and other monitoring data from other networks. Specifically, the new analytical framework should employ the flexibility of Bayesian hierarchical models and the functionality of parameter estimation via Markov chain Monte Carlo (MCMC) methods. Hierarchical models were selected to take advantage of existing count replication within and among years (e.g., multiple count intervals within visits), features of the monitoring protocol that were previously unexploited. The objectives of our proposal were stated as follows:

- Trend analysis; we will use the new modeling framework to assess population trends of landbirds in the NCCN using annual monitoring data collected from 2008-2013.

- Assessing the effects of missing or reduced sampling effort on trend detection; we will assess the sensitivity of trend analyses to missing data, and evaluate inference costs of reduced sampling effort.

- Assessing the effects of weather or other covariates on annual occupancy rates; we will use park and regional climate data to assess the effect of snowpack and other annual weather variables on annual landbird population sizes in the parks.

- Modeling environmental correlates of species occupancy; for a small suite of declining species (e.g., Rufous Hummingbird, Olive-sided Flycatcher) we will use environmental covariates from other park data sets to model species-habitat relationships.

As stated in the funded proposal, this report will also explore whether any changes to field methods or operational aspects of the NCCN landbird monitoring program would be warranted to support the new analytical framework.

2.0 Specification of the new analytical framework

To encourage the adoption of our analytical framework and to facilitate the replication of our results, we used freely available software for all phases of our analysis. Parameter estimation was performed in JAGS (Plummer 2003), which employs MCMC simulation to fit Bayesian hierarchical models. We called JAGS from the R statistical computing environment (R Development Core Team 2015) using the jagsUI package (Kellner 2015), and used R for pre-processing data and post-processing results.

We tested our approach by fitting models to ten-year datasets for many of the species detected in the three mountain parks of the NCCN: North Cascades National Park Complex (NOCA), Olympic National Park (OLYM), and Mount Rainier National Park (MORA). Our three-level hierarchical model allowed estimation of 1) components of individual detection, including effects of covariates on probabilities of individual availability (e.g., would it sing during our count?) and perceptibility (would we hear it?); 2) point-year counts as a function of availability and perceptibility; and 3) spatiotemporal trends in abundance, including effects of climate. Model structure (Fig. 1) is detailed in Ray et al.

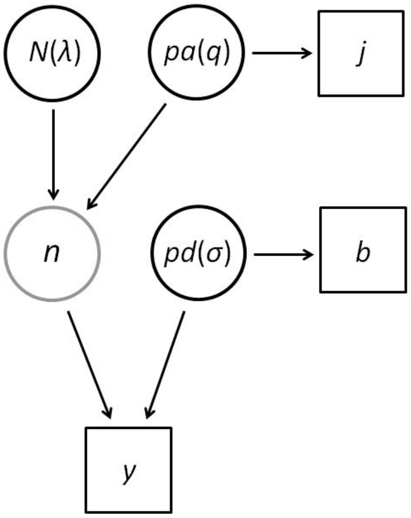


Figure 1. Graphical representation of directed relationships between data (squares) and estimated parameters (circles) in the models presented here. Focal parameters appear in black circles. Arrows indicate process flow. Actual bird abundance (N) was modeled as a Poisson-distributed function of mean abundance (λ) and its covariates (e.g., year, park, elevation or stratum, transect). The per-minute rate of detection failure (q) and its covariates (observer, date, hour) informed availability for detection (pa), which in turn influenced the interval (j) in which each bird was detected and our estimate of the number of birds available for detection (n). Finally, the rate at which detections fall off with distance (σ) and its covariates (tree cover, noise) controlled detectability (pd) and the observed detection distance-bin (b), as well as the number of available birds detected (y).

2.1 Modeling individual observations

Level 1 in this hierarchical modeling framework included removal-based estimation of the probability that individuals were available for detection (aka “availability”) and distance-based estimation of the probability that individuals were perceptible (aka “perceptibility”). In other words, birds detected in earlier intervals (e.g., those that sang more frequently) were more available for detection, and those detected at farther distances (e.g., louder singers) were more perceptible. Building on Amundson et al. (2014), detection probability at a given point k in year t, p_{kt}, was characterized in terms of a bird’s availability for detection, pa_{kt}, and its distance-mediated perceptibility to the observer, pd_{kt}. Note that this approach can account for variation in availability and perceptibility among years and points but, because our data derive from a single visit per point per season, we could not directly estimate variation in *pa* and *pd* throughout the breeding season at a given point. We adopted the Amundson et al. (2014) approach because the scale and topography of our sampling frame combined to preclude multiple visits within the short breeding season, given the resources available for monitoring these species. To minimize the effects of this constraint, we attempted to time our point counts to coincide with the local peak in territorial singing by most songbird species. Each year we began surveying transects in the low-elevation stratum in late May after most passage migrants had likely departed. We then proceeded to conduct surveys in the mid- and finally high-elevation strata as spring-like conditions developed at higher elevations (Siegel et al. 2007, 2010).

Because our data derived from multiple count intervals of different lengths (one interval of three minutes followed by one or two intervals of two minutes each), we followed Farnsworth et al. (2002) in modeling interval-specific probabilities of availability using time-removal data, in which each detection of a unique individual is assigned to one of the three count intervals, and subsequent detections of the same individual are ignored. Defining the per-minute probability of failing to detect a bird as q (Fig. 1), the probability that a bird was never detected across all three intervals (totaling seven minutes) was q_{kt}^{7}, so that $pa_{kt} = 1\text{-}q_{kt}^{7}$. By defining pa_{kt} in terms of $1\text{-}q_{kt}$, we were able to account for unequal interval lengths while modeling bird availability for detection as a function of point- and/or year-specific covariates, x_{kt}, as $\text{logit}(q_{kt}) = \beta_0 + \sum_x \beta_x x_{kt}$. Thus, we were able to estimate effects of date (day-of-year) on availability, as a metric of whether *pa* varied over the season. However, note that our attempts to sample each point at the peak of territorial singing likely reduced our ability to observe effects date on *pa*.

To characterize effects of distance on perceptibility, pd_{kt}, we first dropped about 10% of the farthest (least accurate) observations of each species (per Kéry and Royle 2016) and then sorted this truncated detection-distance data into variable-width bins, equalizing the number of individual detections in each bin (per Amundson et al. 2014; we also explored other binning methods as discussed below). We modeled the probability that a bird was detected in distance bin b within the truncation radius of point k and year t, as $pd_{kt} = \sum_b \pi.d_{ktb}$, where $\pi.d_{ktb}$ was the probability of detection in bin b, a detection-distance function based on the half-normal distribution (Buckland et al. 2001, Farnsworth 2005, Royle et al. 2004, Amundson et al. 2014). The half-normal distribution has shape parameter σ_{kt}, the decay rate of detections by distance, allowing us to model detectability as a function of covariates, x_{kt} (e.g., tree cover, cover density, ambient noise level, group size), as $\log(\sigma_{kt}) = \log(\sigma_0) + \sum_x \beta_x x_{kt}$.

2.2 Modeling the count at each point

Level 2 involved relating each point- and year-specific count of the focal species, y_{kt}, to a specific population size, N_{kt}, through the processes of availability and perceptibility. The number of individuals available for detection was modeled as a random variable drawn from a binomial distribution defined by the actual population size and the proportion available for detection, $n_{kt} \sim \text{Binomial}(N_{kt}, pa_{kt})$. The subsequent count was modeled as a random variable drawn from a binomial distribution defined by the number of individuals available for detection and their detectability, $y_{kt} \sim \text{Binomial}(n_{kt}, pd_{kt})$.

2.3 Modeling the population across points and years

In level 3, we characterized spatial and temporal variation in focal species abundance by modeling the expected population size at point k and year t, λ_{kt}, as a log-linear function of point- and year-specific fixed and random effects, such as $\log(\lambda_{kt}) = park_{p[k]} + year_t + park_{p[k]} : year_t + transect_k + \sum_x \beta_x x_{kt}$. In this example, the intercept $park_{p[k]}$ could vary across the three mountain parks (indexed by p), random effect $year_t$ could interact with $park_{p[k]}$, and random effect $transect_k$ accounted for the spatial pattern of point-count stations, while covariates x_{kt} included year (for estimation of trend) and metrics of climate. Regional trends were estimated by replacing park-specific terms with an overall mean. The simplest (regional trend) model we considered included a single intercept, a linear effect of year, a fixed effect of year and a random effect of transect.

3.0 Application to NCCN landbird monitoring data

We used the new framework to estimate trends for 39 species, including nearly all species detected more than five times per average park-year. Of the species detected more than five times per average park-year, we excluded four species (Pine Siskin, Red Crossbill, Evening Grosbeak and Vaux's Swift); because these four were detected commonly in flocks, they did not conform to model assumptions. The 39 focal species exhibited a variety of nesting and dietary traits that might affect species response to climate, and included 15 "residents" that overwinter in the Pacific Northwest.

3.1 Landbird data preparation and review

Outputs from two queries of the NCCN landbird monitoring database were used for our analyses:

1) "`qs_a121_Distance_export_matrix...csv`" (abbreviated hereafter as `qs_a121`) contained primarily "survey" data consisting of (usually) one record for every bird observed in every point-year, and

2) "`qs_a204_Transect_coordinates_export...csv`" (`qs_a204`) contained point data consisting of one record for every point visited in each year.

The R script "`script4processingQueryData.R`" (Appendix 1) reads the .csv output from both queries and processes these data for use by a subsequent script that performs model fitting. Some of the survey and point data in these two queries overlap, and these overlapping data are processed separately (redundantly) so that we can write stand-alone copies of each query's processed output to file. Specific operations on the survey data (`qs_a121`) include: converting blanks in interval data Int_5_7 and Int_3_5 to zero after 2010 or to "NA" before 2011; changing "-1" in Int_3_5 and Int_5_7 to "1" ("-1" is the code for a positive detection in the database); making sure all blanks have been replaced with "NA" or known values, as appropriate (e.g., for all records of Species=NONE, blanks in Obs_distance_m are set to "NA"); extracting ordinal Day and Hour from Start_date and Start_time; creating "removal" data from interval data (Int_0_3, Int_3_5, Int_5_7) after 2010; handling records for which Nearest_obs = FALSE or zero (more on this below); and writing the processed data to a file for later use in model fitting (e.g., "`nccn.survey.data.2005to2014.csv`"). Specific operations on the point data (`qs_a204`) include: removing transect-origin points (where no counts occur); distributing "best" coordinates and GIS-derived point covariates for point *x* to each year of data for point *x*; transferring point covariates from `qs_a121` to the point data file and trimming redundant columns; distributing the Stratum_name for each point *x* to each year of data for point *x*; and writing the processed data to file for later use (e.g., "`nccn.site.data.2005to2014.csv` ").

The data processing script (Appendix 1) also includes many tools for data review to provide additional quality assurance prior to modeling. Several issues arose during the early application of this script to the data, suggesting the need for further data review and minor revisions to the nature of the data provided by the NCCN landbird monitoring database. For example, some exceptionally large slope values returned by one query led to the discovery that slope and aspect had been calculated incorrectly from elevation data; the calculation was corrected and the database recertified. Each of these issues was explored and exemplified using data from North Cascades National Park Complex (NOCA), and the results were reported in a document, “`DataExplorationNOCA_AllComments.docx`” which is linked directly to the NCCN landbird monitoring database.

Using the tools for data visualization provided in the data processing script (Appendix 1), we further explored the structure of these data by graphing raw counts and counts adjusted by sampling effort for each species against environmental and survey variables suspected as potential covariates. Examples featuring Clark’s Nutcracker (CLNU) appear in Appendix 2.

3.2 Climate data preparation and review

To assess effects of snowpack and other annual weather variables on landbird population density in the parks, we selected ClimateWNA as a source of downscaled covariates. ClimateWNA uses bilinear interpolation and local elevation adjustment to downscale monthly, gridded climate data as scale-free point data (Wang et al. 2016), providing climate metrics directly estimated for each point along our transects. Records from park SNOTEL sites and other local weather data are represented as inputs to the gridded climate product. By using a free and widely reviewed product like ClimateWNA, rather than raw local data, we have increased the accessibility and reproducibility of our analyses.

Outputs from two queries were used for our analyses of ClimateWNA data:

1) “`nccn.wna.pts_2004_2014.csv`” contained annual climate data for each point-year, and

2) “`nccn.wna.pts_Normal_1971_2000.csv`” contained average climate (“normals”) for each point during 1971-2000.

Annual data and 1971-2000 normals were accessed from http://www.climatewna.com using the protocol and code in Appendix 3. The “`script4climateWNAdata.R`” (Appendix 3) prepares query input files and processes query output for use in model fitting.

To characterize spring conditions, we selected mean spring temperature (MST, the average daily temperature from March 1 through May 31) and annual precipitation-as-snow (PAS, millimeters of snow falling between August 1 and July 31). We calculated MST and PAS as anomalies, relative to 1971-2000 normals, for use as predictors of population size, under the expectation that breeding and recruitment would be inversely related to snowfall and directly related to temperature. For surveys in year t, lag-1 MST was the mean temperature anomaly from March 1 to May 31 of year t-1, and lag-1 PAS was the snowfall anomaly from Aug 1 of year t-2 to July 31 of year t-1. We considered both additive and interaction effects of these anomalies and, because the correlation between MST and PAS was moderate (Kendall’s $\tau = -0.45$), we also considered a model based on PAS and “residual MST”, the residuals of a linear regression of MST anomalies on PAS anomalies (Graham 2003).

3.3 Summary of results for 39 landbird species

The R script “`script4trendAnalysis.R`” (Appendix 4) reads the .csv output files described above (from “`script4processingQueryData.R`” and “`script4climateWNAdata.R`”) and further formats the data for use in hierarchical models. Finally, the data are fitted to a model using the JAGS (Plummer 2003) interface package `jagsUI` (Kellner 2015).

Results are detailed in Ray et al. and outlined here. Fitting a model without park or climate effects, we estimated that all but one species were either stable or increasing across the sampled points in these three parks during 2005-2014 (Fig. 3 of Ray et al.). Fitting a park-structured trend model without effects of climate also suggested widespread stability or increase among these species (Fig. 4 of Ray et al.). There was also some evidence of spatial rather than temporal fluctuations over this period, with population decline in one park offset by increase in another (Fig. 2).

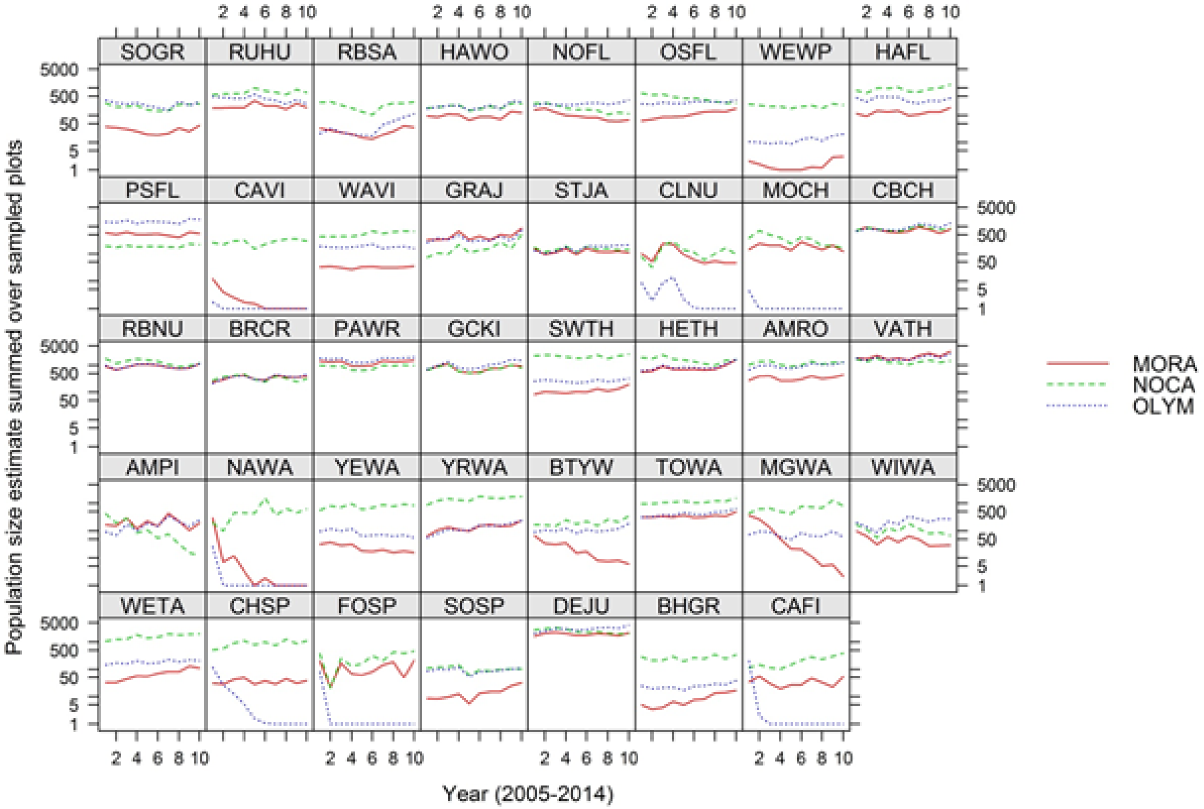


Figure 2. Annual population size estimates for 39 species. Years 2005-2014 are represented as 1-10. Population size, shown on a log scale to accentuate trends in small populations, is summed across the sampled plots in each park, including points not surveyed in every year (e.g., rotating panels). Spurious patterns of rapid decline from intermediate to low numbers (e.g., Nashville and MacGillivray’s Warbler in MORA, Chipping and Fox Sparrow in OLYM) can result when the model uses data from parks where the species is commonly detected to help estimate initial population size in parks where the species is detected only rarely. See text for solutions to this problem.

Appendix 5 summarizes fits to the park-structured model without climate covariates, and provides details on the results of fitting components of the detection model to survey covariates (observer, date, time), and of fitting abundance to site covariates in the database (fixed effects of park, fixed and random effects of year, park-by-year interactions, and random effects of transect). A potential effect was "supported" if its posterior 95% credible interval did not overlap zero, and model fit was "adequate" if the Bayesian p-value was not extreme (<0.1 or >0.9) and the Chi-square lack-of-fit (LOF) ratio was low (<1.1). Park-by-year interactions generally were not supported as predictors of abundance. Within parks, we found little support for effects of day and hour on availability, or for effects of observer on perceptibility. Note that due to logistical constraints in the field, observer and day are often confounded with park. Adding an observer effect at the regional (three-park) level provided little to no improvement in detection model fits.

In these preliminary models, abundance (*N*) was estimated only for the area around each point in which the species was relatively perceptible, which ranged from 30 m (for Rufous Hummingbird) to 250 m (Olive-sided Flycatcher). Park-structured estimates of annual variation in population size for each of 39 species suggested a variety of dynamics (Fig. 2). Of 117 park-specific trends (3 parks by 39 species), only 11 suggested pronounced declines, and these declines were not supported in posterior predictive checks. These declines were associated only with species detected less than twice per park-year, emphasizing the need for caution when fitting models to rare species. Park-structured models can overestimate the initial abundance of a species in parks where it was very rarely detected, by borrowing too heavily from estimates for parks where the species was more commonly detected. To avoid this issue, Ray et al. did not report the trend estimate for a species in any park where its average annual count was less than two. Spurious trend estimates for rare species can also be avoided by exploiting some of the flexibility of the new modeling framework, as discussed below.

An estimate of population size at the scale of each stratum was derived by fitting a stratum-specific model and extrapolating outside the sampling frame (recalling that our sampling frame was limited to ~1-2 km from trails). We did this to illustrate whether stratum-specific trends might combine with stratum-specific population densities to cause declines (or increases) that could be missed by focusing on trends within the surveyed plots. At least some differences in trend among elevational strata would be expected if climate change were forcing trends. However, we found that trends in population size were usually similar across elevational strata (Fig. 3). Supported trends contrasted among elevational strata for only three species (Olive-sided Flycatcher, Western Wood PeeWee and Mountain Chickadee), and any other apparent trends were not supported (e.g., Stellar's Jay, Song Sparrow and Black-headed Grosbeak in Fig. 3). In addition, regional trends based on stratum-specific fits and stratum-specific densities were similar to those based on only the set of surveyed plots (Fig. 4).

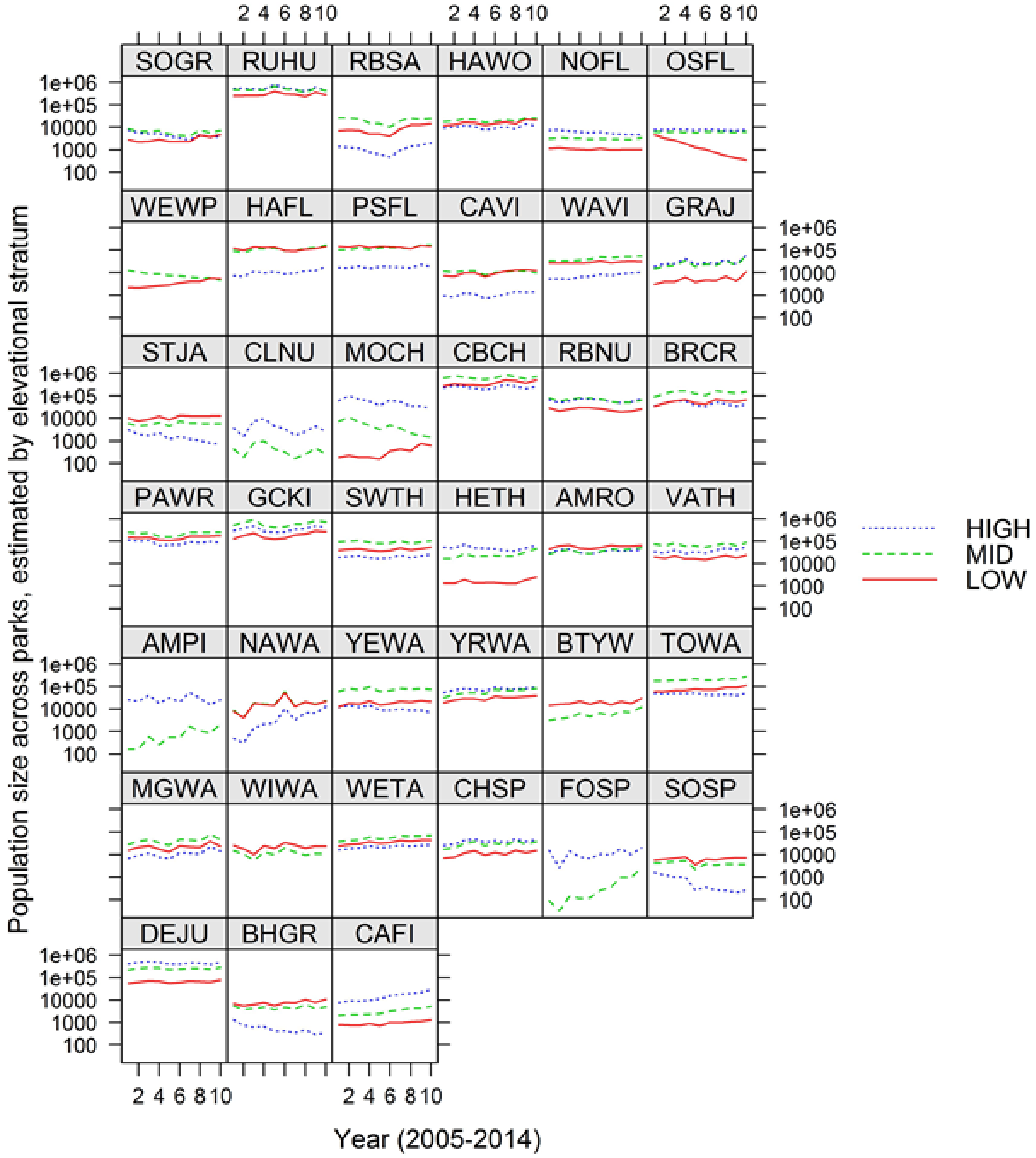


Figure 3. Annual population size estimates for 39 species, organized by elevational stratum (high = above 1350 m; low = below 800 m in MORA or 650 m in NOCA and OLYM). Years 2005-2014 are represented as 1-10. Population size, shown on a log scale to accentuate trends in small populations, was summed across the sampled plots in each stratum (including points not surveyed in every year) and extrapolated to the stratum scale based on sampling effort: (N across surveyed plots/total area surveyed)*total stratum area = N', where N' was the stratum-scale estimate. See text for justification of this approach, and Fig. 4 for calculation of regional population size.

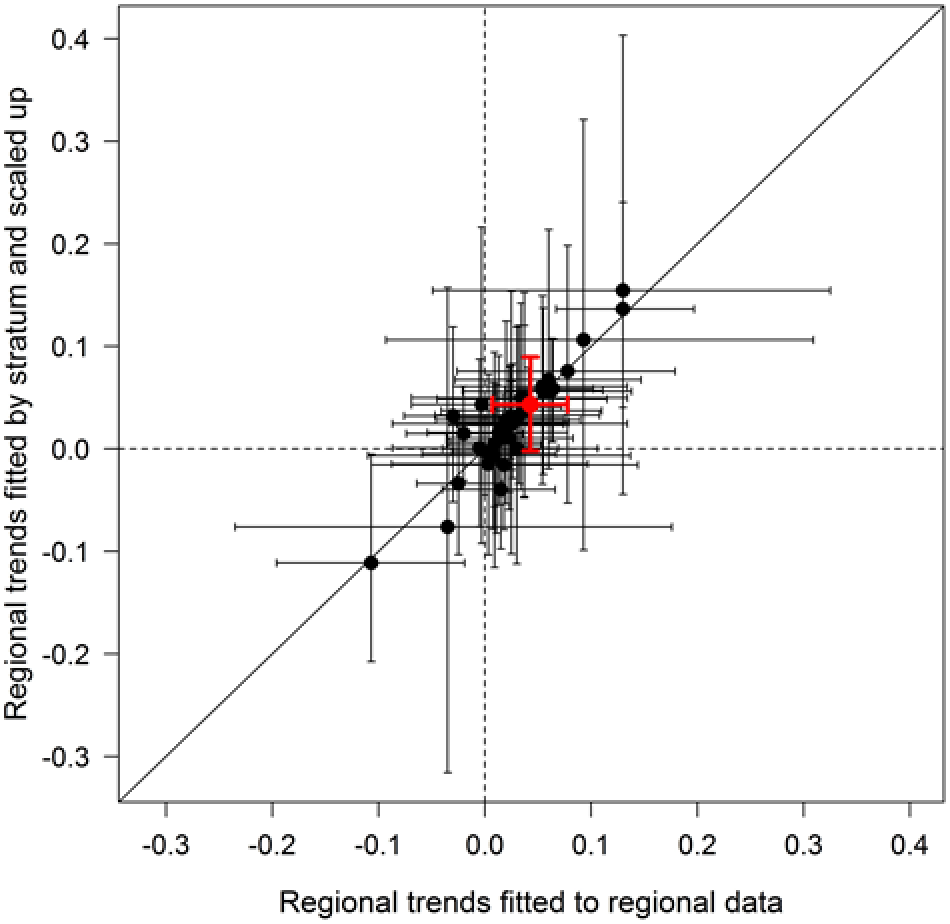


Figure 4. A comparison of regional trends in estimated population size from two analyses: 1) fitting a single, linear effect of year to all data, and not extrapolating trend beyond the surveyed plots (x axis); or 2) fitting stratum-specific trends followed by extrapolation to the stratum level (Fig. 3) and estimation of trend at the regional scale as a weighted average of stratum-level trends, using weights based on the geometric mean of annual population size within each stratum (y axis). Positive values (upper right-hand quadrant) represent trends that were positive in both analyses. Dots represent mean trends for each of the 39 focal species, and error bars represent 95% CIs. One trend in the upper right quadrant, shown in red, represents the only species (Warbling Vireo) for which the 95% CI crossed zero in one analysis but not the other. Bias due to the sampling design appears minimal.

4.0 Implications for future surveys and analysis

The new analytical framework for estimating landbird trends in the NCCN extends the model of Amundson et al. (2014) to accommodate count intervals of varying length (as in Farnsworth et al. 2002), and to estimate population trends from a time series of annual data. This framework has proven quite flexible, accommodating data from the existing point-count protocol as well as variation in the protocol over time (e.g., the retirement of several transects and the 2011 change from removal counts across two intervals to full counts across three intervals). Using Bayesian analysis and MCMC, all missing data are estimated during the modeling process, and we can also obtain a full distribution for each estimated value. This means, for example, that data from point-count stations visited only once can be incorporated into trend analyses, because any missing counts are estimated for every point ever visited using the observed pattern of data from other counts that are not missing.

We can evaluate parameter estimates at any level (from a missing point-count to a trend across point-years) using posterior predictive checks that can be as straightforward as comparing an observed response with the distribution of predicted responses deriving from the many iterations of the MCMC algorithm. At each iteration, model parameters are estimated anew, and a “perfect” dataset (including prediction uncertainty) can be sampled from each new fit to create the posterior predictive distribution of any fitted parameter. Using this approach, we found that relatively simple models were adequate for estimating population trends at park scales and regionally, as long as we detected an average of at least two individuals per park-year. Covariate effects were easily incorporated into submodels of detection and abundance. Based on these observations and further details provided below, continued application of the new modeling framework for population-specific trend analysis in the next synthesis of NCCN landbird monitoring data, as well as in syntheses of monitoring data from SIEN and other networks should provide a relatively robust alternative to the previously envisioned analytical framework of Siegel et al. (2007, 2010). Additional implications of specific survey and analysis questions are discussed below.

Rare species. Estimating trends within the three-park region required more than five detections per averaged park-year, and estimating park-specific trends required further that the distribution of the population support localized trend analysis by offering at least two detections per year in each park. Estimating stratum-specific trends for each elevation band (low, medium and high) will place further demands on the data, likely requiring at least two detections per stratum-year. We recommend caution when estimating trends for rare species. However, for species that are rare only in certain locations (see the examples mentioned in Fig. 2), trends can be estimated in at least three ways. First, it is possible within the new framework to constrain initial population estimates where the species is rare, to avoid inflated estimates of initial population size in these locations. Second, the Bayesian approach explored here can involve the use of an informative prior distribution for the trend parameter. For example, prior information indicating species absence from a park might be used to justify a more informative prior for the park effect. Third, trend analysis can be limited to data from the location where the species is rare, to avoid influence from locations where it is more common. These approaches might be useful when trends must be estimated for rare species targeted for management.

Flocking species. Species detected commonly in flocks do not conform to key assumptions of the models presented here, and we are not aware of related models appropriate for species with pronounced flocking behavior during the breeding period when point-counts are conducted. Only four flocking species were detected commonly enough in NCCN mountain parks to warrant trend estimation (Pine Siskin, Red Crossbill, Evening Grosbeak and Vaux's Swift). For these species, occupancy modeling for a spatially structured investigation of trends in presence, rather than abundance, would be a useful method for determining species response to habitat and climate. Occupancy modeling approaches are well developed and well suited to the data generated by the current NCCN landbird monitoring protocol (MacKenzie et al. 2006).

Covariates of detection. Detection was modeled in two parts, as in Amundson et al. (2014): availability was estimated via removal sampling (birds detected earlier are presumably more available for detection), and perceptibility was estimated via distance sampling (birds detected at greater distances are presumably more perceptible). Within parks, we found little support for effects of day and hour on availability, or for effects of observer on perceptibility. Observer and day were often confounded with park in the data analyzed, because logistical constraints dictate that parks must be visited sequentially or by different observers. Adding an observer effect at the regional level resulted in little improvement over an intercept-only model for most species. Similarly, binning detections by distance rather than by frequency had no appreciable effect on parameter estimates or model fit for most species. For the minority of species with poor fit to the detection models summarized here, further exploration of detection models based on habitat covariates, week and year would be useful, as well as exploration of different functional relationships between detection and distance (characterized here by the half-normal distribution, the most common approach in the literature). Reducing variation in distance estimates through observer training in the field should continue to be emphasized. Distance estimation in steep mountain terrain with complex multilayered canopies extending up to 70 meters high can produce highly variable outcomes (Buckland et al. 2001), especially when compared to ecosystems with less structural complexity, such as shrub-steppe.

Imperfect detection at zero distance. Although distance-sampling methods can reduce bias associated with imperfect detection, an important assumption of many distance-sampling models is perfect detection at zero distance. This assumption may not hold in avian point-count studies due to observer effects on bird behavior or because birds may be concealed in the forest canopy overhead; in our study, canopy layers can extend 75 m above the observer. Fortunately, the assumption of perfect detection at zero distance is obviated in models that incorporate information on both detection time and detection distance (Amundson et al. 2014). Collecting data on detection distance and time is straightforward for point counts, and is broadly encouraged by the research community (Twedt 2015).

Model assumptions. Assumptions that continue to require attention in *N*-mixture models include (1) random placement of survey points with respect to the population's distribution; (2) detection of individuals prior to any movement; (3) accurate estimation of distance to detection; (4) accurate identification of species and unique individuals; (5) independence between probabilities of individual availability and perceptibility; (6) closure, defined as lack of birth, death, immigration and emigration during each survey season; and (7) presence of the entire breeding population during each survey season. Observer training and careful design of survey protocols and analyses can help meet these assumptions (Amundson et al. 2014). For example, after censoring 10% of detections—those at the greatest distances (Kéry and Royle 2016)—we found little evidence of rounding error (clumping) in estimates of dispersal distance. Still, analysis of variance sometimes indicated significant dependence between distance- and time-to-detection, suggesting violation of (5). In these cases, we censored up to 27% (mean = 11%) of detections to avoid significant dependence between detection interval and distance. For several species, we explored the effect of censor rate on estimates of population density; in agreement with Amundson et al. (2014), we found little effect. However, assumptions (2), (6) and (7) may be more difficult to meet (Alldredge et al. 2007, Simons et al. 2009, Hoekman and Lindberg 2012, Schmidt et al. 2013, Mizel et al. 2017). The timing of arrival on the breeding grounds and the duration of peak singing rates can vary importantly across the breeding season (Mizel et al. 2017). Studies like ours, that include no more than one annual visit to each point, cannot directly estimate this source of variation. Although we found little support for an effect of date (day-of-year) on availability, trends in arrival date and availability during the breeding season can be confounded with trends in abundance in single-visit studies (Schmidt et al. 2013). Thus, it is conceivable that a decline in abundance could be countered and masked by a trend in arrival times that increased detectability during our point counts. We found no trends in our two metrics of breeding-season climate, but it will be important to investigate other potentially influential aspects of climate that these populations encounter throughout the year. Finally, the fact that we found little support for an effect of day on availability might be due to the fact that availability was generally quite high among our focal species, such that there was little remnant variance for covariates to explain. In contrast, our estimates of perceptibility were lower, and fit was poor in our sub-model of perceptibility for about ¼ of the species modeled, as discussed above. For similar analyses in the future, such as the 2017 synthesis report on population trends in the NCCN, we suggest further exploration of distance model assumptions, guided by previous studies that have characterized important departures from model assumptions (Simons et al. 2009, Hoekman and Lindberg 2012). Further exploration of covariates or of alternatives to the half-normal distribution commonly used to characterize the decline in detection with distance might also be fruitful. Alternative detection-distance functions might better capture any effects of distance on detection that might be peculiar to this study.

Missing data. Our time series provided a substantial test of whether the new framework can accommodate spatial and temporal variability in survey effort (missing data), an important component of many point-count surveys. Data were missing in each year for the vast majority of points in this dataset, primarily because only two out of every six panels are surveyed annually (Table 1 of Ray et al. in Appendix 1). However, as exemplified in Kéry and Royle (2016), missing data are conveniently estimated in this type of analysis. Parameter estimates for the more common species were surprisingly precise—e.g., credible intervals in Fig. 3 of Ray et al. were often quite narrow—despite the massive amounts of missing data modeled explicitly in our analyses.

If sampling effort were to be deliberately reduced relative to the current protocol, the credible interval for each parameter estimate would become wider. For some species, especially those that are relatively rare, credible intervals are already quite wide for certain parameter estimates (Figs. 4 and 6 of Ray et al. in Appendix 1). Hence, any reduction in sampling effort would affect the power of trend detection, and any reduction in the number of parks sampled would restrict formal inference to the remaining sampled parks. If it became necessary to remove a whole park from the monitoring protocol, retaining North Cascades National Park Complex (NOCA) could perhaps retain the most information for the network, because species trends in NOCA were most representative of trends across the three-park region (Fig. S1 of Ray et al.). If a reduction in annual effort were necessary, dropping whole years from the survey, rather than dropping selected transects within years, would maintain the integrity of the survey design and could avoid or minimize an increase in estimation error within the remaining survey years. However, dropping all point-counts from a given year could reduce our ability to detect effects of time-varying covariates like climate. In this case, instituting a rotating design for whole parks would allow for some continuity across the NCCN in the sampling of such covariates.

Covariates of abundance. We estimated effects of lagged climate on abundance to exemplify model construction based on covariates. Effects of habitat covariates such as tree density would be estimated in the same way. Anticipating declines in the Rufous Hummingbird and Olive-sided Flycatcher, we expected that a discussion of habitat covariates might be especially useful for these species. However, trends estimated for these and many other species were stable at the regional scale, with any park-scale declines offset by increases or at least stability in other parks (Fig. 4 of Ray et al. in Appendix 1). A notable exception was Clark's Nutcracker (CLNU), which appeared to have declined sharply in MORA without any appreciable offset elsewhere. CLNU habitat occurred primarily in MORA and NOCA at high elevations in open areas (Appendix 2, Figs. A2.1 and A2.5). In exploratory analyses featuring univariate models, we identified a negative, lag-1 effect of increasing precipitation-as-snow on CLNU counts from each point-year (Fig. A2.6), but that effect was not supported after accounting for trend, park and transect effects (Fig. 7 of Ray et al. in Appendix 1). Given the noted geographical and evolutionary association between CLNU and cone-bearing whitebark pines (Tomback 1982, Tomback and Linhart 1990, Lanner 1996, McKinney et al. 2009, Barringer et al. 2012), an obvious potential covariate of CLNU density would be the density of this primary food source. The collection and certification of park data on whitebark pines and other white pines that might serve as resources for CLNU is currently underway in both NCCN and SIEN (McKinney et al. 2012). Our preliminary modeling of climate effects on CLNU density demonstrate the potential for modeling CLNU density as a function of foraging habitat, using the new modeling framework to estimate trophic effects on landbird species dynamics. Estimating the specific effect of cone-bearing whitebark pine density on CLNU density in NCCN parks awaits certification of the pine dataset (J. Boetsch, pers. comm.), but a temporal model appears feasible given the upcoming availability of a time-series of whitebark pine surveys in 2004, 2009 and 2015-2016. Recalling the capacity for handling missing data in the new analytical framework, the gaps in temporal data should not be problematic. However, there may be a lack of spatial correspondence between CLNU and whitebark sampling locations, which suggests that data from the covariate and the response should be aggregated spatially prior to analysis, restricting inference regarding the spatial pattern of habitat use by CLNU.

Inference at park and regional levels. Trend estimates can be biased by variations in total area among parks and elevational strata. Although effects of park, transect and year were modeled, and sampling was relatively intensive and well-distributed among parks and elevations each year (see Table 1 in Ray et al.), parameter estimates should also account for variations in sampling density. For the 2017 synthesis report, weighting stratum-level parameter estimates by stratum area (Link and Sauer 2002, 2011) will minimize potential bias in trends at the park level. Similarly, regional trends can be estimated from park-level trends weighted by park area. Accounting for variation in sampling density among elevational strata and parks will be important for making specific inferences regarding park resources and regional trends.

Bias and precision. Bias and precision in parameter estimates from point-count studies are influenced by joint effects of mean species abundance, detection probability, the number of points surveyed and the number of counts within a season at a point (here, the number of count intervals during the single annual visit to a point). Lower values for any of these quantities lead to higher bias and lower precision (Yamaura et al. 2016). As in many avian point-count studies, both detectability and mean population density were low for many species in our study. However, the number of points surveyed and the effective number of counts per point in our study were high enough to ensure low bias and high precision in simulations by Yamaura et al. (2016). Furthermore, our estimates of population density, climate effects, park effects, year effects, and the parameters controlling components of detection were robust to variations in model structure to account for regional or park-specific trends and effects of climate. These patterns should be revisited in the 2017 synthesis report, when data on variation in sampling density among parks and elevational strata will be used to estimate stratum-specific and park-wide parameters.

Bayesian meta-analysis of mean parameter estimates by species group. The analytical framework developed here results in fitted models of population dynamics for individual species. Understanding broader ecological patterns and managing multi-species groups will require aggregating results across species. Mean parameter estimates across species or trait groups can be quantified using Bayesian meta-analysis (*sensu* Kéry and Royle 2016, pp. 680-683), a two-step hierarchical analysis to account for uncertainty in parameter estimates for individual species (Sauer and Link 2002). Example code for a meta-analysis of the average response to mean spring temperature, grouped by nesting habit, is provided in Appendix 6. The need for meta-analysis to estimate parameters by species trait-group is evident in Fig. 5, which contrasts two ways of summarizing the (estimated) effect of temperature on species groups: 1) boxplots of point estimates for each species, and 2) means (filled points) and 95% CIs (shaded areas) for each group as estimated by Bayesian meta-analysis. The boxplots suggest that mean spring temperature has a negative impact only on ground-nesting species. In contrast, results from the meta-analysis suggest that mean spring temperature has no impact on ground-nesters, but has a positive impact on species that do not nest on the ground. Bayesian hierarchical methods provide a convenient analytical framework for propagating uncertainty among levels of analysis. Thus, the credible intervals generated by Bayesian meta-analysis in Fig. 5 provide more appropriate statistical support for inference concerning the effect of MST. However, we must again caution that the nature of the effect of MST is not entirely clear from data based on single-visit studies; MST might affect the actual abundance of non-ground-nesting birds, or it might affect the timing of their arrival (and availability for detection) on the breeding grounds (see examples in Mizel et al. 2017).

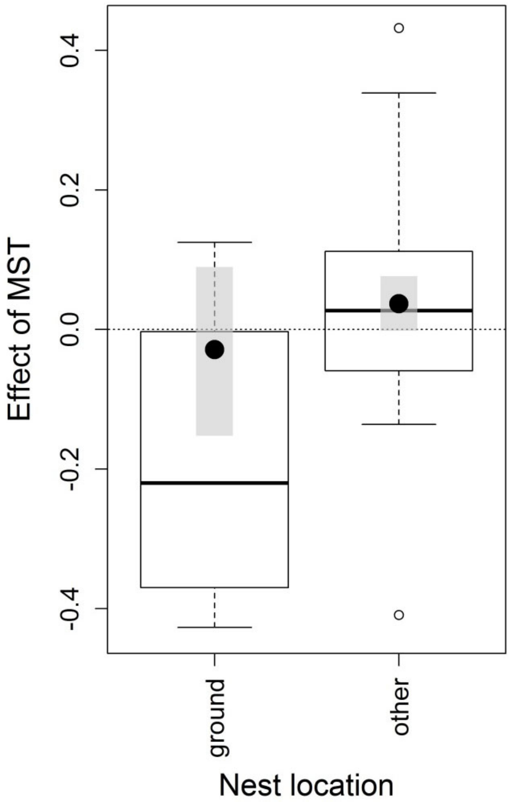


Figure 5. Parameter estimates grouped by nesting habit. Boxplots summarize point estimates of the effect of mean spring temperature (MST) on each species. Filled points (means) and shaded areas (95% CIs) summarize the average response to MST across species according to a Bayesian meta-analysis, which propagates the uncertainty in point estimates to higher level statistics such as species-group means.

Summary

Employing methods in hierarchical modeling of population and community processes will improve our ability to extract meaningful parameter estimates from diverse datasets (Tingley and Beissinger 2013). Combining data on detection distance and time increases the flexibility and utility of models designed to estimate abundance (Amundson et al. 2014). Given within-season replication of counts, as in the NCCN landbird monitoring protocol and others, this analytical framework can be extended to model apparent survival, recruitment and turnover (Dail and Madsen 2011, 2013) to support diverse population management goals.

References

Alldredge, M. W. 2004. Avian point-count surveys: Estimating components of the detection process. Ph.D. dissertation, North Carolina State University, Raleigh.

Alldredge, M. W., K.H. Pollock, T. R. Simons, J. A. Collazo, and S. A. Shriner. 2007. Time-of-detection method for estimating abundance from point-count surveys. Auk 124:653.

Amundson, C. L., J. A. Royle, and C. M. Handel. 2014. A hierarchical model combining distance sampling and time removal to estimate detection probability during avian point counts. 131:476–494.

Barringer, L. E., D. F. Tomback, M. B. Wunder, and S. T. McKinney. 2012. Whitebark pine stand condition, tree abundance, and cone production as predictors of visitation by Clark’s nutcracker. PLoS ONE 7 (5): e37663.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to distance sampling. Oxford University Press, Oxford, United Kingdom.

Dail, D., and M. Madsen. 2011. Models for estimating abundance from repeated counts of an open metapopulation. Biometrics 67:577-587.

Dail, D., and M. Madsen. 2013. Estimating open population site occupancy from presence–absence data lacking the robust design. Biometrics 69:146-156.

Duarte, A., J. S. Hatfield, T. M. Swannack, M. R. J. Forstner, M. C. Green, and F. W. Weckerly. 2016. Simulating range-wide population and breeding habitat dynamics for an endangered woodland warbler in the face of uncertainty. Ecological Modelling 320:52–61.

Farnsworth, G. L., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines, and J. R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. Auk 119:414–425.

Farnsworth, G. L., J. D. Nichols, J. R. Sauer, S. G. Fancy, K. H. Pollock, S. A. Shriner, and T. R. Simons. 2005. Statistical approaches to the analysis of point count data: a little extra information can go a long way. Pages 736–743 *in* C. J. Ralph and T. D. Rich, Editors. Bird Conservation Implementation and Integration in the Americas: Proceedings of the Third International Partners in Flight Conference. USDA Forest Service General Technical Report PSW-GTR-191.

Graham, M. L. 2003. Confronting multicollinearity in ecological multiple regression. Ecology 84:2809-2815.

Hoekman, S. T., and M. S. Lindberg. 2012. Point transect sampling for monitoring passerine birds in Denali National Park and Preserve: an assessment of 2002–2008 pilot data. Natural Resource Technical Report NPS/CAKN/NRTR-2012/589. National Park Service. Fort Collins, Colorado. Published Report-2185664.

Kellner, K. 2015. jagsUI: a wrapper around 'rjags' to streamline 'JAGS' analyses. R package version 1.3.7. Available at: http://CRAN.R-project.org/package=jagsUI

Kéry, M., and J. A. Royle, editors. 2016. Applied hierarchical modeling in ecology: analysis of distribution, abundance and species richness in R and BUGS. Volume 1, Prelude and static models. Elsevier, Amsterdam and Boston. xxiv, 783 pp.

Kéry, M., and M. Schaub. 2012. Bayesian population analysis using WinBUGS: a hierarchical perspective. Academic Press, 535 pp.

Lanner, R. M. 1996. Made for each other: A symbiosis of birds and pines. Oxford, United Kingdom: Oxford Univ. Press.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Academic Press, San Diego, California.

McKinney, S. T., C. E. Fiedler, and D. F. Tomback. 2009. Invasive pathogen threatens bird-pine mutualism: Implications for sustaining a high-elevation ecosystem. Ecological Applications 19:597–607.

McKinney, S.T., T. Rodhouse, L. Chow, A. Chung-MacCoubrey, G. Dicus, L. Garrett, K. Irvine, S. Mohren, D. Odion, D. Sarr, and L. A. Starcevich. 2012. Monitoring white pine (*Pinus albicaulis*, *P. balfouriana*, *P. flexilis*) community dynamics in the Pacific West Region - Klamath, Sierra Nevada, and Upper Columbia Basin Networks: Narrative version 1.0. Natural Resource Report NPS/PWR/NRR—2012/532. National Park Service, Fort Collins, Colorado.

Mizel, J. D., J. H. Schmidt, C. L. McIntyre, and M. S. Lindberg. 2017. Subarctic-breeding passerines exhibit phenological resilience to extreme spring conditions. Ecosphere 8(2):e01680. 10.1002/ecs2.1680

Plummer, M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Pages 1-10 in Hornik, K., F. Leisch, and A. Zeileis (eds.), Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DCS2003), March 20–22, Technische Universität, Vienna, Austria.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/.

Ray, C., J. F. Saracco, A. L. Holmgren, R. L. Wilkerson, R. B. Siegel, K. J. Jenkins, J. I. Ransom, P. J. Happe, J. R. Boetsch, and M. H. Huff. *In revision*. Recent stability of resident and migratory landbird populations in National Parks of the Pacific Northwest. *Ecosphere*.

Royle, J.A. 2004. *N*-mixture models for estimating population size from spatially replicated counts. Biometrics 60:08–115.

Royle, J. A., D. K. Dawson, and S. Bates. 2004. Modeling abundance effects in distance sampling. Ecology 85:1591–1597.

Sauer, J. R. and W. A. Link. 2002. Hierarchical modeling of population stability and species group attributes from survey data. Ecology 86:1743–1751.

Sauer, J. R. and W. A. Link. 2011. Analysis of the North American breeding bird survey using hierarchical models. Auk 128:87–98.

Schmidt, J. H., C. L. McIntyre, and M. C. MacCluskie. 2013. Accounting for incomplete detection: what are we estimating and how might it affect long-term passerine monitoring programs? Biological Conservation 160:130–139.

Siegel, R. B., R. L. Wilkerson, K. J. Jenkins, R. C. Kuntz II, J. R. Boetsch, J. P. Schaberl, and P. J. Happe. 2007. Landbird monitoring protocol for National Parks in the North Coast and Cascades Network. U.S. Geological Survey Techniques and Methods 2-A6. 208 pp.

Siegel, R. B., R. L. Wilkerson, and M. Goldin Rose. 2010. Bird monitoring protocol for national parks in the Sierra Nevada Network: Fort Collins, Colo., U.S. National Park Service, Natural Resource Report NPS/SIEN/NRR–2010/231.

Silsbee, G. G., and D. L. Peterson. 1991. Designing and implementing comprehensive long-term inventory and monitoring programs for National Park System lands. Natural Resources Report NPS/NRUW/NRR-91/04, Denver, CO.

Simons, T. R., K. H. Pollock, J. M. Wettroth, M. W. Alldredge, K. Pacifici, and J. Brewster. 2009. Sources of measurement error, misclassification error, and bias in auditory avian point count data. Pages 237–254 in D. L. Thompson, E. G. Cooch, M. J. Conroy (Eds.), Modeling demographic processes in marked populations. Springer, New York, NY, USA, pp. 237–254.

Tingley, M. W., and S. R. Beissinger. 2013. Cryptic loss of montane avian richness and high community turnover over 100 years. Ecology 94:598-609.

Tomback, D. F. 1982. Dispersal of whitebark pine seeds by Clark’s nutcracker: A mutualism hypothesis. Journal of Animal Ecology 51 (2): 451–67.

Tomback, D.F., and Y.B. Linhart. 1990. The evolution of bird-dispersed pines. Evolutionary Ecology 4:185–219.

Twedt, D. J. 2015. Estimating regional landbird populations from enhanced North American Breeding Bird Surveys. Journal of Field Ornithology 86:352–368.

Wang, T., Hamann, A. Spittlehouse, D., and C. Carroll. 2016. Locally downscaled and spatially customizable climate data for historical and future periods for North America. PLoS ONE 11(6):e0156720. doi:10.1371/journal.pone.0156720

Yamaura, Y., M. Kéry, and J. A. Royle. 2016. Study of biological communities subject to imperfect detection: bias and precision of community N-mixture abundance models in small-sample situations. Ecological Research 31:289–305.

Appendix 1: R script for processing data queried from the NCCN landbird monitoring database.

```
# Process query data for use in hierarchical models

# Code file name “script4processingQueryData.R” available from C. Ray
(cray@birdpop.org)

# Reads the .csv output from two queries, "qs_a121..." (containing survey data) and
# "qs_a204..." (containing point data); addresses issues in these data (see below);
# writes clean/corrected survey and point data to .csv files

# Some of the survey and point data overlap and are processed twice, just so that we
# can write clean, separate copies of survey and point data to file

# Specifically:
# Format survey data derived from query qs_a121... w/one record for every yr/pt/sp:
# - convert blanks in Int_5_7 and Int_3_5 to "0" AFTER 2010, "NA" BEFORE 2011
# - change "-1"s in Int_3_5 and Int_5_7 to "1"s (-1 was db code for sp detection)
# - for all records w/Species=NONE, set blanks in Obs_distance_m to "NA"
# - for transect 1057 (replaced 1030), set Panel_type to "Alternating" and
#     set Panel_name to "3"
# - for transect 3131 (retired), set Panel_type to "Annual" and Panel_name to "1"
# - for Start_date, translate the formatted date (5/25/2005 0:00) into ordinal Day
# - for Start_time, translate the formatted time (12/30/1899 7:57:00) into Hour
# - create removal data from interval data (Int_0_3, Int_3_5, Int_5_7)
# - for Habitat_num=blank, make PMR_type_desc, Detection_class & Is_forested "NA"
# - delete the Flyover column (contains no info--query qs_a121 specifies Flyover=0)
# - remove leading zeros on some pts in q1$Location_code
# - handle records for which Nearest_obs=FALSE/0
# - write nice survey data to .csv file
# Format point data (data derived from query qs_a204...one record for every pt/yr):
# - remove transect-origin points
# - distribute "best" coords and GIS covars for pt "x" to each yr of data for pt "x"
# - for transect 1057 (replaced 1030), set Panel_type to "Alternating" and
#     set Panel_name to “3”
# - for transect 3131 (retired), set Panel_type to "Annual" and Panel_name to "1"
# - transfer pt covars from qs_a121 to the point data file and trim redundant columns
# - distribute "Stratum_name" for pt "x" to each yr of data for pt "x"
# - write nice point data to .csv file
# - merge survey data with point data; preserves pts lost when removing Nearest_obs=F
# - write nice merged file to .csv
# - re-focus survey data onto the focal species (e.g., set Group_size=0 for other spp)

require(dplyr)

#***SEARCH FOR "***" TO FIND ANY LINES THAT MUST BE CUSTOMIZED FOR SPECIFIC ANALYSES

# Format survey data derived from query qs_a121... w/one record for every yr/pt/sp:

NAs <- c(""," ","."," ","..","na","NA","-")

#*** update file name as needed
q1 <- read.csv("qs_a121_Distance_export_matrix_20151023_173639.csv",header=T,
                 na.strings=NAs)
names(q1); dim(q1)

# - convert blanks in Int_5_7 and Int_3_5 to "0" AFTER 2010, "NA" BEFORE 2011

#there are blanks for every non-detection in the 2-interval, removal-sampling before
#2011, and blanks for Int_3_5 and Int_5_7 whenever Species=NONE after 2010; so...
```

```
i <- which(q1$Event_year<2011); length(i) #checking numbers, throughout
q1[i,"Int_5_7"] <- NA #because Int_5_7 wasn't implemented until 2011
i <- which(q1$Event_year>2010&(is.na(q1$Int_5_7)|is.na(q1$Int_3_5))); length(i)
j <- which(names(q1)=="Int_5_7"|names(q1)=="Int_3_5"); j
q1[i,j] <- 0 #next line just checks whether Int_3_5 NAs < Int_5_7 NAs, as expected
i <- which(is.na(q1$Int_3_5&q1$Event_year<2011)); length(i)

# - change "-1"s in Int_3_5 and Int_5_7 to "1"s (-1 was db code for sp detection)

i <- which(q1[,"Int_3_5"]==(-1)); length(i)
q1[i,"Int_3_5"] <- 1
i <- which(q1[,"Int_5_7"]==(-1)); length(i) #Int_5_7 -1s < Int_3_5 -1s, as expected
q1[i,"Int_5_7"] <- 1

# - for all records w/Species=NONE, set blanks in Obs_distance_m to "NA"

i <- which(q1[,"Species"]=="NONE"); length(i)
q1[i,"Obs_distance_m"] #all should be "NA"
sum(is.na(q1$Obs_distance_m)) #should be >= i

# - for transect 1057 (replaced 1030), set Panel_type...and Panel_name...

i <- which(q1$Site_code==1057); i #checking that this is a coherent group of records
q1[i,"Panel_type"] <- "Alternating"
q1[i,"Panel_name"] <- 3 #because transect 1057 is in panel 3
q1[i,c("Site_code","Location_code","Event_year","Panel_type","Panel_name")]

# - for transect 3131 (retired), set Panel_type to "Annual" and Panel_name to "1"

i <- which(q1$Site_code==3131); i
q1[i,"Panel_type"] <- "Annual"
q1[i,"Panel_name"] <- 1
q1[i,c("Site_code","Location_code","Event_year","Panel_type","Panel_name")]

# - for Start_date, translate the formatted date (5/25/2005 0:00) into ordinal Day

odate <- strptime(q1$Start_date,format="%m/%d/%Y %H:%M",tz="America/Los_Angeles");
head(odate)
rdate <- strptime(odate,format="%Y-%m-%d",tz="America/Los_Angeles"); head(rdate)
#check readout
rdate <- as.POSIXlt(odate) #put record date into a useful format
hist(rdate$yday) #check distribution of ordinal/Julian date
Day <- rdate$yday #bind to q1
q1$Start_date <- as.character(rdate); head(q1$Start_date)
q1 <- data.frame(q1,Day); names(q1)

# - for Start_time, translate the formatted time (12/30/1899 7:57:00) into Hour

otime <- strptime(q1$Start_time,format="%m/%d/%Y %H:%M:%S",tz="America/Los_Angeles");
head(otime)
#the default date gives us the wrong time-zone (daylight savings), but that won't
matter
#after we standardize covars
rtime <- strptime(otime,format="%Y-%m-%d %H:%M:%S",tz="America/Los_Angeles");
head(rtime)
rtime <- as.POSIXlt(otime,tz="America/Los_Angeles")
hist(rtime$hour) #check distribution of hour
Hour <- rtime$hour #bind to q1
q1$Start_time <- sub("1899-12-30 ","",rtime)
q1 <- data.frame(q1,Hour); names(q1)

# - for Habitat_num=blank, make PMR_type_desc, Detection_class & Is_forested "NA"
```

```
i <- which(is.na(q1$Habitat_num)); length(i)
j <- which(is.na(q1$PMR_type_desc)); length(j) #if i & j are equal, good
j <- which(is.na(q1$Is_forested)); length(j) #if length(i) > length(j), bad
#Currently, Is_forested defaults to 0 instead of blank/NA in the db
if (length(i) > length(j)) q1[i,"Is_forested"] <- NA
q1[i,c("PMR_type_desc","Detection_class","Is_forested","Habitat_num")] #view context

# - delete the Flyover column (contains no info--query qs_a121 specifies Flyover=0)

j <- which(names(q1)=="Flyover"); j
q1 <- q1[,-j]; names(q1)

# - handle records for which Nearest_obs=FALSE/0

i <- which(q1$Nearest_obs==F); length(i); dim(q1) #~3% of records are Nearest_obs=0
q1$Species[i] <- NA #record is retained but will never be counted as a focal sp
#avoid omitting records; creates problems if Nearest_obs=0 for the only
#bird(s) detected
j <- which(names(q1)=="Nearest_obs"); j #remove Nearest_obs from data frame
q1 <- q1[,-j]; dim(q1)

# - create removal data from interval data (Int_0_3, Int_3_5, Int_5_7)

get.first <- function(x) min(which(x!=0),na.rm=T) #finds first detection interval
                                                  #for record x
j <- which(names(q1)=="Int_0_3"); j #1st of 3 cols of interval data
y <- as.matrix(q1[,j:(j+2)]); levels(as.factor(y)) #interval data = 0, 1 or NA
f <- apply(y,1,get.first) #ignore warnings; create vector f, first occasions
                          #of detection
yi <- matrix(0,nrow=dim(y)[1],ncol=dim(y)[2]) #use f to recreate y
for (i in 1:dim(y)[1]) { if (q1$Group_size[i] > 0) yi[i,f[i]] <- 1 }; head(yi)
yi[q1$Event_year<2011,3] <- NA #3rd interval not implemented until 2011
colnames(yi) <- c("y1","y2","y3")
head(yi); tail(yi); sum(yi[,1:3],na.rm=T)
q1 <- data.frame(q1,yi); names(q1)

# - remove leading zeros on some pts in q1$Location_code

q1[,"Location_code"] <- sub("^[0]+", "", q1[,"Location_code"])

# - write nice survey data to .csv file

#*** update file name as needed
write.csv(q1,file="nccn.survey.data.2005to2014.csv")


# Format point data derived from query qs_a204...w/one record for every pt/yr:

#*** update file name as needed
q2 <- read.csv("qs_a204_Transect_coordinates_export_20151117_164321.csv",header=T)

# - remove transect-origin points

i <- which(q2$Location_code=="TO"); length(i) #~years*x, where x is the number of TOs
q2 <- q2[-i,]; dim(q2) #~years*y, where y is the number of pts, many alternating

# - distribute "best" coords and GIS-derived covars for pt "x" to each yr of data for
pt "x"

#i.e., replace coords collected annually, and covars derived from coords annually,
#with "best" coords, elev, slope and aspect
names(q2) #look at pt covars to see which covars to keep
q2.trim <- q2[,c(1:8)] #keep core data
```

```
ptbest <- q2[q2$Is_best==1,c(5,6,13,15:19)] #keep best coords for ea pt
dim(q2.trim); names(q2.trim); dim(ptbest); names(ptbest)
q2 <- merge(q2.trim,ptbest,all=T); dim(q2) #distrib best to ea pt-yr

# - for transects 1057 & 3131, set appropriate Panel_type and Panel_name

j <- which(names(q2)=="Panel_type"); j
q2[q2$Site_code==1057,j] <- "Alternating"
q2[q2$Site_code==3131,j] <- "Annual"
j <- which(names(q2)=="Panel_name"); j
q2[q2$Site_code==1057,j] <- 3
q2[q2$Site_code==3131,j] <- 1

# - transfer pt covars from qs_a121 to the point data file and trim redundant columns

#divvy q1 data between pt data and survey data
#note: for the Amundson-type model I ended up merging these back together
names(q2); names(q1) #determine which q1 columns should be in q2
pt.d <- unique(q1[,c(1,4,6,7,12:15)]); names(pt.d); dim(pt.d) #store pt data from q1
sv.d <- q1[,c(4,6,7,11,16:25)]; names(sv.d); dim(sv.d) #store survey data from q1

#verify that all pt-yr records in q2 have counterparts in pt.d (q1)
names(q2); dim(q2); names(pt.d); dim(pt.d) #if dims aren't equal, fix

#before looking at which rows in q2 don't occur in pt.d, note that q2$Location_code
#includes leading zeros on some pts, while pt.d$Location_code USUALLY doesn't,
#and fix as follows:
q2[,"Location_code"] <- sub("^[0]+", "", q2[,"Location_code"]) #fixes most
pt.d[,"Location_code"] <- sub("^[0]+", "", pt.d[,"Location_code"]) #fixes more
mismatch <- anti_join(q2,pt.d); dim(mismatch) #rows of q2/qs_a204 not in pt.d/qs_a121
head(mismatch[order(mismatch$Site_code),])
o.mis <- mismatch[order(mismatch$Site_code),]
o.mis[7:12,] #work through examples to understand patterns
o.mis[13:18,]
tail(o.mis)

#pattern: if a pt-yr was missing from qs_a101 but occurred in qs_a204 w/yr-specific
#coords (suggesting it was visited), maybe snow or noise or something precluded the
#survey; this issue is addressed by merging q2 with pt.d
mismatch2 <- anti_join(pt.d,q2); dim(mismatch2) #checking for rows of pt.d not in q2

#merge pt data from pt and survey data files
siteData <- left_join(q2,pt.d,by=c("Event_year","Site_code","Location_code"))
head(siteData); dim(siteData)
nas <- apply(siteData,2,function(x) sum(is.na(x))); nas #checking number of NAs
sum(siteData$Park_code=="LEWI"|siteData$Park_code=="SJHS") #explaining Panel_name NAs

#because NAs in Stratum_name are a special problem, check:
siteData[is.na(siteData$Stratum_name),1:8] #look at the records w/NA in Stratum_name
nas <- apply(q2,2,function(x) sum(is.na(x))); nas #NAs from pt data...
nas <- apply(pt.d,2,function(x) sum(is.na(x))); nas #NAs from survey data

# - distribute "Stratum_name" for pt "x" to each yr of data for pt "x"

a <- q1[!is.na(q1$Stratum_name),c("Stratum_name","Site_code","Location_code")]
a <- unique(a); head(a); dim(a) #Stratum_name associated w/each pt;
                                #dim(ptbest) may differ!
a[,"Location_code"] <- sub("^[0]+", "",a[,"Location_code"]) #omits leading zeros
a <- left_join(siteData,a,by=c("Site_code","Location_code")); names(a)
sum(is.na(a$Stratum_name.x)); sum(is.na(a$Stratum_name.y))
a[is.na(a$Stratum_name.y),] #ptbest w/o Stratum_name

#if we're still missing Stratum_name for some pts, fix as follows: pts w/in a
```

```
#transect share a stratum, so find those other pts & assign shared Stratum_name
for (i in which(is.na(a$Stratum_name.y))) {
  j <- which(a$Site_code==a$Site_code[i]) #indexes all pts w/in this transect
  a[i,"Stratum_name.y"] <- levels(droplevels(a[j,"Stratum_name.y"])) #shared name
}
sum(is.na(a$Stratum_name.x)); sum(is.na(a$Stratum_name.y)) #check that NAs are gone
siteData <- a[,-which(names(a)=="Stratum_name.x")]; names(siteData) #restore df
names(siteData)[dim(siteData)[2]] <- "Stratum_name"; names(siteData)

# - write nice point data to .csv file

#*** update file name as needed
write.csv(siteData,file="nccn.site.data.2005to2014.csv")
```

Appendix 2: Exploration of Clark's Nutcracker (CLNU) data from three mountain parks in the NCCN: Mount Rainier (MORA), North Cascades (NOCA) and Olympic (OLYM).

Raw and effort-adjusted counts were related to potential covariates prior to modeling detection-adjusted trends.

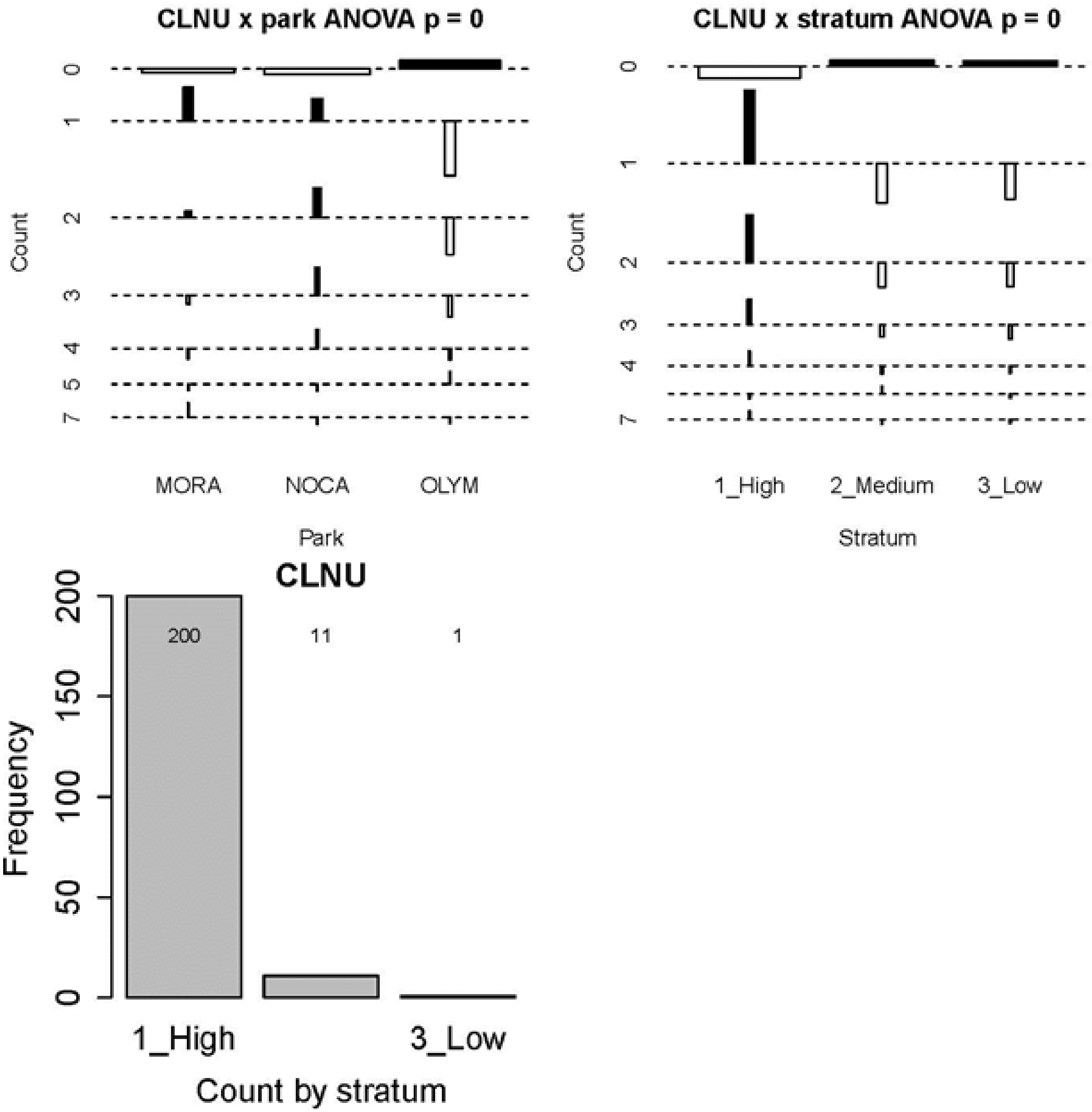


Figure A2.1. Association plots showing raw counts by park (upper left) and by elevational stratum across parks (upper right and lower left). Upper plot titles report the significance (p-value at α = 0.05 by ANOVA) of variance in raw counts explained by each predictor variable.

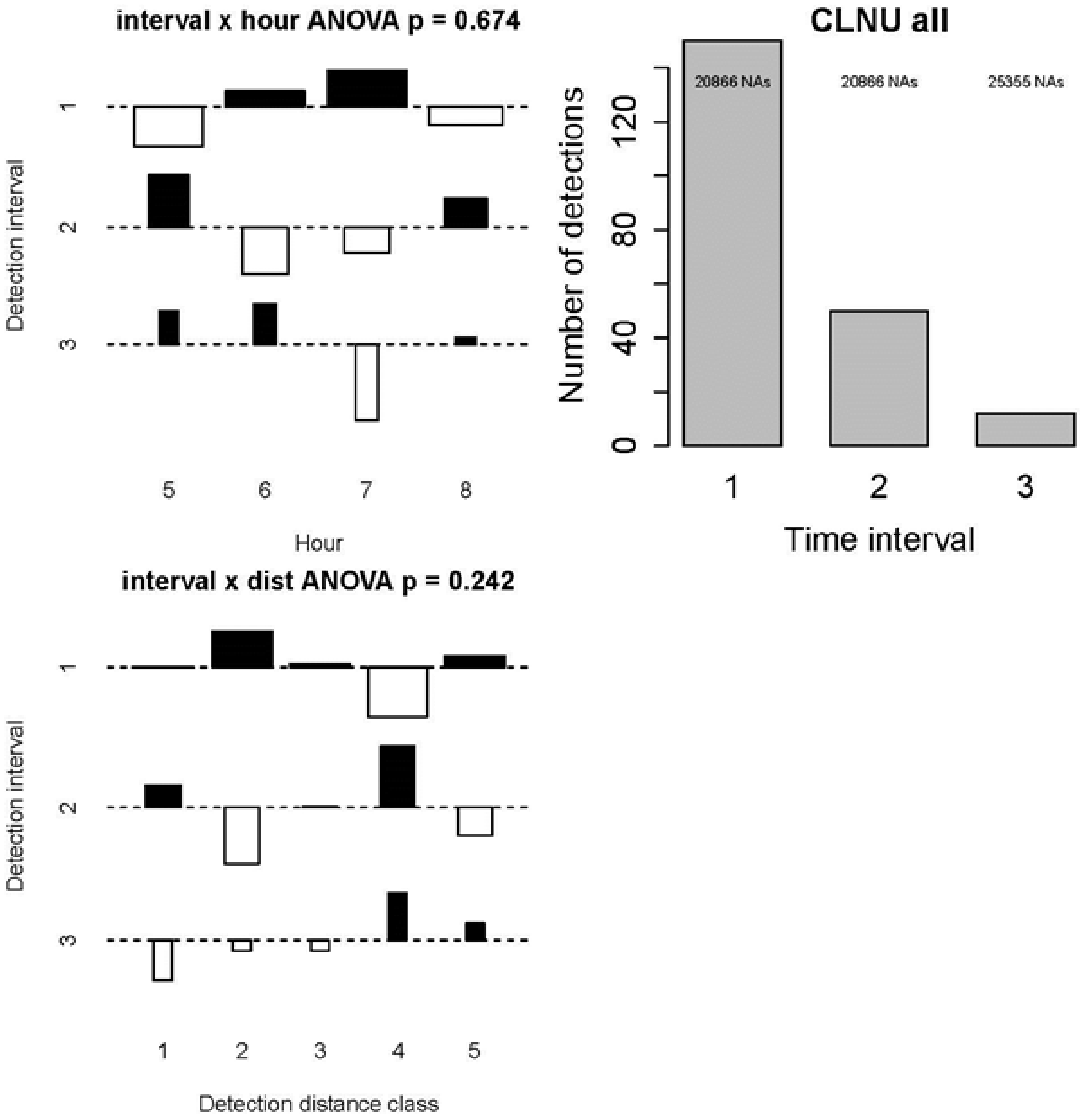


Figure A2.2. Association plots relating CLNU point-level detection interval (based on time-of-detection) to hour of day (upper left) and detection distance (lower left), where plot titles report the significance of variance in raw counts explained by each predictor variable (p-value at α = 0.05 by ANOVA). Detections per time interval are also shown (upper right).

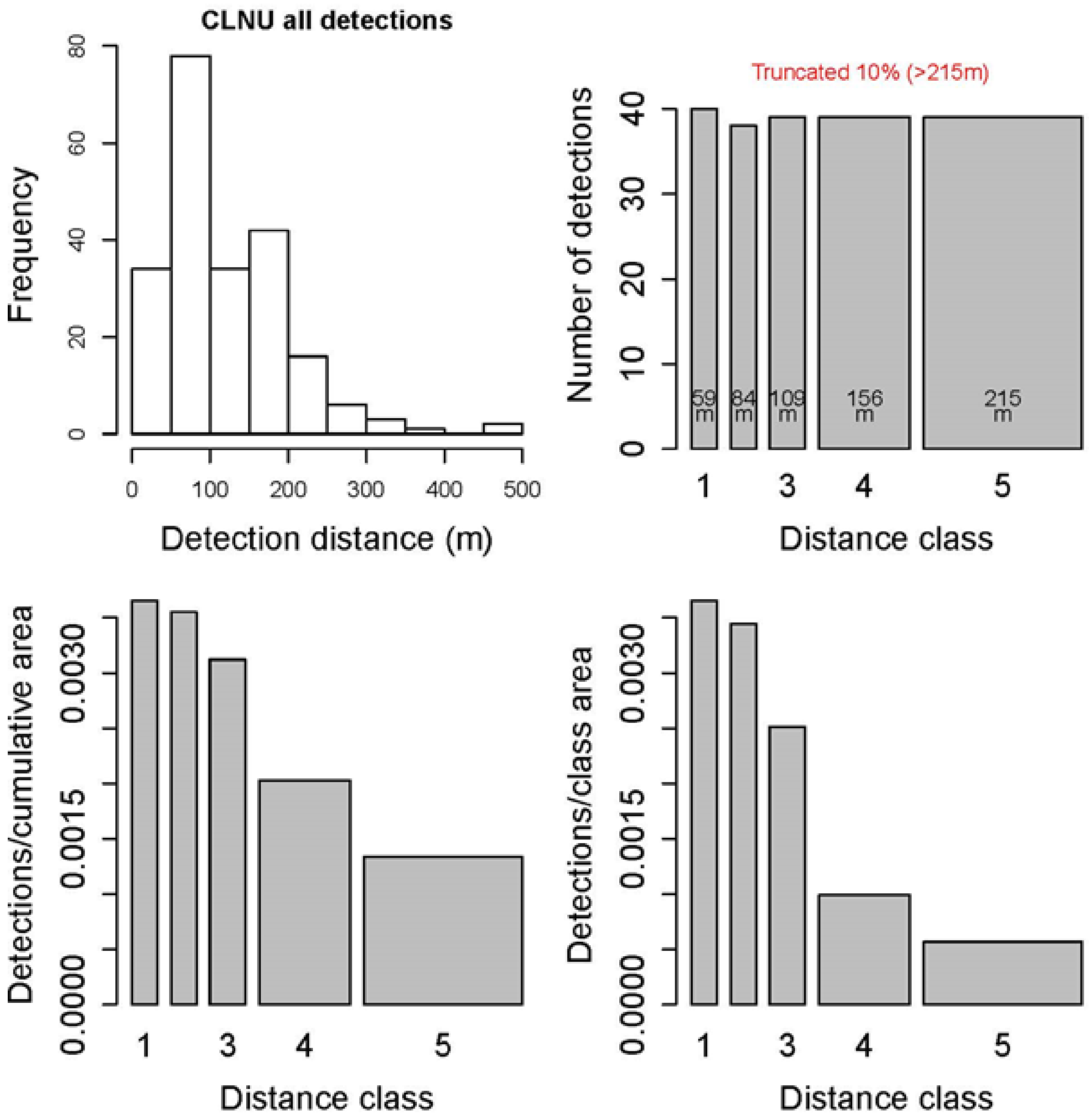


Figure A2.3. CLNU detections by distance (upper left) and by variable distance class (other panels) after truncating the most distant 10% of detections (> 215 meters from the observer) and equalizing the number of detections in each class. Detections per metric of area are explored in the lower panels.

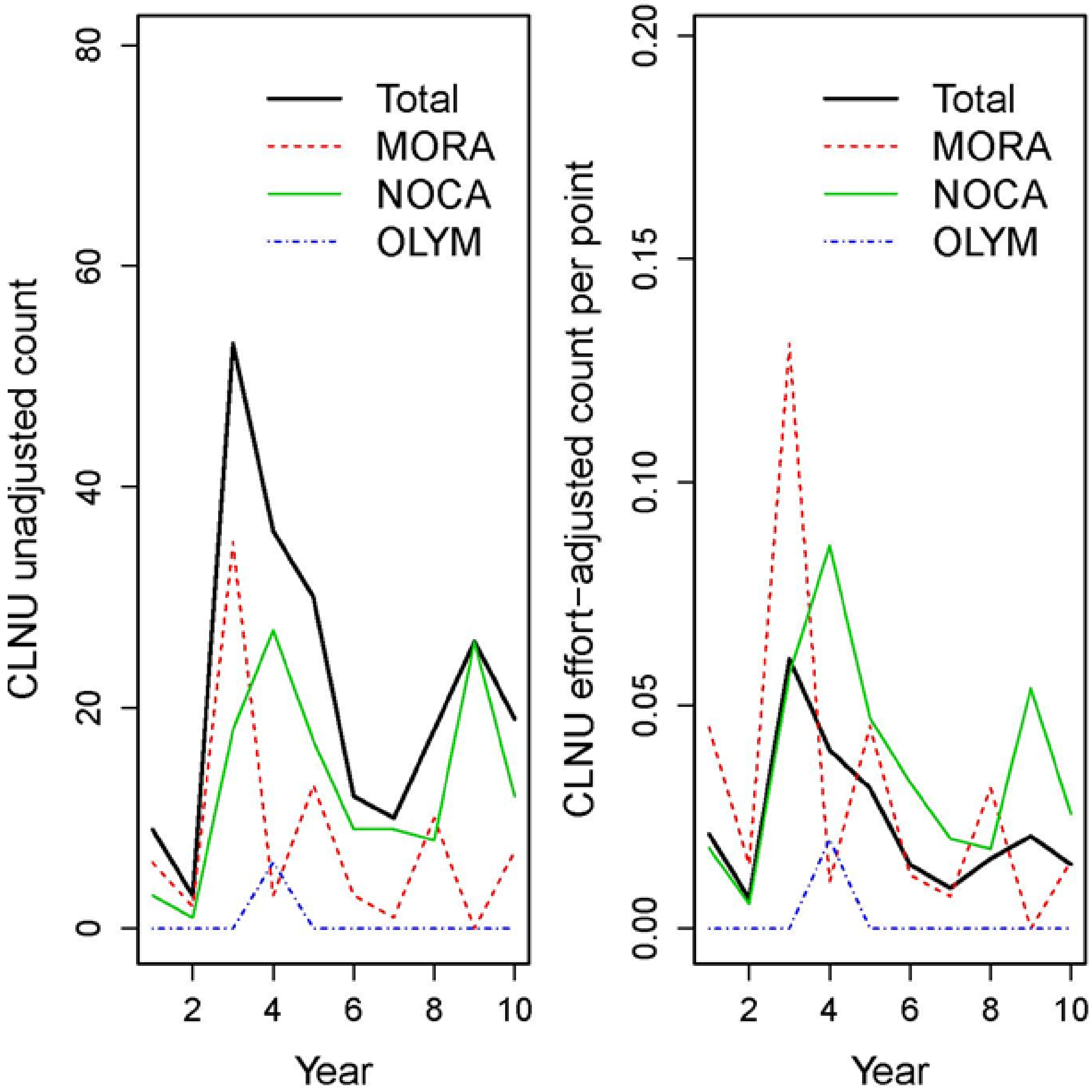


Figure A2.4. CLNU raw counts (left) and effort-adjusted counts (right) for each year 2005-2014 (shown here as year 1-10 of NCCN landbird monitoring), differentiated by park. Note that raw total counts do not approach Ray et al. estimates of Np (where p is detection probability) because N in Ray et al. includes estimates for points that were not surveyed each year (e.g., alternating panels).

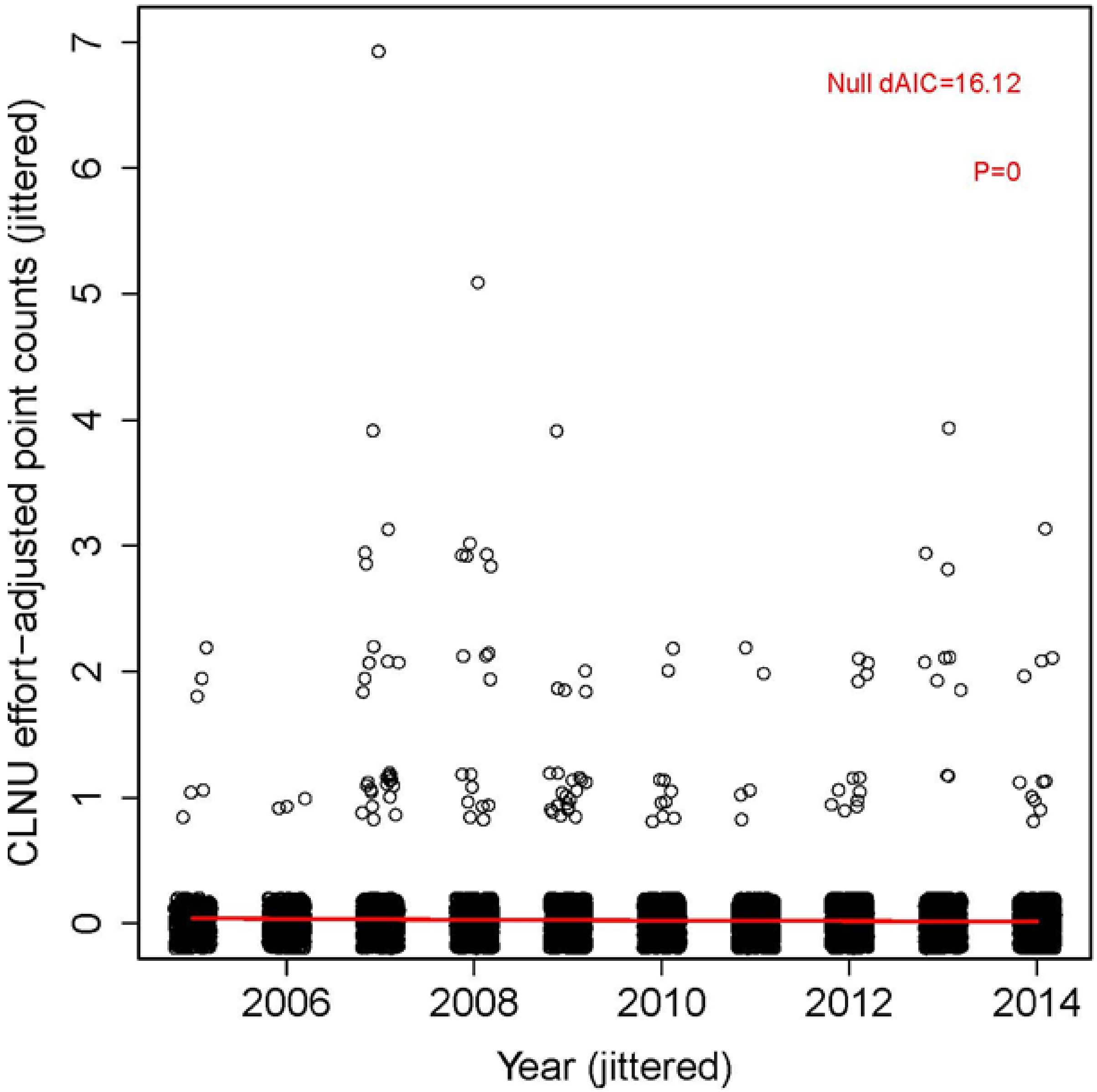


Figure A2.5. Poisson regression of CLNU counts, adjusted for effort within each park-year, on year of landbird monitoring in the NCCN. The slight slope of the negative trend over time (red line) is difficult to detect but garnered much more support than the null model (p = 0 and ΔAIC >> 7 with respect to the null model). Points were jittered for display by adding a small, random value to reduce overlap.

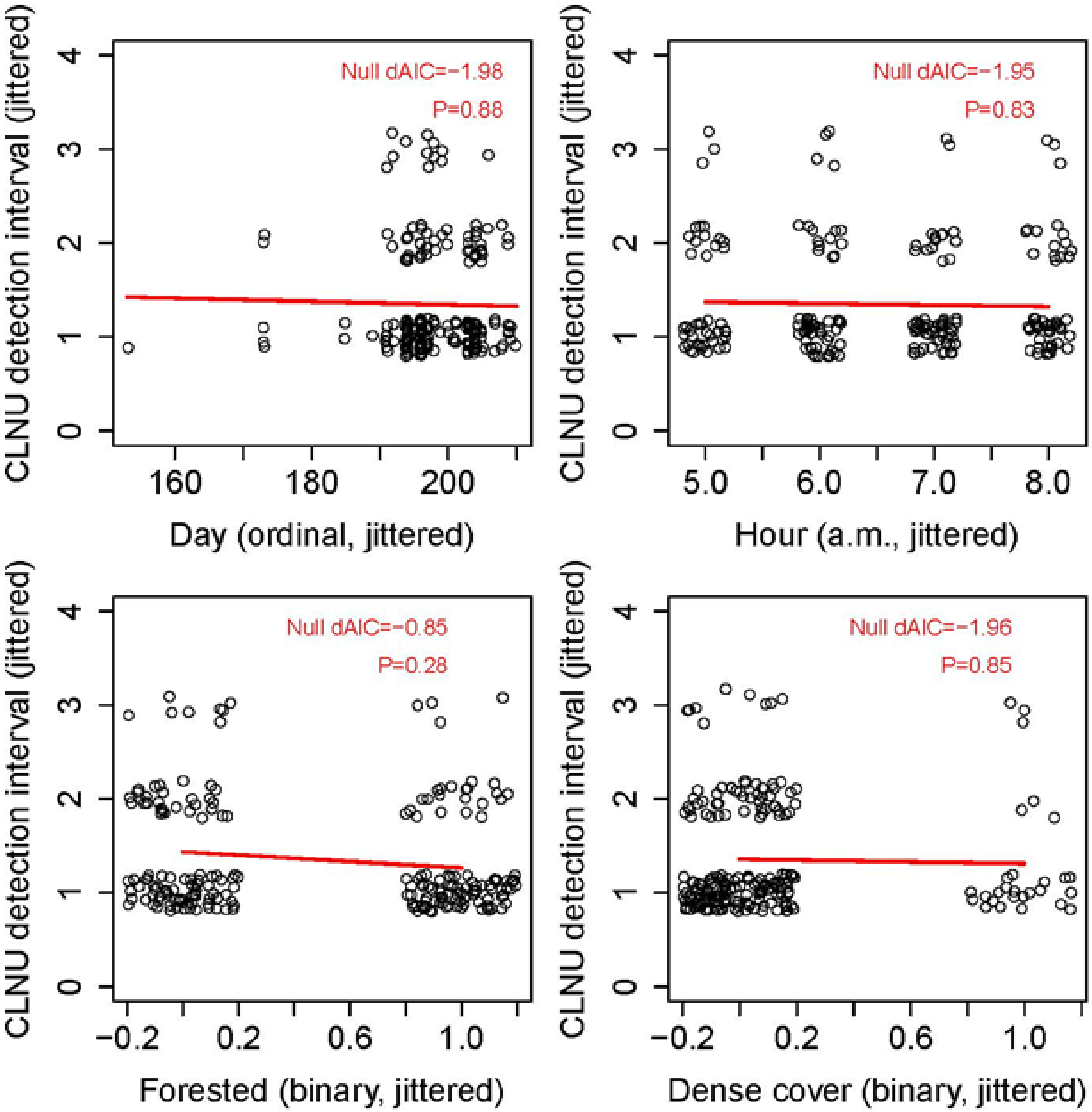


Figure A2.6. CLNU detection interval by day (upper left), hour (upper right), forest presence (lower left) and presence of dense cover (lower right). Results of Poisson regression appear in red; in each case, $P > 0.05$ and $\Delta AIC < 2$ with respect to the null model, so none of these effects on detection interval gained support. Points were jittered for display by adding a small, random value to reduce overlap.

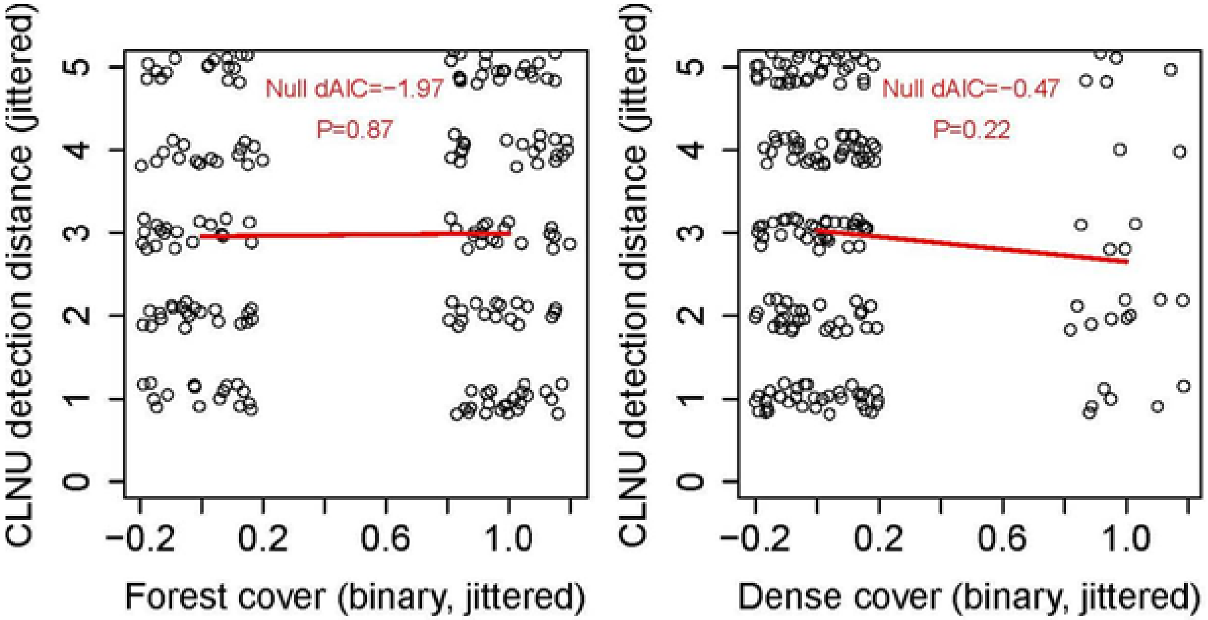


Figure A2.7. CLNU detection distance by forest presence (left) and presence of dense cover (right). Results of Poisson regression appear in red; in each case, p > 0.05 and ΔAIC < 2 with respect to the null model, so these effects on detection distance were not supported. Points were jittered for display by adding a small, random value to reduce overlap.

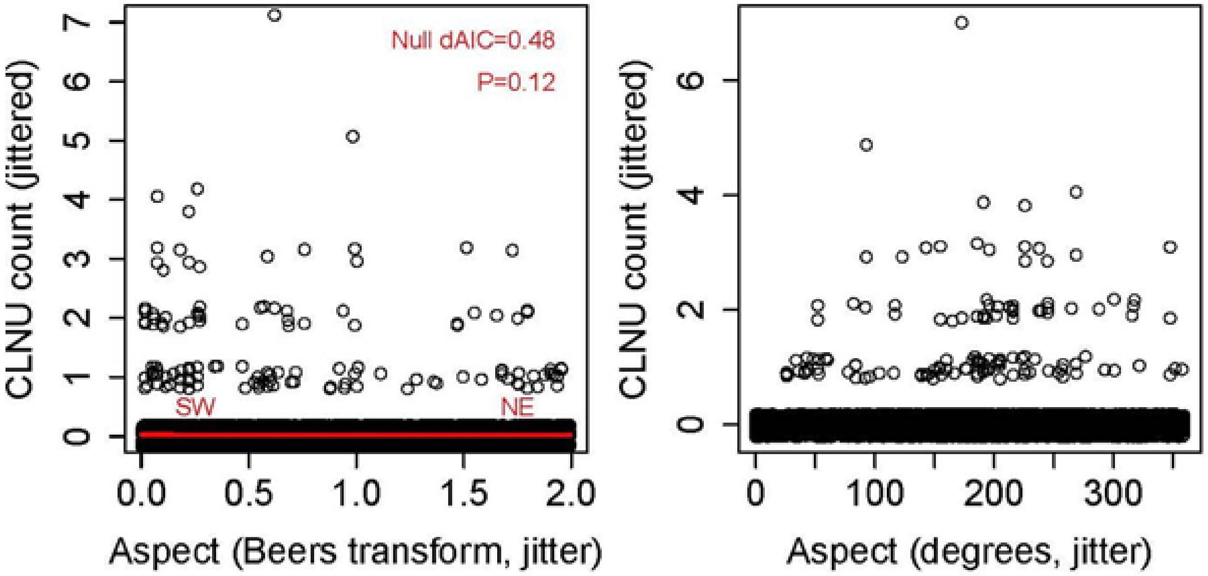


Figure A2.8. CLNU count by transformed aspect (left) untransformed aspect (right). Transformed aspect ranged from zero toward the southwest to 2 toward the northwest. Results of Poisson regression appear in red; p > 0.05 and ΔAIC < 2 with respect to the null model, so an effect of aspect on count was not supported. Points were jittered for display by adding a small, random value to reduce overlap.

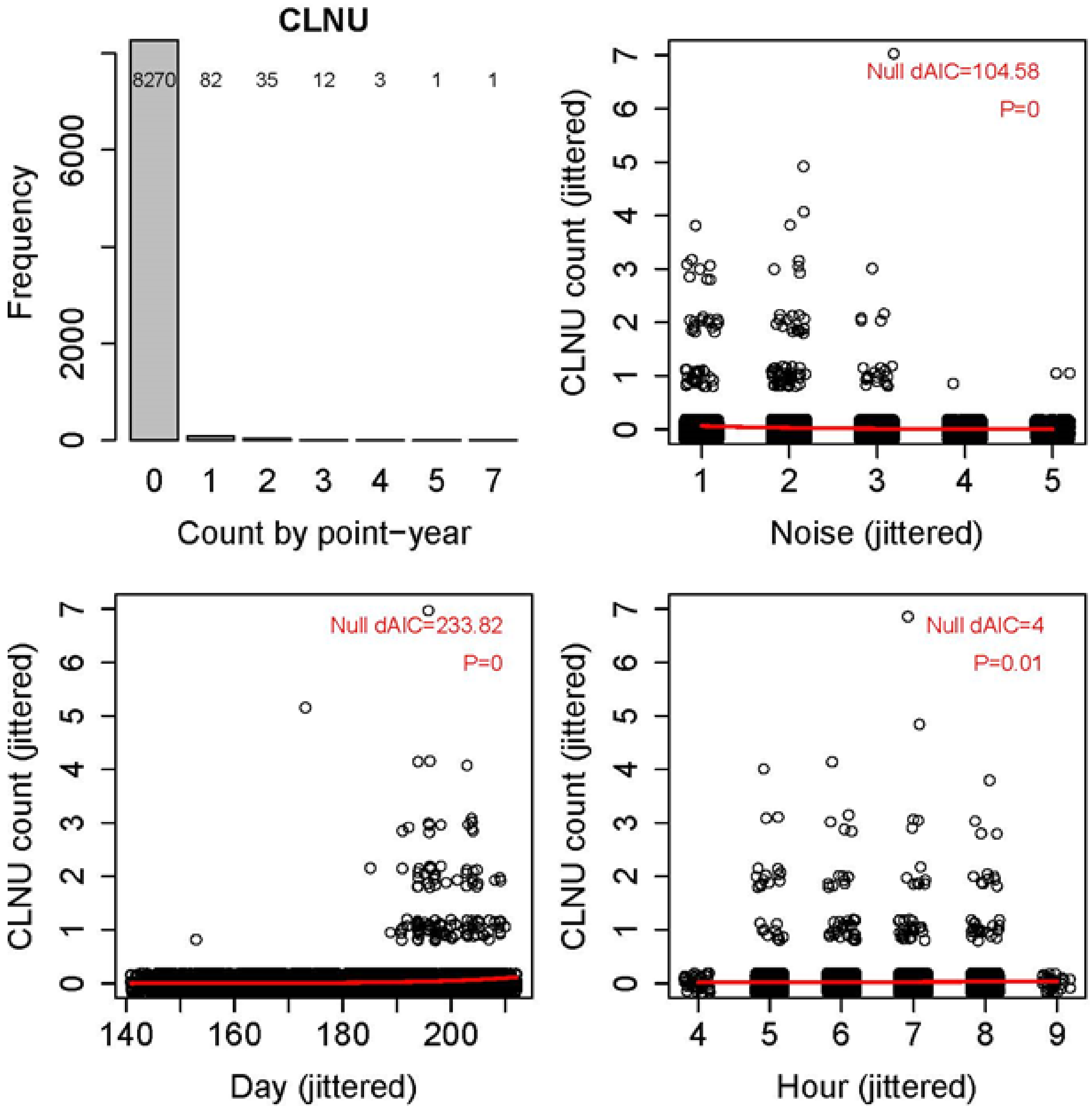


Figure A2.9. CLNU counts summed across parks ranged from zero (n = 8270) to 7 (n = 1) for each point-year (upper left) and were related to noise level during the count (upper right), hour of day (lower right) and day of the year (lower left) during 2005-2014. Each Poisson regression (in red) garnered more support than the null model (ΔAIC > 4). Points were jittered for display by adding a small, random value to reduce overlap.

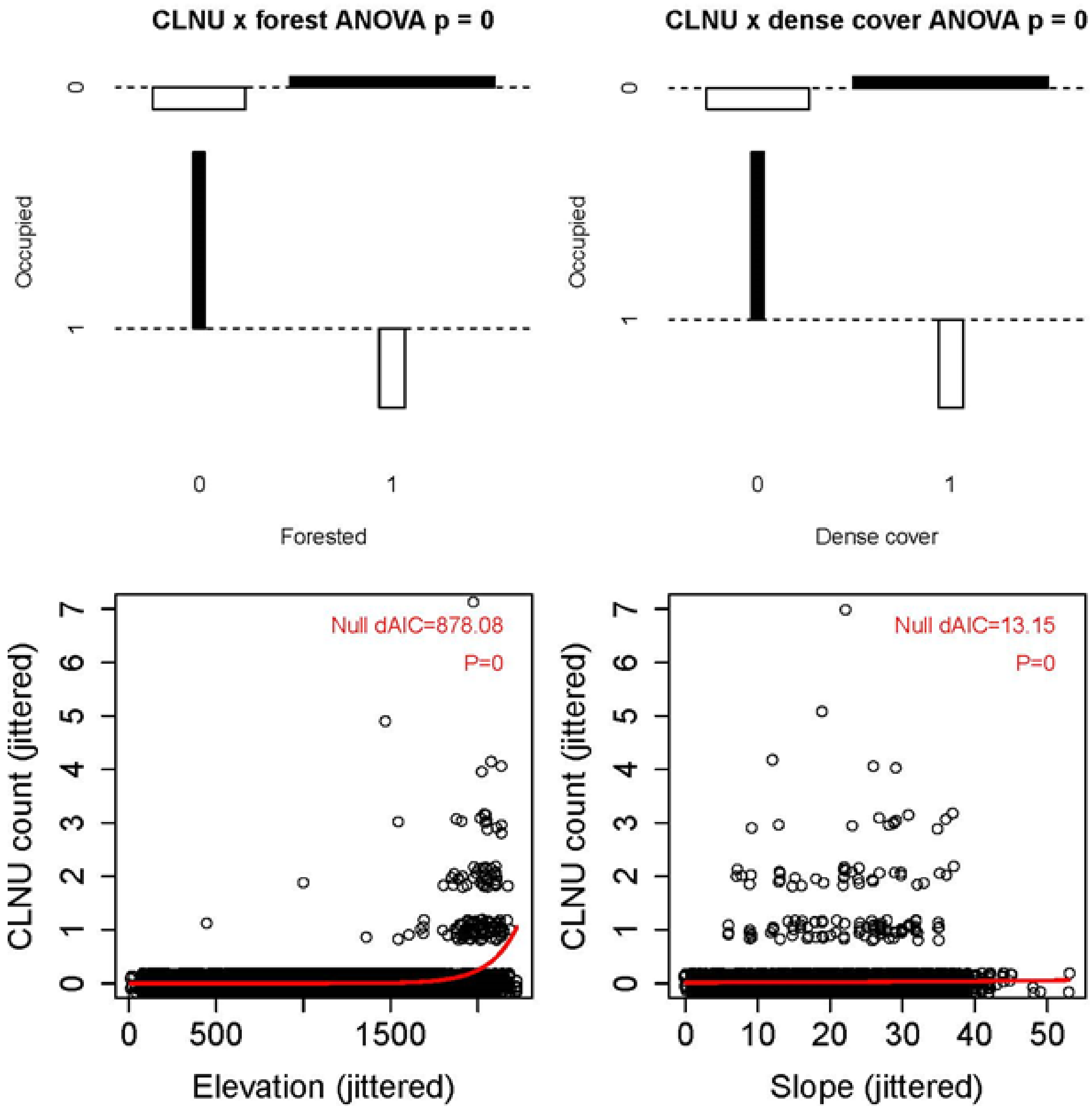
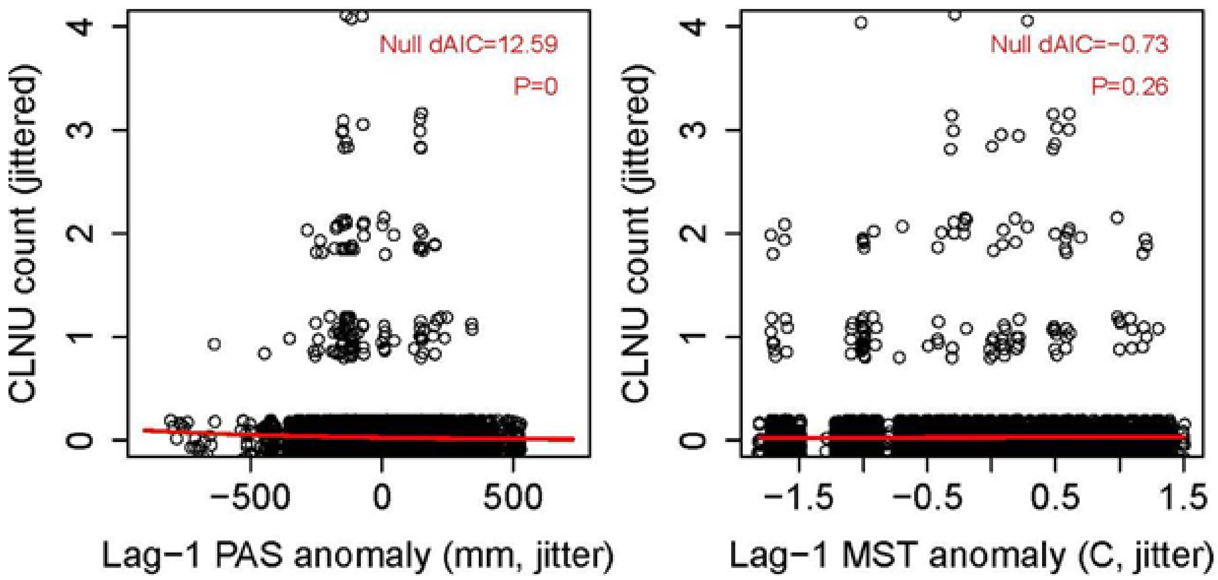


Figure A2.10. Association plots relating CLNU point-level occupancy to forest presence (upper left) and presence of dense cover (upper right), where plot titles report the significance of variance in raw counts explained by each predictor variable (p-value at $\alpha = 0.05$ by ANOVA). Poisson regression of counts on elevation (lower left) and slope (lower right) were also supported more than the null model ($p < 0.05$ and $\Delta AIC > 4$). Points were jittered for display by adding a small, random value to reduce overlap.



FigureA2.11. CLNU counts declined with increasing lag-1 precipitation-as-snow (left; $p < 0.05$, $\Delta AIC > 7$) but did not reflect a univariate effect of lag-1 mean spring temperature (right; $p > 0.05$, $\Delta AIC < 2$). To better display the slopes of these Poisson regressions, graphs were truncated at counts of 4; maximum count was 7. Points were jittered for display by adding a small, random value to reduce overlap.

Appendix 3: Query protocol and R script for accessing and formatting data from ClimateWNA

Prior to running the script below, access ClimateWNA at http://www.climatewna.com and copy the following files and folders into a single folder/working directory: files `ClimateWNA _v*.**.exe` and `help.rtf`, and folders `Prismdat`, `Perioddat` and `GCMdat`. Next, prepare a .csv file containing data on point locations in the following format:

Park_code	Site_code	Location_code	UTM_east	UTM_north
MORA	4001	NN01	583207	5176809
MORA	4001	NN02	583226	5177033
MORA	4001	NN02a	583195.7	5177015
…				

Running the first few lines of the script below will convert your .csv file above into an input file compatible with ClimateWNA, as:

ID1	ID2	Latitude	Longitude	Elevation
4001	NN01	46.73965	-121.911	.
4001	NN02	46.74166	-121.911	.
4001	NN02a	46.7415	-121.911	.
…				

After preparing a compatible input file as shown above, ClimateWNA can be queried as follows for data on means and normals: Open/run the executable file `ClimateWNA_v*.**.exe`. Click on "Select input file" to input your data formatted "`ID1, ID2,` …" Click on "Specify output file" to specify your output folder and file. Finally, click on "Calculate" to have data appended to your input file (as additional columns) and printed to your output file. If your input file includes elevations, the output climate data will be adjusted for elevation. For a time series of climate data, query as follows: Select "Time Series" (for historical years) or select a future time series in the period section drop box. Select a variable category (monthly, seasonal, annual or all variables). Input the starting and ending years in the pop-up boxes. Specify input and output files, and click the "Calculate TS" button. Again, everything is appended to your input file. The script below then reads your appended output, formats the data for use in models, and visualizes the data for review.

```
# Script for accessing and exploring ClimateWNA data

# Code file name “script4climateData.R” available from C. Ray (cray@birdpop.org)

# create input file compatible with ClimateWNA

require(rgdal)
# prepare UTM coordinates matrix
nccn.pts <- read.csv("nccn.pts.csv")
nccn.utm <-
SpatialPoints(cbind(nccn.pts$UTM_east,nccn.pts$UTM_north),proj4string=CRS("+proj=utm
+zone=10"))
```

```
# convert UTM to lon-lat
nccn.lonlat <- spTransform(nccn.utm, CRS("+proj=longlat"))
# prepare as input for ClimateWNA_v5.21
nccn.covars.wna.pts <- data.frame(ID1 = nccn.pts$Site_code, ID2 =
nccn.pts$Location_code, lat = nccn.lonlat$coords.x2, long = nccn.lonlat$coords.x1, el
= ".")
write.csv(nccn.covars.wna.pts, "nccn.wna.pts.csv", row.names=F)

# read ClimateWNA_v5.21 output (for example, 2005-2013 data)

clim <- read.csv("nccn.wna.pts_2005-2013MSYT.csv")
# add labels
nccn.pts <- read.csv("nccn.pts.csv")
clim$Park_code <- rep(nccn.pts$Park_code, 1)
clim <- clim[,c(ncol(clim), 1:ncol(clim)-1)]
names(clim)[3:4] <- c("Site_code", "Location_code")

# view some "30-yr normal" data for some reference periods

norm61 <- read.csv("nccn.wna.pts_Normal_1961_1990MSY.csv")
norm61$Park_code <- nccn.pts$Park_code # add a Park_code column, last
norm61 <- norm61[,c(ncol(norm61),1:ncol(norm61)-1)] # rearrange cols
names(norm61)[2:3] <- c("Site_code", "Location_code")
norm71 <- read.csv("nccn.wna.pts_Normal_1971_2000.csv")
norm71$Park_code <- nccn.pts$Park_code
norm71 <- norm71[,c(ncol(norm71), 1:ncol(norm71)-1)]
names(norm71)[2:3] <- c("Site_code", "Location_code")
ppi <- 300
#jpeg("norm61vs71-PAS.jpg", width=6*ppi, height=6*ppi, res=ppi)
layout(matrix(c(1,2), 1, 2, byrow = FALSE)) # 2 panels
boxplot(norm61$PAS~norm61$Park_code,cex.lab=1.15,ylim=c(0,2500),
        ylab="Precipitation as snow (mm, Aug[t-1]-Jul[t])",
        main="PAS 1961-1990 normals",las=3)
text(3,2500,paste("mean(NOCA) =",
     round(mean(norm61$PAS[norm61$Park_code=="NOCA"]))))
boxplot(norm71$PAS~norm71$Park_code,cex.lab=1.15,ylim=c(0,2500),
        ylab="Precipitation as snow (mm, Aug[t-1]-Jul[t])",
        main="PAS 1971-2000 normals",las=3)
text(3,2500,paste("mean(NOCA) =",
     round(mean(norm71$PAS[norm71$Park_code=="NOCA"]))))
#dev.off()
#jpeg("norm61vs71-MST.jpg", width=6*ppi, height=6*ppi, res=ppi)
boxplot(norm61$Tave_sp~norm61$Park_code,cex.lab=1.15,ylim=c(-2,10),
        ylab="Mean spring temperature (Mar-May)",
        main="MST 1961-1990 normals",las=3)
text(3,10,paste("mean(NOCA) =",
     round(mean(norm61$Tave_sp[norm61$Park_code=="NOCA"]),1)))
boxplot(norm71$Tave_sp~norm71$Park_code,cex.lab=1.15,ylim=c(-2,10),
        ylab="Mean spring temperature (Mar-May)",
        main="MST 1971-2000 normals",las=3)
text(3,10,paste("mean(NOCA) =",
     round(mean(norm71$Tave_sp[norm71$Park_code=="NOCA"]),1)))
#dev.off()

# subtract 1971-2000 normals from 2005-2013 data for PAS and MST

require(dplyr)
c <- left_join(clim[,c(1:7,185,248)],norm71[,c(1:6,184,247)],
               by=c("Park_code","Site_code","Location_code",
                    "Latitude","Longitude","Elevation"))
names(c)
names(c)[8:11] <- c("MST","PAS","MSTnorm","PASnorm")
c <- cbind(c,c$MST-c$MSTnorm,c$PAS-c$PASnorm)
```

```
names(c)[12:13] <- c("anomMST","anomPAS")

# create new mtn-park dataset

i <- c$Park_code=="MORA"|c$Park_code=="NOCA"|c$Park_code=="OLYM"
data <- c[i,c("Year","Park_code","Site_code","Location_code",
              "anomMST","anomPAS")]
names(data)[1] <- "Event_year"
head(data); dim(data)

# write data to file

write.csv(data,file="nccn.climate.data.2005-2013.csv")

# explore distribution

anomPAS.sc <- scale(data$anomPAS)
hist(anomPAS.sc,breaks=100)
shapiro.test(sample(anomPAS.sc,500)) # 5000 = max for shapiro.test
# :. use scaled anomPAS in regressions (sampled W often -> p > 0.05)
anomMST.sc <- scale(data$anomMST)
hist(anomMST.sc,breaks=100)
shapiro.test(sample(anomMST.sc,500))
hist(anomMST,breaks=100)
shapiro.test(sample(anomMST,500))

# plot MST vs PAS

ct <- cor.test(data$anomMST,data$anomPAS,method="kendall"); ct
#jpeg("anomaliesPASvsMST-mtnPks-pointLevel.jpg", width=6*ppi,
#height=6*ppi, res=ppi)
par(mfrow=c(1,1),mar=c(5,5,3,2))
plot(data$anomMST,data$anomPAS,main="All park-years, mountain parks",
     xlab=expression(paste(
          "Anomaly, mean spring temperature ("^"o","C)",sep="")),
     ylab="Anomaly, precipitation as snow (mm)",cex.lab=1.5)
text(-1.0,-750,paste("Kendall's tau = ",round(ct$estimate,2),
                      ", p = ",round(ct$p.value,2),sep=""),cex=1)
#dev.off()

# anomaly plots for data from 3 mtn parks

ppi <- 600
#png("PASanomByYr-mtnPks-wrt71norm-in-line600dpi.png", width=9.4*ppi,
#height=6.74*ppi, res=ppi)
lo <- layout(matrix(c(1,1,1,2), 1, 4, byrow = FALSE)) # 2 panels
par(mar=c(6,6,2,1))
boxplot(data$anomPAS~data$Event_year,
        ylab="Annual precipitation-as-snow anomaly (mm)",
        # main="PAS anomaly based on 1971-2000 normals",
        las=3,cex.lab=2.5,cex.axis=2)
abline(0,0,lty=3)
#text(1,650,"(a)",cex=3)
par(mar=c(6,3,2,2))
boxplot(data$anomPAS,xlab="2005-2013",cex.lab=2,cex.axis=2)
abline(0,0,lty=3)
#dev.off()
#png("MSTanomByYr-mtnPks-wrt71norm-in-line600.png", width=9.4*ppi,
#height=6.74*ppi, res=ppi)
par(mar=c(6,6,2,1))
boxplot(data$anomMST~data$Event_year,
        ylab=expression(paste(
             "Mean spring temperature anomaly ("^"o","C)"),sep=""),
```

```
        las=3,cex.lab=2.5,cex.axis=2)
abline(0,0,lty=3)
#text(1,-1.6,"(b)",cex=3)
par(mar=c(6,3,2,2))
boxplot(data$anomMST,xlab="2005-2013",cex.lab=2,cex.axis=2)
abline(0,0,lty=3)
#dev.off()

# re-calc correlation between PAS and MST
# aggregate by transect to reduce spatial pseudorep
j <- which(data$Park_code=="OLYM"); length(j)
agAnomMST <- aggregate(data$anomMST[j]~data$Site_code[j],FUN=mean)
par(mfrow=c(1,1),mar=c(5,5,3,2))
hist(agAnomMST[,2])
agAnomPAS <- aggregate(data$anomPAS[j]~data$Site_code[j],FUN=mean)
hist(agAnomPAS[,2])
# not normally distributed so use Kendallz tau
cor.test(agAnomPAS[,2],agAnomMST[,2],method="kendall")
# calc correlation via Cochrane-Orcutt procedure
require(orcutt)
reg <- lm(agAnomPAS[,2]~agAnomMST[,2])
summary(reg)
reg2 <- cochrane.orcutt(reg)
reg2
```

Appendix 4: R and embedded JAGS scripts for fitting NCCN landbird monitoring data to *N*-mixture models to estimate trends in population density

```
# Adapts the pre-processed survey and point data for use in hierarchical models, and
# fits Amundson-type models extended to include variation in interval length
# and trends in population density using data from multiple parks and years

# Code file name “script4trendAnalysis.R” available from C. Ray (cray@birdpop.org)

# 1) Adapt the clean survey and point data for the model in part 2 (~line 500)

# Read the .csv output from survey data query "qs_a121..." and point data query
# "qs_a204..."; interpret survey data from the perspective of a given focal species;
# select and display data from a subset of parks, years to guide model construction

# “***#” is used below to throw an error and call attention to any parameters
# requiring updates before fitting data from different species or to different models

require(dplyr)

***# update 5 settings as needed:
species <- "OSFL"
year <- "all" # one of '2005' ... to '2014' or 'all'
park <- "all" # one of 'MORA', 'NOCA', 'OLYM' or 'all' 3 mtn parks
bins <- "equalDensity" # one of 'equalWidth', 'equalDensity' or 'equalArea'
outfile <- paste(species,"out1.eqD",sep="")

***# update input file names as needed
sd <- read.csv(file="nccn.survey.data.2005to2014.csv") #survey data
pd <- read.csv(file="nccn.site.data.2005to2014.csv") #point/site data
cd <- read.csv(file="nccn.climate.data.2005to2014-lagged1at2004-2013.csv")
#ClimateWNA data are lagged 1 yr for both PAS and MST

# - merge survey, point and climate data

#before merging, compare structures to ensure compatibility; e.g., Location_code
#might contain leading zeros for a transect in sd ("069a") but not in pd ("69a")
#note: just replace "069a" with "69a" in the input file before reading it in; to
#no avail I coded fixes (levels(x)[levels(x)=="069a"]<-"69a" or the "plyr" fn
#revalue(x,c("069a"="69a")); though the merge DID work it still gave a warning
names(sd); names(pd); names(cd)
sd <- sd[,-1]; pd <- pd[,-1]; cd <- cd[,-1] #remove first columns
sharedVars <- intersect(colnames(sd),colnames(pd))
str(sd[,sharedVars])
str(pd[,sharedVars])
sharedVars <- intersect(colnames(sd),colnames(cd))
str(cd[,sharedVars])
sharedVars <- intersect(colnames(pd),colnames(cd))
str(cd[,sharedVars])

#Start_date format varies, so omit it from the merge (we'll use Day instead)
j <- c("Event_year","Site_code","Location_code","Park_code","Stratum_name",
 "Species","Observer","Noise_level","Ever_sang","Seen_first","Group_size",
 "Obs_distance_m","Obs_notes","Day","Hour","y1","y2","y3")
sd <- sd[,j]; dim(sd)
j <- c("Event_year","Site_code","Location_code","Park_code","Panel_type","Panel_name",
 "Location_type","Detection_class","Is_forested","Habitat_num",
 "Elevation_m","Slope_deg_90mAvg","Aspect_deg","Beers_90mAvg","UTM_east","UTM_north")
pd <- pd[,j]; dim(pd)
#merge survey data with point data, keeping Park_code.x and Park_code.y separate
```

```
#note: this merge eliminates any problematic rows of pt-yr data w/o survey data
data <- left_join(sd,pd,by=c("Event_year","Site_code","Location_code")); dim(data)

#use Park_code, etc. to inspect the merge
all.equal(data$Park_code.x,data$Park_code.y) #if not TRUE, fix
sum(is.na(data$Park_code.x))
sum(is.na(data$Park_code.y))
sum(is.na(data$Elevation_m))
head(data); tail(data) #everything looks good!

#clean up the data frame
data <- rename(data,Park_code=Park_code.x)
data <- select(data,-Park_code.y)
head(data)

#merge with climate data
data <- left_join(data,cd,by=c("Event_year","Site_code","Location_code")); dim(data)
***problems here can originate with the "069a" Location_code

#inspect the merge
names(data)
all.equal(data$Park_code.x,data$Park_code.y) #if not TRUE, fix
***problems here can originate with the "069a" Location_code
table(data$Park_code.x); table(data$Park_code.y)
sum(is.na(data$Park_code.x))
sum(is.na(data$anomPAS1))
sum(data$Event_year=="2014")
head(data); tail(data) #MAKE SURE everything looks good!

#clean up the data frame
data <- rename(data,Park_code=Park_code.x)
data <- select(data,-Park_code.y)
head(data); tail(data)

# - remove data from small parks LEWI and SAJH

i <- which(data$Park_code=="LEWI"|data$Park_code=="SAJH"); length(i)
data <- data[-i,]; dim(data)

# - set Group_size = 0 for non-focal spp (including species=NA/Nearest_obs=FALSE)

data[which(data$Species!=species),"Group_size"] <- 0
data[is.na(data$Species),"Group_size"] <- 0
table(data$Group_size)
focal_ct <- table(data$Group_size)[2] #keep this as a check below

# - construct focal-sp non-detections, retaining all survey covar data

#this method retains data on observer, etc.; in contrast, using unique()
#is complicated by multiple records of the focal sp at a pt-yr, etc.

#omit column of species names (Group_size will represent focal-sp detections)
data <- data.frame(data[,-which(names(data)=="Species")])
names(data); dim(data)

#separate focal-sp detections from other detections
data1 <- data[which(data$Group_size!=0),]; dim(data1) #focal-sp detections
data0 <- data[which(data$Group_size==0),]; dim(data0) #non-focal sp detections
data0 <- data0[!duplicated(data0[,c("Event_year","Site_code","Location_code")]),]
dim(data0) #unique pt-yrs in which a non-focal sp was detected

#eliminate non-focal sp detections from pt-yrs with focal-sp detections (fsd) by
#stacking unique (nd) pt-yrs w/fsd above unique pt-yrs w/non-fsd to use !duplicated()
```

```
#note: we are not eliminating any focal-sp data in this temporary process
data1.nd <- data1[!duplicated(data1[,c("Event_year","Site_code","Location_code")]),]
dim(data1.nd) #unique pt-yrs in which the focal sp was detected
temp <- rbind.data.frame(data1.nd,data0); dim(temp) #unique pt-yrs w/fsd over w/o
nd <- temp[!duplicated(temp[,c("Event_year","Site_code","Location_code")]),]
dim(nd) #a single event (fsd or non-fsd) for every pt-yr surveyed
data0 <- nd[nd$Group_size==0,]; dim(data0) #non-fsd only at pt-yrs w/no fsd

#remove interval counts and associated data from pt-yrs w/no fsd, retaining NAs
data0[,c("Obs_distance_m","Obs_notes")] <- NA
j <- c("Ever_sang","Seen_first","y1","y2","y3")
apply(data0[,j],2,function(x) sum(is.na(x))) #display NAs for the check below
data0[,j][data0[,j]>0] <- 0 #set all numbers to zero, retaining NAs
apply(data0[,j],2,function(x) sum(x,na.rm=T)) #check - all zeros?
apply(data0[,j],2,function(x) sum(is.na(x))) #check - NAs retained?

#create compact data frame of survey data, restoring all focal-sp data
data <- rbind.data.frame(data1,data0); dim(data); head(data)
table(data$Group_size)
focal_ct #should agree with table above

# - index the focal data

ip <- iy <- 1:dim(data)[1] #in case park=="all" & year=="all"
if (park!="all") {ip <- which(data$Park_code==park)}
if (year!="all") {iy <- which(data$Event_year==year)}
data.i <- intersect(ip,iy)

# - determine extent of NAs within the useful set of survey covars

j <- which(names(data)=="Obs_distance_m"|names(data)=="Obs_notes"|names(data)=="y3")
j
sum(complete.cases(data[data.i,-j])); sum(is.na(data[data.i,-j]))

# - write focal data to file

dat <- data[data.i,]
dat <- dat[order(dat$Event_year,dat$Site_code,dat$Location_code),]
str(droplevels(dat))
write.csv(dat,file=paste(species,".",year,".",park,".csv",sep="")) #*** update file
name as needed

# - if year=all, pad data w/NAs for points not visited in certain years

if (year=="all") {

    #find unique points and unique point-years
    uniq.pts <- dat[!duplicated(dat[,c("Site_code","Location_code")]),]
    uniq.pt.yrs <-
        dat[!duplicated(dat[,c("Event_year","Site_code","Location_code")]),]
    dim(dat); dim(uniq.pts); dim(uniq.pt.yrs)

    #create "PY" data frame to hold all pt-yr data, including pt-yrs not surveyed
    a <- uniq.pts; names(a) #set time-varying covars to NA
    a[,c("Detection_class","Is_forested","Habitat_num")] <- NA
    yrs <- range(a$Event_year,na.rm=T); totyrs <- yrs[2]-yrs[1]+1; totyrs
    PY <- a #stack "a" into PY "totyrs" times
    for (i in 2:totyrs) PY <- rbind(PY,a)
    ey <- rep(yrs[1]:yrs[2],each=dim(uniq.pts)[1]) #create Event_year column
    length(ey)
    PY[,"Event_year"] <- ey; head(PY); tail(PY$Event_year)
    #join by static point attributes
    PY <- left_join(PY,dat,by=c("Event_year","Site_code","Location_code",
```

```
                                "Park_code","Stratum_name","Panel_type","Panel_name",
                                "Location_type","Elevation_m","Slope_deg_90mAvg",
                                "Aspect_deg","Beers_90mAvg","UTM_east","UTM_north"))
    apply(PY,2,function(x) sum(is.na(x))) #NAs should vary between .x and .y columns
    names(PY); j <- grep(".x",names(PY)); j #j identifies .x columns to delete
    PY <- PY[,-j]; names(PY)
    names(PY) <- sub("\\.y","",names(PY)); head(PY) #re-name .y columns
    #add climate vars for every pt-yr regardless of sampling
    PY <- left_join(PY,cd,by=c("Event_year","Site_code","Location_code","Park_code"))
    apply(PY,2,function(x) sum(is.na(x))) #NAs should vary between .x and .y columns
    names(PY); j <- grep(".x",names(PY)); j #j identifies .x columns to delete
    PY <- PY[,-j]; names(PY)
    names(PY) <- sub("\\.y","",names(PY)); head(PY) #re-name .y columns
    write.csv(PY,file=paste(species,".with.NAs.csv",sep="")) #*** update file name
    dat <- PY

} # end: if (year=="all")

# - calculate y = count of individual birds by point-year

dat <- mutate(dat,ptyr=paste(Site_code,Location_code,Event_year,sep="."))
dat <- mutate(dat,PTYR=ptyr)
by_ptyr <- group_by(dat,ptyr) #group by pt-yr
#re-code Detection_class to allow averaging across records below (for now)
dat$Detection_class <- abs(as.numeric(dat$Detection_class)-2)
by_ptyr_v <- summarize(by_ptyr,
                         y=sum(Group_size,na.rm=T),
                         PTYR=first(PTYR),
                         noise=first(Noise_level),
                         day=first(Day),
                         hour=mean(Hour,na.rm=T),
                         obs=first(Observer),
                         forest=mean(Is_forested,na.rm=T),
                         dense=mean(abs(as.numeric(Detection_class)-2),na.rm=T),
                         elev=first(Elevation_m),
                         slope=first(Slope_deg_90mAvg),
                         aspect.d=first(Aspect_deg),
                         aspect.b=first(Beers_90mAvg),
                         strat=first(Stratum_name),
                         tran=first(Site_code),
                         Park=first(Park_code),
                         Year=first(Event_year),
                         PASanom=first(anomPAS1),
                         MSTanom=first(anomMST1))

#re-name the vars
y <- by_ptyr_v$y
PTYR <- by_ptyr_v$PTYR
noise <- by_ptyr_v$noise #covars will be rescaled in data step
day <- by_ptyr_v$day
hour <- by_ptyr_v$hour
obs <- droplevels(by_ptyr_v$obs)
forest <- by_ptyr_v$forest
dense <- by_ptyr_v$dense
elev <- by_ptyr_v$elev
slope <- by_ptyr_v$slope
aspect.d <- by_ptyr_v$aspect.d #Aspect_deg
aspect.b <- by_ptyr_v$aspect.b #Beers_90mAvg
strat <- droplevels(by_ptyr_v$strat)
tran <- droplevels(by_ptyr_v$tran)
Park <- droplevels(by_ptyr_v$Park) #lowercase "park" is used elsewhere to
                                   #define the dataset (park==all)
Year <- by_ptyr_v$Year #lowercase "year" used elsewhere
```

```
PASanom <- by_ptyr_v$PASanom
MSTanom <- by_ptyr_v$MSTanom

#get rid of zeros in y that are really NAs
i <- which(is.na(day)); length(i)
y[i] <- NA; table(y)
#get rid of NaNs in hour, which resulted from NAs when taking the mean
i <- which(is.nan(hour)); length(i)
hour[i] <- NA; table(hour)

# display data and relationships

page <- 1
pdf(file=paste("pt-yr.plots",page,".pdf",sep=""))

par(mfrow=c(2,2),mar=c(4,4,1,1)+0.1)
layout(matrix(c(1,2,3,3), 2, 2, byrow = TRUE))

#day
h <- hist(day,plot=F); h
tb <- h$counts; labText <- "Day of count (ordinal)"
names(tb) <- round(h$mids)
bp <- barplot(tb,ylab="Point-counts completed",cex.axis=1.5,cex.lab=1.5,
              cex.names=1.25,xlab=labText,las=3)
legend(bp[round(length(bp)/2)],max(tb),xjust=0.5,bty="n",
       legend=c(paste("Min",min(day,na.rm=T)),paste("Max",max(day,na.rm=T))))

#hour
tb <- table(hour); labText <- "Hour of count (a.m.)"
bp <- barplot(tb,ylab="Point-counts completed",cex.axis=1.5,cex.lab=1.5,
              cex.names=1.5,xlab=labText)

#observers
tb <- table(droplevels(obs)); labText <- "Observer (initials)"
firstNames <- sub("^.*_","",names(tb)); firstNames
firstInitials <- substring(firstNames,1,1); firstInitials
lastNames <- sub("_.*$","",names(tb)); lastNames
lastInitials <- substring(lastNames,1,1); lastInitials
initials <- paste(firstInitials,lastInitials,sep=""); initials
names(tb) <- initials; names(tb)
bp <- barplot(tb,ylab="Point-counts completed",cex.axis=1.5,cex.lab=1.5,
              cex.names=1.25,xlab=labText,las=3)

dev.off()
page <- 2
pdf(file=paste("pt-yr.plots",page,".pdf",sep=""))

par(mfrow=c(2,2),mar=c(4,8,2,8)+0.1)

#surveys X park
tb <- table(droplevels(Park[hour>0])); labText <- "Park"
bp <- barplot(tb,ylab="Point-counts completed",cex.axis=1.5,cex.lab=1.5,space=0.75,
              cex.names=1.5,las=3)

#strat
tb <- table(droplevels(strat[hour>0])); labText <- "Stratum"
names(tb) <- sub("^.*_","",names(tb))
bp <- barplot(tb,ylab="Point-counts completed",cex.axis=1.5,cex.lab=1.5,space=0.75,
              cex.names=1.5,xlab=labText)

#forest
tb <- table(forest); labText <- "Forested?"
names(tb) <- c("No","Yes")
```

```
bp <- barplot(tb,ylab="Frequency",cex.axis=1.5,cex.lab=1.5,space=1,
              cex.names=1.5,xlab=labText)

#dense
tb <- table(dense); labText <- "Dense cover?"
names(tb) <- c("No","Yes")
bp <- barplot(tb,ylab="Frequency",cex.axis=1.5,cex.lab=1.5,space=1,
              cex.names=1.5,xlab=labText)

dev.off()
page <- 3
pdf(file=paste("pt-yr.plots",page,".pdf",sep=""))

par(mfrow=c(1,1),mar=c(4,8,1,8)+0.1)

#noise
tb <- table(noise); labText <- "Noise level during count"
bp <- barplot(tb,ylab="Frequency",cex.axis=1.5,cex.lab=1.5,
              cex.names=1.5,xlab=labText)
for (i in 0:max(tb)) {text(bp[i+1],0.9*max(tb),tb[i+1],cex=0.9)}

dev.off()
page <- 4
pdf(file=paste("pt-yr.plots",page,".pdf",sep=""))

par(mfrow=c(2,2),mar=c(4,4,1.5,1)+0.1)

#elev
h <- hist(elev[hour>0],plot=F); h
tb <- h$counts; labText <- "Elevation (x100 m)"
names(tb) <- as.numeric(h$mids)/100
bp <- barplot(tb,ylab="Frequency",cex.axis=1.5,cex.lab=1.5,
              cex.names=1,xlab=labText,las=3)

#slope
h <- hist(slope[hour>0],plot=F); h
tb <- h$counts; labText <- "Slope (degrees)"
names(tb) <- round(as.numeric(h$mids))
bp <- barplot(tb,ylab="Frequency",cex.axis=1.5,cex.lab=1.5,
              cex.names=1.25,xlab=labText,las=3)

#aspect.d
h <- hist(aspect.d[hour>0],plot=F); h
tb <- h$counts; labText <- "Aspect (degrees)"
names(tb) <- h$mids
bp <- barplot(tb,ylab="Frequency",cex.axis=1.5,cex.lab=1.5,
              cex.names=1,xlab=labText,las=3)

#aspect.b
h <- hist(aspect.b[hour>0],plot=F); h
tb <- h$counts; labText <- "Aspect (Beers transform)"
names(tb) <- h$mids
bp <- barplot(tb,ylab="Frequency",cex.axis=1.5,cex.lab=1.5,
              cex.names=1,xlab=labText,las=3)
#Beers_90mAvg=1+cos((45-Aspect_deg[averaged over 90-m radius])*pi/180) (Beers 1966)
#can't back-transform uniquely, so grab original Aspect_deg, below

par(new=TRUE,mar=c(15,13,1.5,1)+0.1)
plot(aspect.d[hour>0],aspect.b[hour>0],xlab="Aspect",ylab="Beers transform",pch=".")

dev.off()
page <- 5
pdf(file=paste("pt-yr.plots",page,".pdf",sep=""))
```

```
#returns a p-value from ANOVA of y~x
aov.p.val <- function(y,x) {
                aov.yx <- aov(y~x)
                return(unlist(summary(aov.yx))["Pr(>F)1"])
}

assoc <- function(x,y,mtxt,xtxt,ytxt,rtxt) {
            tb <- table(x,y)
            if (!missing(rtxt)) rownames(tb) <- rtxt
            assocplot(tb[,1:dim(tb)[2]],main=mtxt,xlab=xtxt,ylab=ytxt,
                      col=c("black","white"))
}

par(mfrow=c(2,2),mar=c(4,4,1,1)+0.1)

if (park=="all") {

    #noise x park
    xa <- droplevels(Park); ya <- noise
    mtxt <- paste("ANOVA p =",round(aov.p.val(ya,xa),digits=3))
    assoc(xa,ya,mtxt,"Park","Noise")

    #park x observer ("initials" for each observer defined above)
    xa <- droplevels(obs); ya <- droplevels(Park)
    mtxt <- paste("ANOVA p =",round(aov.p.val(as.numeric(ya),as.numeric(xa)),
                  digits=3))
    assoc(xa,ya,mtxt,"Observer","Park",initials)

} #end: if (park=="all")

dev.off()

#display focal-species patterns - page 1

page <- 1
pdf(file=paste(outfile,".plots",page,".pdf",sep=""))

par(mfrow=c(2,2),mar=c(4,4,1.5,1)+0.1)

if (park=="all") {

    #count by park ANOVA
    xa <- droplevels(Park[hour>0]); ya <- y
    mtxt <- paste(species,
                  " x park ANOVA p = ",round(aov.p.val(ya,xa),digits=3),sep="")
    assoc(xa,ya,mtxt,"Park","Count")

} #end: if (park=="all")

#ct x stratum
xa <- droplevels(strat); ya <- y
mtxt <- paste(species," x stratum p = ",round(aov.p.val(ya,xa),digits=3),sep="")
assoc(xa,ya,mtxt,"Stratum","Count")

dev.off()
page <- 2
pdf(file=paste(outfile,".plots",page,".pdf",sep=""))

plotpoisson <- function(x,y,mtxt,xtxt,ytxt) {
    plot(jitter(x),jitter(y),main=mtxt,xlab=xtxt,ylab=ytxt,
         cex.main=1.5,cex.axis=1.5,cex.lab=1.5)
    i <- !is.na(x) #index used to limit null fit to same dataset as model fit
```

```
    mod <- glm(y[i]~x[i],family=poisson)
    x.rng <- seq(min(x[i]),max(x[i]),length.out=sum(i))
    y.pred <- exp(mod$coefficients[1]+mod$coefficients[2]*x.rng) #predict()
    lines(x.rng,y.pred,lwd=2,col="red")
    null <- glm(y[i]~1,family=poisson)
    null.dAIC <- AIC(mod,null)[2,2]-AIC(mod,null)[1,2]
    text(max(x.rng),max(y,na.rm=T)-0.25,col="red",cex=1,pos=2,
         paste("Null dAIC=",round(null.dAIC,digits=2),sep=""))
    p.val <- summary(mod)$coefficients[2,"Pr(>|z|)"]
    text(max(x.rng),max(y,na.rm=T)-0.5,col="red",cex=1,pos=2,
         paste("P=",round(p.val,digits=2),sep=""))
}

par(mfrow=c(2,2),mar=c(4,4,1.5,1)+0.1)

#count by pt-yr
tb <- table(y); labText <- "Count by point"
if (year=="all") {yr <- "-year"} else {yr <- paste(" (",year,")",sep="")}
labText <- paste(labText,yr,sep="")
bp <- barplot(tb,main=species,ylab="Frequency",cex.main=1.5,cex.axis=1.5,
              cex.lab=1.5,cex.names=1.5,xlab=labText)
for (i in 0:max(tb)) {text(bp[i+1],0.9*max(tb),tb[i+1],cex=0.9)}

#count by noise
plotpoisson(noise,y,"","Noise (jittered)",paste(species,"count (jittered)"))

#count by pt-yr-day
plotpoisson(day,y,"","Day (jittered)",paste(species,"count (jittered)"))

#count by pt-yr-hour
plotpoisson(hour,y,"","Hour (jittered)",paste(species,"count (jittered)"))

dev.off()
page <- 3
pdf(file=paste(outfile,".plots",page,".pdf",sep=""))

#raw count by park and year
par(mfrow=c(1,2),mar=c(4,4,1,1)+0.1)
yraw <- mutate(data.frame(y),pkyr=paste(Park,Year,sep="."))
by_pkyr <- group_by(yraw,pkyr) #group by park-yr
by_pkyr_v <- summarize(by_pkyr,toty=sum(y,na.rm=T))
toty <- by_pkyr_v$toty; toty
yraw_yr <- mutate(data.frame(y),Year)
by_yr <- group_by(yraw_yr,Year) #group by yr
by_yr_v <- summarize(by_yr,toty_yr=sum(y,na.rm=T))
toty_yr <- by_yr_v$toty_yr; toty_yr
ht <- 1.5*max(toty_yr,na.rm=T)
plot(1:10,toty_yr,ylim=c(0,ht),xlab="Year",ylab=paste(species,"unadjusted count"),
     pch="",cex.lab=1.5,cex.axis=1.25)
lines(1:10,toty_yr,lwd=2)
if (park=="all") {
    lines(1:10,toty[1:10],lty=2,col="red")
    lines(1:10,toty[11:20],lty=1,col="green3")
    lines(1:10,toty[21:30],lty=4,col="blue")
    legend(3,ht,legend=c("Total","MORA","NOCA","OLYM"),lwd=c(2,1,1,1),
           lty=c(1,2,1,4),col=c("black","red","green3","blue"),bty="n",cex=1.5)
} else {
    text(3,ht,park,cex=1.5)
} #end: if (park=="all")

#effort-adjusted count by park and year (count per pt surveyed omitting interval 3)
datprime <- select(dat,y1,y2,Day,Park_code,Event_year,ptyr) #omit interval 3
datprime <- mutate(datprime,yprime=y1+y2,pkyr=paste(Park_code,Event_year,sep="."))
```

```
#group by pt and then by park to avoid over-counting pts surveyed
by_ptyr_prime <- group_by(datprime,ptyr) #group by pt 1st
by_ptyr_v_prime <- summarize(by_ptyr_prime,
                             dayprime=first(Day),
                             yprime=sum(yprime,na.rm=T),
                             pkyr=first(pkyr),
                             yearprime=first(Event_year))
by_pkyr <- group_by(by_ptyr_v_prime,pkyr) #group by park 2nd
by_pkyr_v <- summarize(by_pkyr,
                       yprime=sum(yprime,na.rm=T),
                       surveyed_pts=sum(dayprime>0,na.rm=T),
                       yearprime=first(yearprime))
ct_pkyr <- by_pkyr_v$yprime #total count by park-year
pts_pkyr <- by_pkyr_v$surveyed_pts #total pts surveyed by park-year
tot_pk <- ct_pkyr/pts_pkyr
by_yr_prime <- group_by(by_pkyr_v,yearprime) #group by year to graph
by_yr_v_prime <- summarize(by_yr_prime,
                           totct=sum(yprime,na.rm=T),
                           totpts=sum(surveyed_pts,na.rm=T))
toty <- by_yr_v_prime$totct/by_yr_v_prime$totpts
ht <- 1.5*max(tot_pk,na.rm=T)
plot(1:10,toty,ylim=c(0,ht),xlab="Year",
     ylab=paste(species,"effort-adjusted count per point"),
     pch="",cex.lab=1.5,cex.axis=1.25)
lines(1:10,toty,lwd=2)
if (park=="all") {
    lines(1:10,tot_pk[1:10],lty=2,col="red")
    lines(1:10,tot_pk[11:20],lty=1,col="green3")
    lines(1:10,tot_pk[21:30],lty=4,col="blue")
    legend(3,ht,legend=c("Total","MORA","NOCA","OLYM"),lwd=c(2,1,1,1),
           lty=c(1,2,1,4),col=c("black","red","green3","blue"),bty="n",cex=1.5)
} else {
    text(3,ht,park,cex=1.5)
} #end: if (park=="all")

dev.off()
page <- 4
pdf(file=paste(outfile,".plots",page,".pdf",sep=""))

#count by year
par(mfrow=c(1,1),mar=c(4,4,1.5,1)+0.1)
y_prime <- by_ptyr_v_prime$yprime[hour>0] #defined for plot above
par(mfrow=c(1,1),mar=c(4,4,1,1)+0.1)
plotpoisson(Year,y_prime,"","Year (jittered)",
            paste(species,"effort-adjusted point counts (jittered)"))

dev.off()
page <- 5
pdf(file=paste(outfile,".plots",page,".pdf",sep=""))

par(mfrow=c(2,2),mar=c(4,4,1.5,1)+0.1)

#occ x forest
occ <- y>0; occ <- as.numeric(occ)
xa <- forest; ya <- occ
mtxt <- paste(species,"x forest ANOVA p =",round(aov.p.val(ya,xa),digits=3))
assoc(xa,ya,mtxt,"Forested","Occupied")

#occ x dense
xa <- dense; ya <- occ
mtxt <- paste(species,"x dense cover ANOVA p =",round(aov.p.val(ya,xa),digits=3))
assoc(xa,ya,mtxt,"Dense cover","Occupied")
```

```
#count by elevation
#no need for "hour>0" here, because y contains NAs
plotpoisson(elev,y,"","Elevation (jittered)",paste(species,"count (jittered)"))

#count by slope
plotpoisson(slope,y,"","Slope (jittered)",paste(species,"count (jittered)"))

dev.off()
page <- 6
pdf(file=paste(outfile,".plots",page,".pdf",sep=""))

par(mfrow=c(2,2),mar=c(4,4,1.5,1)+0.1)

#count by pt-yr-aspect (Beers)
#this regression is ok because y contains NAs (no pseudorep of zeros)
plotpoisson(aspect.b,y,"","Aspect (Beers transform, jitter)",
            paste(species,"count (jittered)"))
text(0.25,0.5,"SW",col="red"); text(1.75,0.5,"NE",col="red")

#count by pt-yr-aspect (untransformed)
plot(jitter(aspect.d),jitter(y),main="",
     xlab="Aspect (degrees, jitter)",ylab=paste(species,"count (jittered)"),
     cex.main=1.5,cex.axis=1.5,cex.lab=1.5)

#count by PAS anomaly
plotpoisson(PASanom,y,"","Lag-1 PAS anomaly (mm, jitter)",
            paste(species,"count (jittered)"))

#count by MST anomaly
plotpoisson(MSTanom,y,"","MST anomaly (C, jitter)",
            paste(species,"count (jittered)"))

dev.off()
page <- 7
pdf(file=paste(outfile,".plots",page,".pdf",sep=""))

#find tinterval = time-of-removal per detection - function used below
get.first <- function(x) {
  if (sum(x,na.rm=T)>0) {posn <- min(which(x!=0))} else {posn <- NA}
}

#plot effects of distance and time

# nbreaks = number of distance bins
nbreaks <- 5 #this is the number Andy Royle suggested

#find a sample for which detection distance and interval are independent
samp <- 0.9 #initial sample = 90% of all detections at this pt-yr

repeat{

  #find dclass = distance class per detection
  d <- dat$Obs_distance_m
  par(mfrow=c(2,2),mar=c(4,4,1.5,1)+0.1)
  hist(d,breaks=8,main=paste(species,year,"detections"),cex.lab=1.5,
       xlab="Detection distance (m)")
  maxd <- trunc(quantile(d,probs=samp,type=8,na.rm=T)); maxd
  if (bins=="equalDensity") {
    breaks <- quantile(d[d<=maxd],probs=seq(0,1,(1/nbreaks)),type=7,na.rm=TRUE)
    breaks[1] <- 0
  }
  if (bins=="equalArea") { #radii=sqrt((x/nbreaks)*maxd^2) where x = number of bins
    breaks <- trunc(c(0,sqrt((1:nbreaks)/nbreaks*(maxd^2))))
```

```
  }
  if (bins=="equalWidth") {
    breaks <- round(seq(0,maxd,maxd/nbreaks))
  }
  dclass <- cut(d,breaks,labels=F)
  areas <- pi*((breaks[2:(nbreaks+1)]^2)-(breaks[1:nbreaks]^2))

  #freq of detections in each dclass
  tb <- table(dclass)
  bp <- barplot(tb,width=areas,ylim=c(0,1.2*max(tb)),ylab="Number of detections",
                cex.main=1.5,cex.axis=1.5,cex.lab=1.5,cex.names=1.5,
                xlab="Distance class")
  for (i in 1:nbreaks) {text(bp[i],0.15*max(tb),round(breaks[i+1]),cex=1)}
  for (i in 1:nbreaks) {text(bp[i],0.1*max(tb),"m",cex=1)}
  #pct of detection-distances truncated to avoid fitting noise, non-independence
  yj <- subset(dat,select=c(y1,y2,y3)); yj[1:50,]
  truncPct <- sum(yj[d>maxd,],na.rm=T)/sum(yj,na.rm=T)*100; print(truncPct)
  text(round(bp[nbreaks]),(1.1*max(tb)),col="red",cex=1,pos=2,
       paste("Truncated ",round(truncPct),"% (>",maxd,"m)",sep=""))
  #detections per total (cumulative) area
  density <- cumsum(tb)/(pi*(breaks[2:(nbreaks+1)]^2))
  bp <- barplot(density,width=areas,ylab="Detections/cumulative area",cex.main=1.5,
                cex.axis=1.5,cex.lab=1.5,cex.names=1.5,xlab="Distance class")
  #detections per area of each class
  density <- tb/areas
  bp <- barplot(density,width=areas,ylab="Detections/class area",cex.main=1.5,
                cex.axis=1.5,cex.lab=1.5,cex.names=1.5,xlab="Distance class")

  #find tinterval = time-of-removal per detection
  f <- apply(yj,1,get.first)
  tinterval <- as.numeric(f) #plotted way below

  #find interval x distance association; looking for independence
  xa <- dclass; ya <- tinterval
  pval <- aov.p.val(ya,xa)
  mtxt <- paste("Assoc: interval x dist, p =",round(pval,digits=3))
  assoc(xa,ya,mtxt,"Detection distance class","Detection interval")

  #find slope of interval x distance; do not truncate distant detections if
  #detection interval is negatively related to distance
  IxD <- lm(ya~xa); slope <- coef(IxD)[2]

  if (pval > 0.05 | samp < 0.6 | slope < 0) {break} else {samp <- samp - 0.025}

} #repeat

dev.off()

if (samp < 0.6) {stop("Detection interval and distance are not independent")}

# delta = distance width for each bin (set to be equal intervals in this example)
delta <- breaks[-1] - breaks[-length(breaks)]; delta

# mdpts = midpoints of distance bins
mdpts <- breaks[-1] - delta/2; mdpts

#find nsurveys = number of points surveyed (less than or equal to rows of data)
nsurveys <- length(y); nsurveys #y=num of focal-sp indivs detected in each survey

# find nobs = number of times the species was detected (as an individual or group)
#NOTE: To use nobs the way Amundson did, I sort a data frame of dclass,
#tinterval and surveyid such that detections appear in the top 1:nobs rows;
#nobs is used to index (i <- 1:nobs) surveyid[i], dclass[i] and interval[i]
```

```
#for each observation of a species in the Amundson model, and surveyid[i]
#should store a number of unique ids equal to nsurveys, because it is used
#to fill pi.pd.c[j,k] and pi.pa.c[j,k], where k = 1:nsurveys, in the
#observation-level model, as:
#for(i in 1:nobs) {
#  dclass[i] ~ dcat(pi.pd.c[,surveyid[i]])
#  tinterval[i] ~ dcat(pi.pa.c[,surveyid[i]]) }

#set surveyid = point ID (site ID) for each individual (or group) detected
surveyid <- as.numeric(as.factor(dat$ptyr)); tail(surveyid)

#look at correlations between detection interval, hour, day and group size
hr <- dat$Hour
dy <- dat$Day
gs <- dat$Group_size
nobs <- sum(gs>0,na.rm=T); nobs
obsData <- data.frame(surveyid,dclass,tinterval,hr,dy,gs)
head(obsData); dim(obsData)
obsData <- obsData[order(dat$Group_size,decreasing=T),]; obsData[(nobs-2):(nobs+2),]
#the line above should show the transition between high survey-id #s and low ones
#if nobs is correct, because all detections are listed first, followed by
#non-detections
obsDat <- obsData[1:nobs,] #grab the data associated with detections
surveyid <- obsDat$surveyid
dclass <- obsDat$dclass
tinterval <- obsDat$tinterval
hr <- obsDat$hr
dy <- obsDat$dy
gs <- obsDat$gs
cor(obsDat[,-1],use="complete.obs")
cor.test(dy,hr,use="complete.obs",method="kendall")
cor.test(gs,dclass,use="complete.obs",method="kendall")
cor.test(dy,hr,use="complete.obs",method="kendall")
#other correlations are examined below

#display any dependence of time-of-removal on detection-distance

page <- 8
pdf(file=paste(outfile,".plots",page,".pdf",sep=""))

par(mfrow=c(2,2),mar=c(4,4,1.5,1)+0.1)

#interval by hour
xa <- hr; ya <- tinterval
mtxt <- paste("Assoc: interval x hour ANOVA p =",round(aov.p.val(ya,xa),digits=3))
assoc(xa,ya,mtxt,"Hour","Detection interval")
#note: even if the raw count is strongly related to hour, the model will find
#no effect of hour if the interval of detection is not related to hour

#display freq of detections in each time interval
tb <- table(tinterval); tb
bp <- barplot(tb,ylab="Number of detections",main=paste(species,year),cex.main=1.5,
              cex.axis=1.5,cex.lab=1.5,cex.names=1.5,xlab="Time interval",space=0.5)
nas <- is.na(yj); nas <- apply(nas,2,sum); nas
for (i in 1:3) {text(bp[i],0.9*max(tb),paste(nas[i],"NAs"),cex=0.75)}

#display interval x distance association (hidden in loop above)
xa <- dclass; ya <- tinterval
pval <- aov.p.val(ya,xa)
mtxt <- paste("Assoc: interval x dist, p =",round(pval,digits=3))
assoc(xa,ya,mtxt,"Detection distance class","Detection interval")

dev.off()
```

```
# ntrans = total number of transects
ntrans <- length(unique(dat$Site_code)); ntrans

# nyears = number of years/primary surveys
nyears <- totyrs

# nobservers = number of observers over the season/period modeled
nobservers <- length(initials)

# Nst = initial value for estimate of N - must be close to N or model will not run
Nst <- y+1; table(Nst); sum(is.na(Nst))

# J = number of time intervals within each point survey
J <- 3


# 2) run Amundson-type models but with variable-length count intervals
#    and trends fitted to data from multiple years and parks

require(jagsUI)

# R/JAGS code adapted from Amundson et al. (2014) for estimation of
# population trends from multi-region point-count data using an
# N-mixture model to fit covariate effects on trend and on each of
# two components of detection probability: availability based on
# time-removal data from intervals varying in length (Farnsworth
# et al. 2002), and perceptibility based on detection distance
#
# References:
# Amundson, C. L., J. A. Royle, and C. M. Handel. 2014. A hierarchical
#   model combining distance sampling and time removal to estimate
#   detection probability during avian point counts. The Auk 131:476–494.
# Farnsworth, G. L., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines,
#   and J. R. Sauer. 2002. A removal model for estimating detection
#   probabilities from point-count surveys. The Auk 119:414–425.

# VARIABLES
# y = count of birds per point (*associated w/surveyid, not tinterval)
# surveyid = survey point/site ID for each detection (individual or group)
# dclass = distance class per detection
# tinterval = time interval per detection
# nsurveys = number of points surveyed (*number of pt-yrs)
# nobs = number of individuals (or groups) detected
# delta = distance width for each bin (set as equal intervals in this example)
# nbreaks = number of distance bins
# mdpts = midpoints of distance bins
# maxd = maximum truncated distance (e.g., 200 m)
# J = number of time periods
# ... = standardized covariates (e.g., cover w/in 200m)
# ntrans = total number of transects
# tran = transect ID for each point
# day = ordinal date of survey
# Nst = initial value for estimate of N - must be close to N

# INDICES
# k in 1:nsurveys # surveys
# b in 1:nbreaks # distance bins
# j in 1:J # time intervals
# i in 1:nobs # detections (individual or group), each having a distance category
# t in 1:ntrans # transects

# PARAMETERS TO ESTIMATE
```

```
# meansig = mean scale parameter across sites (half normal shape in this example)
# meanpdet = mean probability of perceptibility
# meanpavail = mean probability of availability
# a0 = intercept for availability
# b.a1, b.a2, ... = coefficients of covars in availability model
# sigma0 = intercept for perceptibility
# b.p1, b.p2, ... = coefficients of covars in perceptibility model
# b1, b2, ... = coefficients of covars in abundance model
# meanN = mean site-level abundance
# mu.tran = mean abundance intercept across transects
# sd.tran = SD of random transect effect
# totN = population size of total area surveyed
# bayesp.pd = Bayesian p-value for pd model
# bayesp.pa = Bayesian p-value for pa model
# dens.ha = density of birds per hectare = totN/area surveyed

***# update file name as needed
file.name <- paste(species,"-fixPk-rndYr-rndTran-trend-climate.jagsUI",sep="")

cat("

    model {

    # PRIORS

    a0 ~ dnorm(0,0.01) # availability intercept
    b.a1 ~ dnorm(0,0.01) # coefs of availability covars (detection interval)
    b.a2 ~ dnorm(0,0.01)
    sigma0 ~ dunif(0,200) #scale of detection, bounded for covergence
    b.p1 ~ dnorm(0,0.01) # coefs of perceptibility covars (perception distance)
    b.p2 ~ dnorm(0,0.01)
    b0 ~ dnorm(0,0.01) # abundance intercept
    b1 ~ dnorm(0,0.01) # coefs of abundance covars
    b2 ~ dnorm(0,0.01)
    b3 ~ dnorm(0,0.01)
    b4 ~ dnorm(0,0.01)

    for (p in 1:3) { # park-specific effects
    parkeff[p] ~ dnorm(0,0.001)T(-15,15) # intercepts bounded for convergence
    for (l in 1:3) { # strata
    strateff[p,l] ~ dnorm(0,tau.strat[p,l]/wt[p,l]) # weighted stratum effect
    tau.strat[p,l] <- pow(sd.strat[p,l],-2)         # (to be implemented)
    sd.strat[p,l] ~ dunif(0,10)
    } # strata
    }

    for (p in 1:3) { # park-specific trends
    bp[p] ~ dnorm(0,0.001)T(-10,10) # slopes bounded for covergence
    }

    # random transect effect on abundance intercept nested pt w/in tran
    for (t in 1:ntrans) {
    traneff[t] ~ dnorm(0,tau.tran)
    }
    tau.tran <- pow(sd.tran,-2)
    sd.tran ~ dunif(0,10)

    # random year effect on abundance
    for (t in 1:nyears) {
    yeareff[t] ~ dnorm(0,tau.year)
    }
    tau.year <- pow(sd.year,-2)
    sd.year ~ dunif(0,10)
```

```
# random observer effect on perception
for (i in 1:(nobservers+1)) {
obseff[i] ~ dnorm(0,tau.obs)
}
tau.obs <- pow(sd.obs,-2)
sd.obs ~ dunif(0,10)

# random transect effect on perception
for (t in 1:ntrans) {
traneff.pd[t] ~ dnorm(0,tau.tr.pd)
}
tau.tr.pd <- pow(sd.tr.pd,-2)
sd.tr.pd ~ dunif(0,10)

# random year effect on perception
for (i in 1:10) {
yeareff.pd[i] ~ dnorm(0,tau.yr.pd)
}
tau.yr.pd <- pow(sd.yr.pd,-2)
sd.yr.pd ~ dunif(0,10)

# random pt-year effect on perception
for (k in 1:nsurveys) {
ptyreff.pd[k] ~ dnorm(0,tau.py.pd)
}
tau.py.pd <- pow(sd.py.pd,-2)
sd.py.pd ~ dunif(0,10)

# DETECTION PROBABILITY FUNCTIONS

for (k in 1:nsurveys) {

# covariates of availability and perceptibility***
log(sigma[k]) <- log(sigma0) #+ obseff[observer[k]] #+ ptyreff.pd[PTYR[k]]
logit(q[k]) <- a0 #+ b.a1*day[k] + b.a2*hour[k]

# distance-based estimation of detection probability

for (b in 1:nbreaks) {

log(g[b,k]) <- -mdpts[b]*mdpts[b]/(2*sigma[k]*sigma[k]) #half-normal
f[b,k] <- (2*mdpts[b]*delta[b])/(maxd*maxd) # bin widths vary***
pi.d[b,k] <- g[b,k]*f[b,k]  # p(detection) per pt-bin
pi.d.c[b,k] <- pi.d[b,k]/pd[k] # conditional form stdized by p(detected)

} # b in 1:nbreaks

pd[k] <- sum(pi.d[,k])  # pd is a sum over all dclass bins

# removal-based estimation of availability given unequal intervals

pi.a[1,k] <- 1-pow(q[k],3) # p(available in interval j=1)
pi.a[2,k] <- pow(q[k],3)*(1-pow(q[k],2)) # p(avail in j=2)
pi.a[3,k] <- pow(q[k],5)*(1-pow(q[k],2)) # p(avail in j=3)
# p(available in each interval j | available in at least one interval)
for (j in 1:J) {pi.a.c[j,k] <- pi.a[j,k]/pa[k]}
pa[k] <- sum(pi.a[,k]) # p(available in at least one interval)

} # k in 1:nsurveys

# OBSERVATION-LEVEL MODEL
```

```
for (i in 1:nobs) {

# single binomial trial with categorical dist linking dclass & tinterval to pt
dclass[i] ~ dcat(pi.d.c[,surveyid[i]]) # p(outcome = 1 to nbreaks dclasses)
tinterval[i] ~ dcat(pi.a.c[,surveyid[i]]) # p(outcome = 1 to J intervals)

} # i in 1:nobs

# ABUNDANCE ESTIMATION

for (k in 1:nsurveys) {

# counts as a function of number available and detection probability
y[k] ~ dbin(pd[k],n.a[k])
# number available for sampling as a function of abundance and p(available)
n.a[k] ~ dbin(pa[k],N[k])

# abundance model
N[k] ~ dpois(lambda[k]) # predicted abundance per survey (point-year)

# covariates of abundance***
log(lambda[k]) <- parkeff[Park[k]] + yeareff[Year[k]] + bp[Park[k]]*Year[k] +
                  traneff[tran[k]] + b2*PASanom[k] + b3*MSTanom[k]

} # k in 1:nsurveys

# GOODNESS OF FIT STATS

for (k in 1:nsurveys) {

n.a.fit[k] ~ dbin(pa[k],N[k]) # create new realization of model
y.fit[k] ~ dbin(pd[k],n.a[k])

e.pd[k] <- pd[k]*n.a[k] # original model prediction
E.pd[k] <- pow((y[k]-e.pd[k]),2)/(e.pd[k]+0.5) #dev from obs
E.New.pd[k] <- pow((y.fit[k]-e.pd[k]),2)/(e.pd[k]+0.5) #dev from pred

e.pa[k] <- pa[k]*N[k]
E.pa[k] <- pow((n.a[k]-e.pa[k]),2)/(e.pa[k]+0.5)
E.New.pa[k] <- pow((n.a.fit[k]-e.pa[k]),2)/(e.pa[k]+0.5)

} # k in 1:nsurveys

fit.pd <- sum(E.pd[]) #dev from obs
fit.new.pd <- sum(E.New.pd[]) #dev from pred

fit.pa <- sum(E.pa[])
fit.new.pa <- sum(E.New.pa[])

# SUMMARY STATS

mupavail <- mean(pa[]) # mean probability of availability
mupdet <- mean(pd[]) # mean probability of perceptibility
musigma <- mean(sigma[]) # mean scale parameter across sites
bayesp.pd <- step(fit.new.pd-fit.pd) # p-value for perceptibility model
bayesp.pa <- step(fit.new.pa-fit.pa) # p-value for availability model
totN.m <- sum(N[mora])  # population size of area surveyed in MORA
totN.n <- sum(N[noca])
totN.o <- sum(N[olym])
totN <- sum(N[])
for (i in 1:10) { # N per year and park-year
N.yr.m[i] <- sum(N[yr.m[,i]]) #in MORA
N.yr.n[i] <- sum(N[yr.n[,i]])
```

```
    N.yr.o[i] <- sum(N[yr.o[,i]])
    N.yr[i] <- sum(N[yr[,i]])
    }
    muN.mora <- mean(N[mora]) # mean point-level abundance in MORA
    muN.noca <- mean(N[noca])
    muN.olym <- mean(N[olym])
    muN <- mean(N[])
    ha <- maxd*maxd*3.14159/10000 # hectares surveyed at each pt
    dens.ha.m <- muN.mora/ha # density
    dens.ha.n <- muN.noca/ha
    dens.ha.o <- muN.olym/ha
    dens.ha <- mean(N[])/ha

    } #end: model

    ",file=file.name) #end: cat

observer <- as.numeric(droplevels(obs))
observer[is.na(observer)] <- max(observer,na.rm=T)
sum(is.na(observer)); unique(observer); hist(observer)
PTYR <- as.numeric(as.factor(PTYR))

#create some useful indices to pass to jags
yr <- matrix(0,length(Year)/10,10); dim(yr); str(yr)
for (i in 2005:2014) {yr[,(i-2004)] <- which(Year==i)}
if (park=="all") {
    mora <- which(Park=="MORA"); noca <- which(Park=="NOCA")
    olym <- which(Park=="OLYM")
    yr.m <- matrix(0,length(which(Park=="MORA"))/10,10); dim(yr.m)
    yr.n <- matrix(0,length(which(Park=="NOCA"))/10,10); dim(yr.n)
    yr.o <- matrix(0,length(which(Park=="OLYM"))/10,10); dim(yr.o)
    for (i in 2005:2014) {
      yr.m[,(i-2004)] <- which(Year==i&Park=="MORA")
      yr.n[,(i-2004)] <- which(Year==i&Park=="NOCA")
      yr.o[,(i-2004)] <- which(Year==i&Park=="OLYM")
    }
} #end: if (park=="all")

win.data <- list(y=y,
                   surveyid=surveyid,
                   maxd=maxd,
                   dclass=dclass,
                   tinterval=tinterval,
                   nsurveys=nsurveys,
                   nobs=nobs,
                   nbreaks=nbreaks,
                   delta=delta,
                   mdpts=mdpts,
                   J=J,
                   ntrans=ntrans,
                   tran=as.numeric(tran),
                   nobservers=nobservers,
                   observer=observer, #as.numeric(droplevels(obs)),
                   PTYR=PTYR,
                   day=as.vector(scale(day)),
                   forest=forest,
                   dense=dense,
                   elev=as.vector(scale(elev)),
                   noise=as.vector(scale(noise)),
                   hour=as.vector(scale(hour)),
                   slope=as.vector(scale(slope)),
                   aspect=as.vector(scale(aspect.b)),
                   Park=as.numeric(Park), # factors 1=MORA, 2=NOCA, 3=OLYM
```

```
                   nyears=nyears,
                   Year=Year-2004,
                   PASanom=as.vector(scale(PASanom)),
                   MSTanom=as.vector(scale(MSTanom)),
                   yr=yr, # indices for Year
                   mora=mora, noca=noca, olym=olym, # indices for N in each park
                   yr.m=yr.m, yr.n=yr.n, yr.o=yr.o #indices for Park-Year
)

inits <- function() list(N=Nst) #,a0=runif(1,-1,1))

params <- c("musigma",
              "mupdet",
              "mupavail",
              "bayesp.pd",
              "bayesp.pa",
              "sigma0",
              #"obseff",
              #"sd.obs",
              #"sd.py.pd",
              #"sd.yr.pd",
              #"sd.tr.pd",
              #"b.p1", #"b.p2",
              "a0",
              #"b.a1", #"b.a2",
              #"b0","b1",
              "b2","b3", #"b4",
              "parkeff",
              "yeareff",
              "sd.year",
              "bp",
              "muN.mora","muN.noca","muN.olym",
              "sd.tran",
              "totN.m","totN.n","totN.o",
              "totN",
              "dens.ha.m","dens.ha.n", "dens.ha.o",
              "dens.ha",
              "N.yr.m","N.yr.n","N.yr.o",
              "fit.pd","fit.new.pd","fit.pa","fit.new.pa",
              "N.yr"
)

ni <- 40000
nt <- 30
nb <- 10000
nc <- 3

***#update jags output file name as needed
system.time(OSFLout1.jags <- jags(win.data,inits,params,file.name,n.chains=nc,
            n.thin=nt,n.iter=ni,n.burnin=nb,parallel=TRUE))

sink(file=paste(outfile,".txt",sep=""),split=T)
print("Fitted models of detection, abundance")
print(paste("Maximum detection distance (maxd) =",maxd))
print(paste("% discarded to ensure independence between detection interval”,
            “& distance (truncPct) =",truncPct))
print(paste("Finding global effects and trends with climate effects"))

***#update models as needed
print("#log(lambda[k]) <- parkeff[Park[k]] + yeareff[Year[k]] + bp[Park[k]]*Year[k] +
       traneff[tran[k]] + b2*PASanom[k] + b3*MSTanom[k] # + b4*PASanom[k]*MSTanom[k]")
print("#log(sigma[k]) <- log(sigma0)") #+ ptyreff.pd[PTYR[k]]") #+ obseff[observer[k]]
print("#logit(q[k]) <- a0")
```

```
***#update jags output file name as needed
print(OSFLout1.jags,digits=3)
sink()

#evaluate fit
e.obs.pd <- OSFLout1.jags$sims.list$fit.pd
e.pred.pd <- OSFLout1.jags$sims.list$fit.new.pd
mu.bayespd <- mean(OSFLout1.jags$sims.list$bayesp.pd)
e.obs.pa <- OSFLout1.jags$sims.list$fit.pa
e.pred.pa <- OSFLout1.jags$sims.list$fit.new.pa
mu.bayespa <- mean(OSFLout1.jags$sims.list$bayesp.pa)
mean(e.pred.pd > e.obs.pd) # SA bayesp.pd
mean(e.obs.pd/e.pred.pd)
mean(e.pred.pa > e.obs.pa) # SA bayesp.pd
mean(e.obs.pa/e.pred.pa)
par(mfrow=c(2,2))
ppi <- 300
jpeg(file=paste(outfile,".fitStats-%d.jpg",sep=""),width=6*ppi,height=6*ppi,res=ppi)
plot(e.obs.pd,e.pred.pd,frame.plot=F,main=paste(species,"detected"),cex.lab=1.25,
     xlab="Discrepancy of observed data",ylab="Discrepancy of simulated data")
abline(0,1,lwd=2)
mu.obs <- mean(e.obs.pd); mu.pred <- mean(e.pred.pd)
legend(min(e.obs.pd),max(e.pred.pd),title="Mean discrepancy",bty="n",
       legend=c(paste("Obs:",round(mu.obs)),paste("Sim:",round(mu.pred))))
legend(max(e.obs.pd),min(e.pred.pd),xjust=1,yjust=0,bty="n",
       legend=c(paste("Bayes p:",round(mu.bayespd,3)),
       paste("LOF:",round(mean(e.obs.pd/e.pred.pd),3))))
plot(e.obs.pa,e.pred.pa,frame.plot=F,main=paste(species,"available"),cex.lab=1.25,
     xlab="Discrepancy of observed data",ylab="Discrepancy of simulated data")
abline(0,1,lwd=2)
mu.obs <- mean(e.obs.pa); mu.pred <- mean(e.pred.pa)
legend(min(e.obs.pa),max(e.pred.pa),title="Mean discrepancy",bty="n",
       legend=c(paste("Obs:",round(mu.obs)),paste("Sim:",round(mu.pred))))
legend(max(e.obs.pa),min(e.pred.pa),xjust=1,yjust=0,bty="n",
       legend=c(paste("Bayes p:",round(mu.bayespa,3)),
       paste("LOF:",round(mean(e.obs.pa/e.pred.pa),3))))
dev.off()

pdf(file=paste(outfile,"Chains.pdf",sep="")) #save all plots in 1 hi-res file
plot(OSFLout1.jags,ask=FALSE)
print("ok") #could use 30 of these to handle "Hit <Return> to see next plot:"
dev.off()
```

Appendix 5: Summary of park-structured models without climate covariates, including exploration of models based on covariates of detection

This summary of preliminary results for four species is an excerpt from the unpublished report "NCCN_landbird_trends_2005-2014_preliminary_summary.docx", on file with the NPS NCCN.

NCCN Landbird Trends 2005-2014: Preliminary Summary

Prepared by Chris Ray, The Institute for Bird Populations, cray@birdpop.org

Executive summary

Here, we document the results of our first model fits, showing raw and effort-adjusted trends in counts for each of several species, followed by modeled estimates of total population size across the sampled plots. In these preliminary models, we estimate detection probability including observer effects for some species, and explore temporal trends in conjunction with fixed effects of park, random effects of year, park-by-year interactions, and random effects of transect. Park-by-year interactions were not generally supported, and trends were rare among the commonly observed species.

Counts

Raw counts were affected by interannual variation in the number of point-counts and the number of counts per point: 2 counts per point during 2005-2010, 3 counts per point during 2011-2014. For each species, unadjusted (raw) annual counts are plotted next to effort-adjusted annual counts per point (adjusted for number of point-counts and counts per point). Years 1-10 represent 2005-2014.

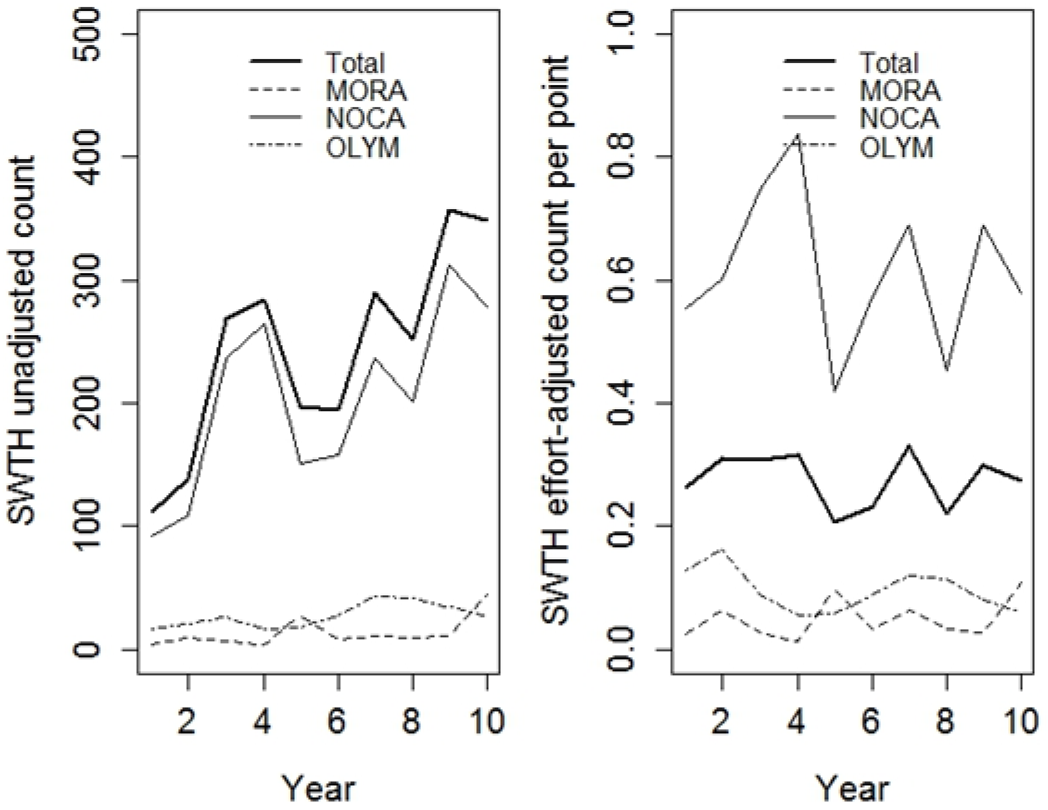
Detection models

Detection was modeled in two parts: availability was estimated via removal sampling (birds detected earlier are presumably more available for detection), and perceptibility was estimated via distance sampling (birds detected at greater distances are presumably more perceptible). At the level of individual parks, we found little support for effects of day and hour on availability, or for effects of observer on perceptibility. Observer and day are often confounded with park. Adding an observer effect at the 3-park level made little to no improvement over an intercept-only model, even when the intercept-only model was inadequate.

Abundance models

In the models summarized here, abundance (N) was estimated only for the points surveyed, and only for the area around each point in which the species was relatively perceptible. For the species presented here, the radius of this area varied from 30 m (RUHU) to 250 m (OSFL). Raw total counts don't approach Np (where p is detection probability) because N includes estimates for points that were not surveyed each year (including alternating panels). Mean annual N is $1/10^{th}$ of "totN" in the model outputs here, because totN is summed across 10 years.

Swainson’s thrush – raw and effort-adjusted data



Swainson’s thrush – model results

Summary: The simplest detection model (no covariates) appears adequate according to Bayesian p-values. The scale parameter for the half-normal detection function was estimated at 47.383 (m). Detection probability was 0.466 within 90 m of the survey point. Species availability was high. Park, year and transect effects on abundance were supported, but there was no support for a temporal trend overall or for effects of a park-by-year interaction. Species density per hectare was estimated at 0.254, for a mean annual population size (w/in 90 m of all survey points, annual and alternating) of 1892 birds.

Abundance model:

log(lambda[k]) <- parkeff[Park[k]]+yeareff[Year[k]]+traneff[tran[k]]+b1*Year[k]

Detection models:

log(sigma[k]) <- log(sigma0)

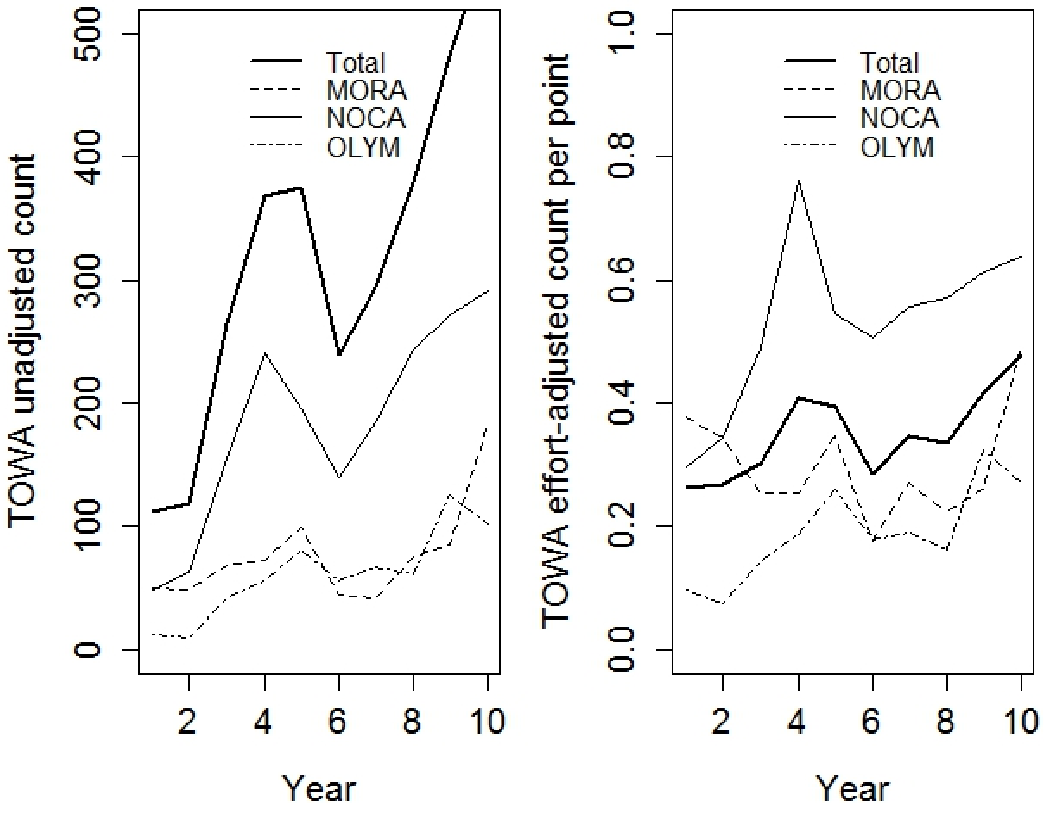
logit(q[k]) <- a0

Parameter estimates:

JAGS output for model 'SWTH-fixPk-rndYr-rndTran-trend', generated by jagsUI. Estimates based on 3 chains of 25000 iterations, burn-in = 5000 iterations and thin rate = 5, yielding 12000 total samples from the joint posterior.

SWTH	**mean**	**sd**	**2.50%**	**50%**	**97.50%**	**overlap0**	**F**	**Rhat**	**n.eff**
meansig	47.383	1.135	45.31	47.301	49.854	FALSE	1	1	4214
meanpdet	0.466	0.014	0.44	0.465	0.497	FALSE	1	1	4216
meanpavail	0.948	0.005	0.938	0.949	0.958	FALSE	1	1	9467
a0	0.639	0.04	0.561	0.638	0.717	FALSE	1	1	9384
sigma0	47.383	1.135	45.31	47.301	49.854	FALSE	1	1	4214
parkeff[1]	-4.284	0.454	-5.184	-4.28	-3.386	FALSE	1	1.016	240
parkeff[2]	-0.489	0.303	-1.094	-0.489	0.1	TRUE	0.946	1.008	243
parkeff[3]	-3.647	0.383	-4.428	-3.638	-2.925	FALSE	1	1.034	66
b1	0.008	0.021	-0.035	0.008	0.05	TRUE	0.671	1.001	2628
meanN	0.646	0.027	0.59	0.647	0.697	FALSE	1	1.005	456
sd.tran	2.146	0.183	1.806	2.138	2.513	FALSE	1	1.072	33
sd.year	0.171	0.063	0.083	0.159	0.327	FALSE	1	1.003	1225
totN	18915	777	17261	18938	20395	FALSE	1	1.005	456
bayesp.pd	0.37	0.483	0	0	1	TRUE	1	1	12000
bayesp.pa	0.478	0.5	0	0	1	TRUE	1	1	7390
dens.ha	0.254	0.01	0.232	0.254	0.274	FALSE	1	1.005	456
deviance	14878	104	14653	14883	15072	FALSE	1	1.002	1343

Townsend's warbler – raw and effort-adjusted data



Townsend's warbler – model results

Summary: The simplest detection model (especially the perceptibility model) might not be adequate. The scale parameter for the half-normal detection function was estimated at 44.447 (m). Detection probability was 0.428 within 90 m of the survey point. Species availability was high. Park, year and transect effects on abundance were supported, as well as a positive temporal trend overall, but there was no support for effects of a park-by-year interaction (not shown). Species density per hectare was estimated at 0.378, for a mean annual population size (w/in 90 m of all survey points, annual and alternating) of 2812 birds.

Abundance model:

log(lambda[k]) <- parkeff[Park[k]]+yeareff[Year[k]]+traneff[tran[k]]+b1*Year[k]

Detection models:

log(sigma[k]) <- log(sigma0)

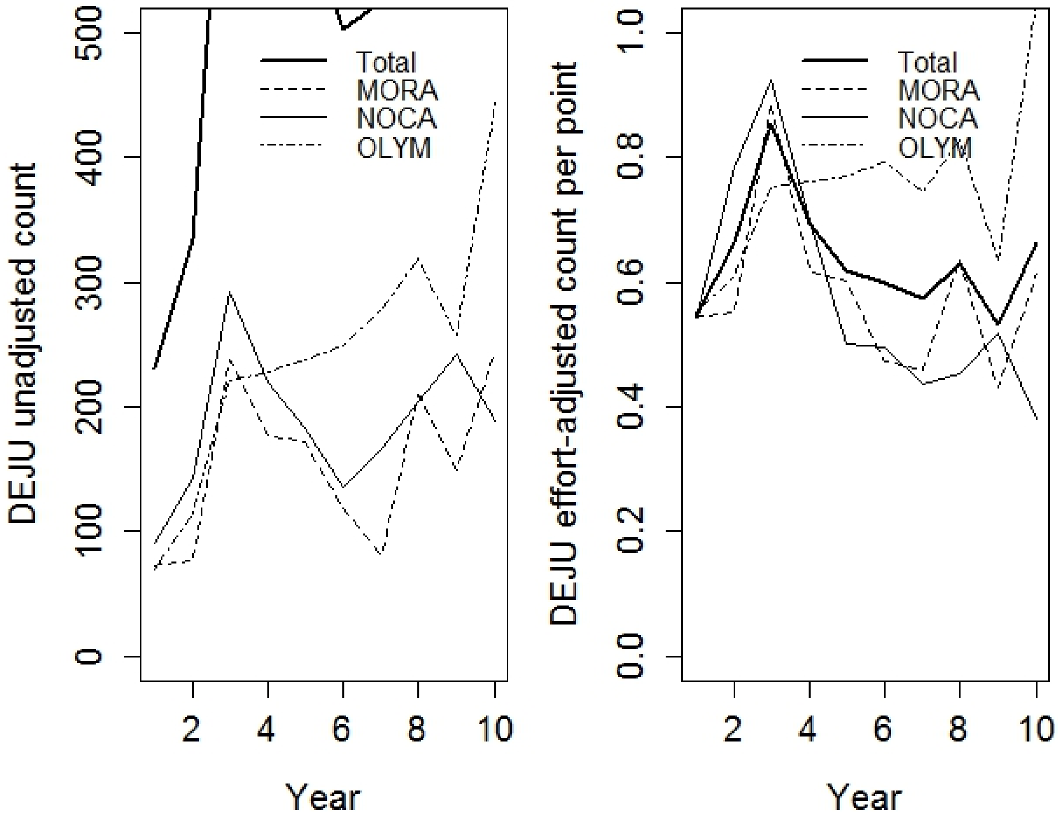
logit(q[k]) <- a0

Parameter estimates:

JAGS output for model 'TOWA-fixPk-rndYr-rndTran-trend', generated by jagsUI. Estimates based on 3 chains of 25000 iterations, burn-in = 5000 iterations and thin rate = 5, yielding 12000 total samples from the joint posterior.

TOWA	Mean	Sd	2.50%	50%	97.50%	overlap0	f	Rhat	n.eff
meansig	44.447	0.776	42.984	44.419	46.076	FALSE	1	1.02	109
meanpdet	0.428	0.01	0.408	0.428	0.45	FALSE	1	1.02	108
meanpavail	0.968	0.003	0.962	0.968	0.973	FALSE	1	1	3336
a0	0.452	0.033	0.387	0.452	0.518	FALSE	1	1.001	3197
sigma0	44.447	0.776	42.984	44.419	46.076	FALSE	1	1.02	109
parkeff[1]	-1.686	0.273	-2.233	-1.681	-1.167	FALSE	1	1.01	228
parkeff[2]	-0.356	0.234	-0.811	-0.357	0.1	TRUE	0.937	1.002	1195
parkeff[3]	-2.398	0.255	-2.908	-2.397	-1.904	FALSE	1	1.01	201
b1	0.04	0.015	0.01	0.04	0.071	FALSE	0.992	1.001	2314
meanN	0.961	0.033	0.896	0.96	1.028	FALSE	1	1.018	139
sd.tran	1.684	0.115	1.472	1.68	1.92	FALSE	1	1.017	139
sd.year	0.112	0.05	0.034	0.105	0.231	FALSE	1	1.001	12000
totN	28122	962	26233	28111	30094	FALSE	1	1.018	139
bayesp.pd	0.176	0.38	0	0	1	TRUE	1	1	12000
bayesp.pa	0.453	0.498	0	0	1	TRUE	1	1	5632
dens.ha	0.378	0.013	0.352	0.377	0.404	FALSE	1	1.018	139
deviance	20374	108	20139	20378	20584	FALSE	1	1.038	60

Dark-eyed junco – raw and effort-adjusted data



Dark-eyed junco – model results

Summary: The simplest detection model was not adequate even after including an observer effect. The scale parameter for the half-normal detection function was estimated at 33.83 (m). Detection probability was 0.277 within 90 m of the survey point. Species availability was high. Park, year and transect effects on abundance are supported, but there was no support for temporal trend or for effects of a park-by-year interaction (not shown). Species density per hectare was estimated at 1.014, for a mean annual population size (w/in 90 m of all survey points, annual and alternating) of 7549 birds.

Abundance model:

log(lambda[k]) <- parkeff[Park[k]]+yeareff[Year[k]]+traneff[tran[k]]+b1*Year[k]

Detection models:

log(sigma[k]) <- log(sigma0)+obseff[observer[k]]

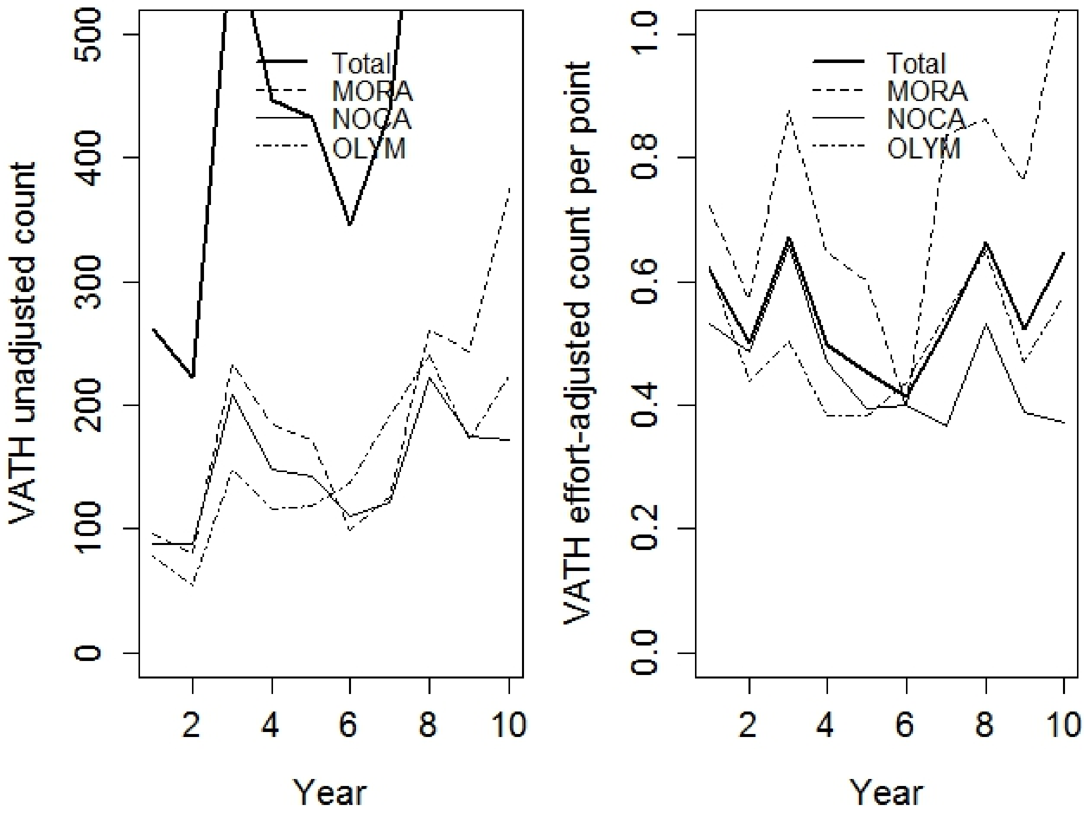
logit(q[k]) <- a0

Parameter estimates:

JAGS output for model 'DEJU-fixPk-rndYr-rndTran-trend-obs', generated by jagsUI. Estimates based on 3 chains of 30000 iterations, burn-in = 10000 iterations and thin rate = 10, yielding 6000 total samples from the joint posterior.

DEJU	**mean**	**sd**	**2.50%**	**50%**	**97.50%**	**overlap0**	**f**	**Rhat**	**n.eff**
meansig	33.83	0.887	32.186	33.798	35.606	FALSE	1	1	5778
meanpdet	0.277	0.013	0.254	0.277	0.303	FALSE	1	1	5789
meanpavail	0.936	0.004	0.929	0.936	0.944	FALSE	1	1	6000
a0	0.73	0.027	0.676	0.73	0.781	FALSE	1	1	6000
sigma0	35.816	0.597	34.677	35.805	37.028	FALSE	1	1.001	1237
yeareff[1]	-0.223	0.134	-0.492	-0.22	0.03	TRUE	0.96	1	4428
yeareff[2]	0.254	0.116	0.039	0.25	0.489	FALSE	0.989	1.004	533
yeareff[3]	0.148	0.098	-0.033	0.145	0.346	TRUE	0.949	1.003	579
yeareff[4]	0.006	0.087	-0.163	0.004	0.182	TRUE	0.522	1.001	1327
yeareff[5]	-0.153	0.08	-0.314	-0.151	0	FALSE	0.975	1.002	903
yeareff[6]	-0.073	0.084	-0.242	-0.071	0.094	TRUE	0.818	1.001	1494
yeareff[7]	0.007	0.092	-0.177	0.007	0.186	TRUE	0.534	1	6000
yeareff[8]	-0.037	0.096	-0.231	-0.037	0.147	TRUE	0.667	1.001	2001
yeareff[9]	-0.033	0.11	-0.252	-0.031	0.185	TRUE	0.627	1.002	1151
yeareff[10]	0.096	0.125	-0.15	0.098	0.338	TRUE	0.797	1.001	2434
parkeff[1]	0.621	0.191	0.245	0.623	0.993	FALSE	0.999	1.006	307
parkeff[2]	0.635	0.185	0.264	0.638	0.997	FALSE	0.999	1.003	777
parkeff[3]	0.907	0.18	0.547	0.911	1.249	FALSE	1	1.001	1780
b1	-0.017	0.023	-0.062	-0.017	0.03	TRUE	0.8	1.004	563
meanN	2.579	0.068	2.446	2.579	2.713	FALSE	1	1.009	417
sd.obs	0.083	0.013	0.06	0.082	0.111	FALSE	1	1.002	1488
sd.tran	0.825	0.052	0.729	0.823	0.932	FALSE	1	1.002	867
sd.year	0.189	0.069	0.1	0.175	0.358	FALSE	1	1.006	976
totN	75489	1987	71588	75491	79424	FALSE	1	1.009	417
bayesp.pd	0.058	0.234	0	0	1	TRUE	1	1.001	6000
bayesp.pa	0.452	0.498	0	0	1	TRUE	1	1	6000
dens.ha	1.014	0.027	0.961	1.014	1.066	FALSE	1	1.009	417
deviance	39776	107	39563	39777	39987	FALSE	1	1.005	570

Varied thrush – raw and effort-adjusted data



Varied thrush – model results

Summary: Detection distance was higher for this species, and the detection model was not adequate even after including an observer effect. The scale parameter for the half-normal detection function was estimated at 72.595 (m). Detection probability was 0.313 within 180 m of the survey point. Species availability was high. Park, year and transect effects on abundance are supported, but there was no support for a trend in abundance or for effects of a park-by-year interaction. Species density per hectare was estimated at 0.174, for a mean annual population size (w/in 180 m of all survey points, annual and alternating) of 5182.

Abundance model:

log(lambda[k]) <- parkeff[Park[k]]+yeareff[Year[k]]+traneff[tran[k]]+b1*Year[k]

Detection models:

log(sigma[k]) <- log(sigma0)+obseff[observer[k]]

logit(q[k]) <- a0

Parameter estimates:

JAGS output for model 'VATH-fixPk-rndYr-rndTran-trend-obs', generated by jagsUI. Estimates based on 3 chains of 30000 iterations, burn-in = 10000 iterations and thin rate = 10, yielding 6000 total samples from the joint posterior.

VATH	**mean**	**sd**	**2.50%**	**50%**	**97.50%**	**overlap0**	**f**	**Rhat**	**n.eff**
meansig	72.595	2.218	68.408	72.508	77.166	FALSE	1	1.004	476
meanpdet	0.313	0.016	0.283	0.312	0.346	FALSE	1	1.004	476
meanpavail	0.982	0.001	0.979	0.982	0.985	FALSE	1	1.002	1262
a0	0.254	0.027	0.201	0.255	0.307	FALSE	1	1.002	1234
sigma0	71.601	1.566	68.575	71.583	74.696	FALSE	1	1.007	362
obseff[1]	-0.032	0.084	-0.189	-0.033	0.139	TRUE	0.659	1.003	731
obseff[2]	-0.058	0.045	-0.143	-0.059	0.03	TRUE	0.902	1.008	285
obseff[3]	0.23	0.067	0.107	0.227	0.373	FALSE	1	1.004	435
…									
obseff[45]	0.009	0.045	-0.077	0.009	0.097	TRUE	0.58	1	6000
obseff[46]	-0.002	0.124	-0.241	-0.004	0.244	TRUE	0.514	1.001	2538
parkeff[1]	0.605	0.209	0.199	0.605	1.017	FALSE	0.997	1.001	2267
parkeff[2]	-0.472	0.21	-0.888	-0.473	-0.059	FALSE	0.988	1.001	1604
parkeff[3]	0.076	0.197	-0.302	0.073	0.472	TRUE	0.65	1.002	906
b1	0.015	0.023	-0.032	0.015	0.06	TRUE	0.758	1	3537
meanN	1.771	0.046	1.684	1.77	1.863	FALSE	1	1.016	163
sd.tran	1.037	0.072	0.905	1.034	1.189	FALSE	1	1.002	1072
sd.year	0.187	0.07	0.094	0.173	0.365	FALSE	1	1.002	4005
totN	51824	1350	49288	51804	54527	FALSE	1	1.016	163
bayesp.pd	0.022	0.146	0	0	0	FALSE	1	1.002	6000
bayesp.pa	0.466	0.499	0	0	1	TRUE	1	1	6000
dens.ha	0.174	0.005	0.165	0.174	0.183	FALSE	1	1.016	163
deviance	30545	102	30341	30546	30748	FALSE	1	1.008	304

Appendix 6: R script for Bayesian meta-analysis of parameter estimates averaged by species trait-group

```
# Meta-analysis to determine whether trends and climate effects
# differ by trait group

# Exerpted from code file name “script4traitGroupEffectsMetaAnal.R”; complete code
# including data prep is available from C. Ray (cray@birdpop.org

# meta-analysis of trend by migration habit (e.g., ndf$trend~ndf$use.NOCA)
library(jagsUI)
modelType <- "trendXmig-means"
file.name <- paste("metaAnal-",modelType,".jagsUI",sep="")
cat("
    model {

    # VARIABLES
    # stat = the statistic passed to the meta-analysis (e.g., trend
    #        or effect of PAS), previously estimated with error
    # sd = the standard deviation of previous estimates for stat
    # beta[i] = the mean of stat for trait group i (e.g., the mean
    #          trend for migrants)

    # PRIORS
    for (i in 1:2) { # for 1 = migrants and 2 = residents...
      beta[i] ~ dnorm(0,0.0001) # ...mean effects of migration habit
    }
    tauSp <- pow(sdSp,-2) # species residual
    sdSp ~ dunif(0,10)

    # LIKELIHOOD
    for (i in 1:n) { # for each of n species...
      stat[i] ~ dnorm(mu[i],tau.sd[i]) # measurement err model
      tau.sd[i] <- pow(sd[i],-2)
      mu[i] <- beta[mig[i]] + eps.sp[i] # full model
      eps.sp[i] ~ dnorm(0,tauSp) # residual err
    }

    } # end model
    ",file=file.name) # end cat

mig <- ndf$use.NOCA
levels(mig)[levels(mig)=="MIG"] <- 1 # migrants = 1 so the effect
levels(mig)[levels(mig)=="RES"] <- 2 # is a resident effect
jags.data <- list(mig=as.numeric(mig[-1]), # migration habit
                  stat=ndf$trend[-1],      # fitted sp trend in N
                  sd=trendNsd[-1],         # fitted sp sd in trend
                  n=length(trendNsd)-1)    # count of fitted sp trends
str(jags.data)

inits <- function() list(beta=rnorm(2))
params <- c("beta")
ni <- 12000; nt <- 10; nb <- 2000; nc <- 3
system.time(out <- jags(jags.data,inits,params,file.name,
                        n.iter=ni,n.thin=nt,n.burnin=nb,n.chains=nc))
sink(file="metaAnalOut.txt",split=T,append=T)
cat("trend x mig NOCA no climate","\n","\n")
print(out,digits=3)
cat("\n")
```

```
sink()

# (see “script4traitGroupMetaAnal.R” for additional examples and code)
```

NPS 963/139316, July 2017

National Park Service
U.S. Department of the Interior

Natural Resource Stewardship and Science
1201 Oakridge Drive, Suite 150
Fort Collins, CO 80525

www.nature.nps.gov

EXPERIENCE YOUR AMERICA ™