Applying occupancy estimation and modelling to the analysis of atlas data

Giancarlo Sadoti\textsuperscript{1,4}, Benjamin Zuckerberg\textsuperscript{2}, Marta A. Jarzyna\textsuperscript{3} and William F. Porter\textsuperscript{3}

\textsuperscript{1}Department of Geography, University of Nevada, Reno, NV, USA, \textsuperscript{2}Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, 1630 Linden Drive, Madison, WI, 53706-1598, USA, \textsuperscript{3}Department of Fisheries and Wildlife, Michigan State University, 480 Wilson Road, East Lansing, MI, 48824, USA

ABSTRACT

Aim Biological atlases are a globally widespread and effective means for documenting the distribution of numerous taxa and have been used to study many macroecological relationships. A common assumption when analysing atlas data is that species are detected perfectly ($p = 1$). This assumption is likely incorrect, but the application of methods to account for heterogeneous detectability ($p < 1$) has been difficult to implement. We provide an application of current methods of occupancy estimation and modelling to account for imperfect detection in the analysis of atlas data.

Location New York, USA.

Methods We employed multiseason occupancy models substituting spatial replicates for temporally repeated surveys to examine changes in distribution of the Canada Warbler (\textit{Cardellina canadensis}) using breeding bird atlases from 1980–1985 and 2000–2005. We compared estimates from models accounting for $p < 1$ versus those assuming $p = 1$ in assessing statewide patterns of occupancy, colonization and extinction.

Results We found forest cover, observer effort, information on previous detections and the sampling year were important predictors of detection. Environmental predictors of statewide occupancy dynamics were similar among models accounting for $p < 1$ versus those assuming $p = 1$. Despite these similarities, site-level estimates of occupancy from the model accounting for imperfect detection indicated 14\% and 19\% higher site occupancy in the 1980–1985 and 2000–2005 Atlases, respectively. In addition, relative to the model accounting for $p < 1$, the model assuming perfect detectability underestimated persistence and overestimated extinction between atlases. The model accounting for $p < 1$ had 0.3\% and 7.9\% higher accuracy in predicting occupancy in the 1980–1985 and 2000–2005 Atlases, respectively, than the models assuming $p = 1$.

Main conclusions Occupancy modelling and estimation can be successfully applied to broad-scale surveys, such as atlases, that do not explicitly implement repeated visits to a survey block. Occupancy modelling allows for a more rigorous analysis of atlas data for exploring species–environmental relationships and modelling species distributions while accounting for imperfect detection.

Keywords Citizen science, detectability, forest birds, New York, observer effort, occupancy modelling, spatial autocorrelation, species distribution.
from 0.06 to 14,400 km² (Gibbons et al., 2007), atlas data are increasingly used for assessing ecological responses to habitat fragmentation (Trzcinski et al., 1999; Zuckerberg & Porter, 2010), climate change (Thomas & Lennon, 1999; Brommer, 2004, 2008; Lemoine et al., 2007; Zuckerberg et al., 2009) and gradients of primary productivity (Coops et al., 2009). Many of these atlases, however, were developed and conducted before the formal introduction of occupancy estimation and modelling (MacKenzie et al., 2006). Consequently, analyses of most atlas data make the assumption that if a species occupies a survey block during the time of the atlas, it will be detected without error given reasonable survey effort. Unfortunately, when modelling species distributions using data from atlases or other sources, this assumption may have one of more undesired consequences including the following: (1) underestimation of species distributions, (2) overestimation of changes in species distribution, (3) biased estimates of environmental covariate relationships and (4) model predictions may incorporate systematic biases in detectability (Tyre et al., 2003; MacKenzie et al., 2006; Kéry, 2010; Rota et al., 2011). In this study, we provide an application of current methods of occupancy estimation and modelling to account for imperfect detection in the analysis of atlas data.

Contemporary methods of occupancy modelling and estimation generally depend on temporally replicated observations of a sampling location, but this replicated design is not explicit in the protocols of many past atlases. Although most atlas programs encourage repeated visits to survey blocks, this is not a requirement for data collection, and atlas participants (‘atlasers’ hereafter) are often instructed to avoid re-recording a species unless it is observed with a higher level of breeding confirmation. Space-for-time approaches have been suggested (Mackenzie & Royle, 2005) as a method of using spatially clustered sampling locations within a larger, primary sampling unit as a substitute for temporally replicated visits (Fig. 1). In these cases, there is a general, if not always necessary, assumption that all sampling locations within the primary unit have the same state of occupancy if sampled within a period of population closure (e.g. a breeding season; MacKenzie et al., 2006; Guillera-Arroita, 2011). Space-for-time approaches to modelling occupancy and estimation have been advocated when dictated by study logistics or objectives (Mackenzie & Royle, 2005) and have been applied in several recent studies (Francois et al., 2008; Sirami et al., 2008; Wenger et al., 2008; Anderson et al., 2012). Space-for-time approaches have also been applied in studies investigating species richness (Nichols et al., 1998; Jiguet et al., 2005) and, more recently, using a Bayesian approach (Royle & Kéry, 2007) in modelling dynamics of a single species. To our knowledge, using spatial replicates as a surrogate for repeated visits has not previously been applied to atlas data.

We demonstrate an approach that can be widely applied to many grid-based atlas programs such as those underway for herpetofauna and birds. To illustrate the use of this method, we model the dynamic occupancy patterns of the Canada Warbler (Cardellina canadensis), while accounting for imperfect detection, using the first and second New York State BBA (Andrle & Carroll, 1988; McGowan & Corwin, 2008). New York was the first state in the United States to have two repeated atlases, but many regions have completed or are in the process of completing their second atlas. Our objectives were to (1) demonstrate the usefulness of space-for-time occupancy estimation and modelling method in quantifying regional changes in the distributions of a forest songbird and (2) assess the application of this approach for quantifying broad-scale variation in detectability using atlas data. We met these objectives using a hierarchical occupancy model for quantifying species-specific patterns of detectability, occupancy and 20 years of distributional change between two atlas surveys.

**METHODS**

**Atlas Data (New York State Breeding Bird Atlas)**

The New York State Breeding Bird Atlas (hereafter BBA) is a comprehensive, statewide survey with the objective of documenting the distribution of breeding birds in New York. The BBA was conducted in two time periods. The first atlas project was conducted from 1980 to 1985 (hereafter 1980 Atlas; Andrle & Carroll, 1988), and the second atlas was conducted from 2000 to 2005 following the same protocol of the first BBA (hereafter 2000 Atlas; McGowan & Corwin, 2008). The BBA used a grid system, which stratified New York State (128,402 km²) into 1398 primary sampling units (hereafter sites) measuring 10 km × 10 km. These sites were divided into four 5 km × 5 km secondary samplings units (hereafter blocks) resulting in a total 5332 blocks (Fig. 1). The BBA implemented a set of protocols to achieve consistent and adequate coverage for each atlas survey (McGowan & Corwin, 2008). In brief, atlasers were assigned to survey one or more blocks and were instructed to spend at least 8 h in the block, visiting each habitat present and, with some.

![Figure 1](attachment:image.png)
exceptions, recording at least 76 species. Atlasers recorded evidence of breeding for the birds they observed using defined breeding codes within possible (e.g. species observed in appropriate habitat or singing male), probable (e.g. pair observed in suitable habitat in breeding season) and confirmed (e.g. nest with young) classifications. For each species, the highest level of confirmation of breeding recorded during the survey was retained in the final database. For the purposes of our analysis, we considered any level of breeding evidence to be a detection and summarized the atlas data as detections or non-detections (1 or 0) at the block scale. Atlasers recorded the amount of effort they spent atlasing as the number of survey hours multiplied by the number of atlasers (McGowan & Corwin, 2008). We considered each 5 km x 5 km block as a spatially replicated observation within larger 10 km x 10 km sites (Fig. 1). As such, we were interested in quantifying estimates of site-level occupancy and occupancy dynamics.

**Study species**

The Canada Warbler is forest-associated, Neotropical migrant species with a breeding range from the southern Appalachian Mountains through the north-east and north-central United States into southern Canada. It is associated with a wide variety of wooded habitats where it nests near or on the ground and forages in and from shrubs and lower tree branches (Reitsma et al., 2010). This species is one of conservation concern in parts of its range (Dettmers, 2003) and has experienced an overall 40-year decline in the United States, with greatest declines occurring in the north-eastern United States (Sauer et al., 2011).

We selected the Canada Warbler due to its (1) regional conservation concern, (2) intermediate level of observed occupancy and (3) association with a wide range of forested habitats that are well-described by remotely sensed land cover data. Forest passerines, including the Canada Warbler, are well-represented in New York and have been the focus of previous atlas-based studies on landscape-scale avian occupancy and dynamics in north-eastern North America (Trzcinski et al., 1999; Zuckerberg & Porter, 2010).

**Data subsampling**

To reduce biases due to potential violations of the assumption of population closure across blocks within sites (Kendall & White, 2009; Guillera-Arroita, 2011), we subsampled atlas data to remove blocks that had little or no habitat availability or had atlasing spread over several years within atlas periods. Given the habitat requirements of our study species (see Study Species), we addressed habitat availability by including only those blocks having at least 50% forest cover (MacKenzie et al., 2006; Guillera-Arroita, 2011). We addressed within-block temporal closure by retaining only those blocks in which at least 90% of all species detected in a block within an atlas were detected in one primary year of sampling. Blocks were removed if, by chance, the Canada Warbler was detected in a year outside the primary sampling year. We also removed any blocks lacking information on survey effort (total observer hours per block) because we assumed effort would strongly influence species detection probabilities (McGowan & Zuckerberg, 2008; Zuckerberg & Porter, 2010; Kujala et al., 2012).

**Sampling and environmental covariates**

In addition to the effects of effort, we suspected information on past detections available to atlasers (Andrle & Carroll, 1988; McGowan & Corwin, 2008) might inform the expectation of future detections, resulting in an increased probability of detecting Canada Warblers. This information came in two forms: (1) 1980 Atlas information was available to 2000 Atlas participants in the form of species lists, and (2) annual progress data produced by coordinators for each atlas presented as preliminary maps of occurrence. To address these potential biases, we generated two sampling variables. The first, detect80, was a binary covariate describing detection of Canada Warblers in the 1980 Atlas. We gave all blocks in the 1980 Atlas a value of 0, while blocks in the 2000 Atlas were assigned a value of 1 if the Canada Warbler was observed in the 1980 Atlas and 0 if not. The second sampling variable, neigh3, was a continuous covariate describing the proportion of blocks in a 3-block x 3-block neighbourhood having within-atlas detections in years prior to the year of sampling. Finally, we included the primary year of block sampling as a factor covariate in all models accounting for imperfect detection. This was performed to (1) allow us to relax the assumption of population closure within atlas (Kéry et al., 2010) and (2) account for potential variation in detection probabilities due to different atlasers or population responses (e.g. abundance) across years. In terms of environmental covariates that might influence detectability, we included the percentage of forest cover (forest) at the block level assuming that forest cover would positively influence the abundance of Canada Warblers and, in turn, their detectability (Betts et al., 2008; De Wan et al., 2009).

At the block level, we quantified elevation (ELEV), percentage forest cover (FOREST) and length of forest-non-forest edges (EDGE). We subsequently generated site-level means of these values from subsamples retained (described previously) for use in modelling site-level processes. We used block boundaries provided by the New York Department of Environmental Conservation to derive Universal Transverse Mercator (UTM) coordinates and we derived elevation from the national elevation data set (Gesch et al., 2002). We used the 2001 National Land Cover Data set (NLCD; Homer et al., 2007) to quantify forest cover. While the 1992 NLCD (Vogelmann et al., 2001) was conducted closer to the midpoint between atlases, the 2001 NLCD had higher overall classification accuracy (Wickham et al., 2010). Examinations of the NLCD change product (Fry et al., 2008) indicated only small changes in the percentage area covered by forest.
classes between the 1992 and 2001 NLCD when measured at the scale of an atlas block (mean = −0.5%, SE = 0.1). Although Canada Warblers are often associated with forested wetlands (Reitsma et al., 2010), we aggregated deciduous, mixed, evergreen and woody wetland NLCD classes into a generalized forest cover classification as woody wetland was poorly classified in the north-eastern United States (Wickham et al., 2010).

Canada Warblers are often absent from or in low densities in mature, undisturbed forests (Reitsma et al., 2010). While we lacked information on forest age or structure at the scale of New York State, we used the length of forest–non-forest edges as a covariate in occupancy models as we expected forest edge effects (McCullin, 1998) to be associated with higher occupancy and lower extinction probabilities among Canada Warblers at the site level (Hallworth et al., 2008). We removed the correlation ($r < −0.6$) between forest edge and forest cover using a simple linear regression (Trzcinski et al., 1999), retaining residuals as an index of forest edge uncorrelated with forest cover. We predicted Canada Warblers may be elevation-limited in the cooler high-elevation areas of New York and used a quadratic elevation effect in occupancy models. In considering occupancy dynamics, we expected lower extinction and higher colonization probabilities in sites having more forest, having longer forest-non-forest edge, and located in the more northern regions of New York (Zuckerberg et al., 2009; Zuckerberg & Porter, 2010).

**Occupancy modelling**

We used the likelihood-based occupancy modelling approach of MacKenzie et al. (2003) to model patterns of Canada Warbler distributional dynamics between the 1980 and 2000 Atlases. We modelled occupancy ($\psi$), colonization ($\gamma$), extinction ($\epsilon$) and detection ($p$) using a parsimonious set of environmental covariates (see section Environmental Covariates). We followed a hierarchical four-step approach to building a set of models (Kéry et al., 2010; McClure & Hill, 2012). We first constructed a global (all-covariate) model and (1) holding all other covariates constant, we sequentially removed covariates predicting $p$ to construct an initial set of 16 models, (2) using each competitive ($\Delta AIC < 2$) model from step one, we constructed an additional seven models by removing covariates predicting $\psi$, (3) using each competitive model from step two, we constructed an additional three models by removing covariates predicting $\gamma$, and (4) using each competitive model from step three, we constructed an additional 16 models by removing covariates predicting $\epsilon$ to identify competitive models combining all components of $\psi$, $\gamma$, $\epsilon$ and $p$. We only constructed three models in step three due to model convergence issues when either forest or forest edge variables were used to predict colonization. We assessed competitive models at each model-building step and in the final candidate set to ensure none contained non-informative parameters (Arnold, 2010).

The global model describing initial occupancy was

$$\text{logit}(\psi_i) = \hat{\beta}_0 + \hat{\beta}_1 \times \text{NORTH}_i + \hat{\beta}_2 \times \text{FOREST}_i + \hat{\beta}_3 \times \text{EDGE}_i + \hat{\beta}_4 \times \text{ACOV}_i,$$

where $\psi_i$ is the probability of site $i$ being occupied in the 1980 Atlas and $\hat{x}$ are estimated parameters associated with each covariate. ACOV is a spatial autocovariate describing the naïve occupancy (i.e. detected = occupied) of sites in the neighbourhood of each focal site. We added this autocovariate to account for spatial autocorrelation in occupancy (Augustin et al., 1996; Klute et al., 2002; Moore & Swihart, 2005; Betts et al., 2008), and we determined the value of ACOV for each site by examining the distance of significant spatial autocorrelation (15 km, or a 3-site × 3-site window) in residuals of initial occupancy from the global model. We calculated residuals following Moore & Swihart (2005) and assessed significance following Legendre & Legendre (1998) within a distribution of 1000 permutations of expected Moran’s $I$. We used the 15-km distance to generating an inverse distance-weighted average (excluding the focal site) of naïve 1980 Atlas occupancy in the neighbourhood of each site.

The most complex model describing colonization ($\gamma$) was

$$\text{logit}(\gamma_i) = \hat{\beta}_0 + \hat{\beta}_1 \times \text{NORTH}_i + \hat{\beta}_2 \times \text{NEIGHBOURS}_i,$$

while the most complex model describing extinction ($\epsilon$) was

$$\text{logit}(\epsilon_i) = \hat{\beta}_0 + \hat{\beta}_1 \times \text{NORTH}_i + \hat{\beta}_2 \times \text{FOREST}_i + \hat{\beta}_3 \times \text{EDGE}_i + \hat{\beta}_4 \times \text{NEIGHBOURS}_i,$$

where $\gamma$ or $\epsilon_i$ is the probability of site $i$ being colonized or becoming extinct, respectively, between the 1980 and 2000 Atlases and $\hat{\beta}$ are parameter estimates associated with each covariate. Similar to ACOV, NEIGHBOURS is a spatial autocovariate we included to account for the influences of neighbourhood occupancy (in the 1980 Atlas) on the probabilities of colonization and extinction. We determined the radius of this neighbourhood by first examining the Mantel spatial cross-correlation (Oden & Sokal, 1986; Koenig & Knops, 1998) of naïve occupancy in the 1980 and 2000 Atlases. The distance of significant correlation (following Legendre & Legendre, 1998) was 50 km, and we used this value to generate an inverse distance-weighted average of naïve 1980 Atlas occupancy in the neighbourhood of each site. We expected occupancy dynamics to be related to this index of past neighbourhood occupancy, forest cover and forest edge (Zuckerberg & Porter, 2010), and latitudinal range shifts expected under climate change (Zuckerberg et al., 2009).

The most complex model describing the detection process ($p$) was
logit(p_{ijt}) = \hat{\delta}_0 + \hat{\delta}_1 \times \text{forest}_{ijt} + \hat{\delta}_2 \times \text{effort}_{ijt} + \hat{\delta}_3 \times \text{detect1980}_{ijt} + \hat{\delta}_4 \times \text{detect3} \times 3_{ijt} + \hat{\delta}_5 \times \text{yr1980}_{ijt} + \hat{\delta}_6 \times \text{yr1981}_{ijt} + \hat{\delta}_7 \times \text{yr1982}_{ijt} + \hat{\delta}_8 \times \text{yr1983}_{ijt} + \hat{\delta}_9 \times \text{yr1985}_{ijt} + \hat{\delta}_{10} \times \text{yr2000}_{ijt} + \hat{\delta}_{11} \times \text{yr2001}_{ijt} + \hat{\delta}_{12} \times \text{yr2002}_{ijt} + \hat{\delta}_{13} \times \text{yr2003}_{ijt} + \hat{\delta}_{14} \times \text{yr2004}_{ijt} + \hat{\delta}_{15} \times \text{yr2005}_{ijt},

where \( p_{ijt} \) is the probability of detecting the Canada Warbler in site \( i \), block \( j \) and atlas \( t \), and \( \hat{\delta} \) are parameter estimates associated with each block-level covariate. \( \hat{\delta}_5 \) through \( \hat{\delta}_{15} \) are the estimated sampling year effects referenced by 1984 (chosen as a reference as it was the year most represented in our sample).

Model comparisons

To meet our second objective, we compared parameter estimates and estimates of occupancy and occupancy processes (absence, colonization, persistence and extinction) between occupancy models that accounted for imperfect detection (\( p < 1 \) models) and those that assumed species were detected perfectly (\( p = 1 \) models). To generate comparable information from \( p = 1 \) models, we fixed the assumption that species were detected perfectly by replacing all detection values with 1 if the Canada Warbler was detected at least once in site \( i \) and atlas \( t \) (J.A. Royle, pers. comm.). We created a set of \( p = 1 \) counterpart models by omitting block-level detection (\( p \)) covariates and by following the same hierarchical model-building process as we employed for \( p < 1 \) models.

We compared estimates of the latent occupancy state (\( \hat{z} \)) for \( p < 1 \) and \( p = 1 \) models at each site \( i \) in atlas \( t \) generated using empirical Bayes methods (Fiske et al., 2012) described in MacKenzie et al. (2006) and Royle & Dorazio (2008). The probability of \( \hat{z} \) can be represented as

\[
\Pr(\hat{z}_{it} = 1 | \gamma_{ijt}, \hat{\psi}; \hat{z}_i, \hat{\bar{p}}_{ijt})
\]

where \( \hat{z} \) is conditional upon \( y \) observations in across \( j \) blocks in site \( i \) in atlas \( t \) and model-predicted probabilities of \( \psi \) in site \( i \) in the 1980 Atlas, \( \gamma \) and \( \epsilon \) in site \( i \) between the 1980 and 2000 Atlases, and \( p \) across \( j \) blocks in site \( i \) in atlas \( t \). For example, \( \hat{\psi}_i \) (the model-predicted probability of 1980 Atlas occupancy) of a site in which the Canada Warbler was detected at least once may be < 1 due to effects of site-level covariates and parameter estimates associated with occupancy. However, given detection of the Canada Warbler in at least one block within this site, \( \hat{z} \) would be 1. Using competitive \( p < 1 \) and \( p = 1 \) models, we derived the mode value of the posterior distribution of \( \hat{z} \) (\( \hat{z}_{\text{mode}} \)) in the 1980 and 2000 Atlases. We generated model-weighted averages of \( \hat{z} \) in the event multiple models were competitive.

Consistent with estimates of latent occupancy states, we estimated the occupancy process of Canada Warblers between the 1980 and 2000 Atlases in \( p < 1 \) and \( p = 1 \) models using values of \( \hat{z}_{\text{mode}} \). If \( \hat{z}_{\text{mode}} \) was 0 in the 1980 Atlas and 1 in the 2000 Atlas, we categorized the site-level occupancy process as colonization. Similarly, if \( \hat{z}_{\text{mode}} \) was 1 in the 1980 Atlas and 0 in the 2000 Atlas, we categorized the occupancy process as extinction. The occupancy process of sites in which \( \hat{z}_{\text{mode}} \) remained 0 was categorized as absence, while the occupancy process of sites in which \( \hat{z}_{\text{mode}} \) remained 1 was categorized as persistence. We calculated \( \hat{z}_{\text{mode}} \) from multiple competitive models using a model-weighted average and rounded values to the nearest integer. We tested differences in proportions of occupancy and occupancy processes between \( p < 1 \) and \( p = 1 \) models using two-sample chi-square tests for equality of proportions. While \( \hat{z}_{\text{mode}} \) is estimable for sites with incomplete sampling histories, we excluded estimates of \( \hat{z}_{\text{mode}} \) from these tests for site-atlas combinations removed during subsampling.

Testing model fit and accuracy

We used 1000 bootstrap randomizations (resampling detection data and refitting models each time) to generate a distribution of chi-square values (Fiske et al., 2012). We used the rank of the observed chi-square value within this distribution as a measure of goodness-of-fit of global \( p < 1 \) and \( p = 1 \) models. To assess the overall ability for models to accurately classify occupancy, we first calculated model-predicted probabilities of detecting the Canada Warbler at least once in a site (\( D \)) given its probability of site-level occupancy and block-level detection

\[
\hat{D}_{it} = \hat{\psi}_{it} \hat{p}_{ijt},
\]

where \( \hat{\psi}_{it} \) is the predicted probability of occupancy in site \( i \) in atlas \( t \), \( \hat{p}_{ijt} \) is the probability of detecting a species at least once in site \( i \) in atlas \( t \) as calculated by

\[
1 - \prod_{j=1}^{J} (1 - \hat{p}_{ijt}),
\]

\( \hat{p}_{ijt} \) is the predicted probability of detection in site \( i \) and observation \( j \) in atlas \( t \), and \( J \) is the number of blocks within a site (Moore & Swihart, 2005; Rota et al., 2011). We followed MacKenzie et al. (2003) in calculating values of \( \psi \) in site \( i \) in the 2000 Atlas as

\[
\hat{\psi}_{i2000} = \hat{\psi}_{i1980} (1 - \hat{e}_i) + \gamma_i \left( 1 - \hat{\psi}_{i1980} \right).
\]

We then compared these probabilities to the naïve occupancy of a site using the receiver operating characteristic curve (sensitivity plotted against 1 minus specificity at multiple cut-off thresholds) to calculate the area under the curve (AUC; Fielding & Bell, 1997) as an overall measure of classification accuracy.

We conducted all statistical analyses in \( R \) (R Development Core Team, 2012). We used the SPDEP package (Bivand,
to examine autocorrelation in occupancy residuals and to construct spatial autocovariates, the ncf package (Bjørnstad, 2012) to generate Mantel cross-correlograms, the ROCR package (Sing et al., 2009) to calculate AUC of models and the unmarked package (Fiske et al., 2012) to model occupancy, generate estimates of latent occupancy and calculate goodness-of-fit statistics. We derived forest covariates using ARCGIS (ESRI 2009). No site-level covariates were correlated at |r| > 0.6. We scaled all continuous covariates to a mean of 0 and a standard deviation of 1 prior to modelling.

RESULTS

Subsamples

We identified 331 and 322 sites in the 1980 and 2000 Atlases, respectively, that contained at least two blocks (mean = 2.7, SE = 0.03) from a total 1790 blocks meeting our subsampling criteria. The total number of unique sites was 504. Of these, 149 sites were sampled in both atlases. Characteristics of these sites and blocks indicated they were representative of all sites and blocks in New York having at least 50% forest cover. The notable exception is lower effort in our subsample, an expected difference given sites sampled in multiple years (and mostly excluded from our analysis) were also those with more observer effort (r > 0.36, P < 0.0001 in both atlases; Table 1).

Detection, occupancy, colonization and extinction

Only one model was competitive at each step of model-building (see Table S1 in Supporting Information for a complete list of models). A similar model was competitive among p < 1 and p = 1 model sets (Table 2), the only difference (aside from the exclusion of p covariates in the p = 1 model) being the exclusion of the NEIGHBOURS covariate from the p < 1 colonization model. While other models had ΔAIC values < 2, the additional parameters in these models were uninformative (Arnold, 2010). Mean detection probability for the p < 1 model was 54.0% (SE = 0.5%). Higher probabilities of detection were associated with greater effort, previous detection in the 1980 Atlas (detect80), detection in neighbouring blocks within the same atlas and prior to the sampling year (neigh3), higher forest cover (forest) and the year of sampling. Sampling years in the 1980 Atlas generally showed higher detection probabilities that those in the 2000 Atlas.

The probability of site occupancy was associated with elevation (ELEV), forest cover (FOREST) and neighbourhood occupancy of the site (ACOV). Sites more likely to be occupied were located at intermediate elevations, had higher forest cover and, in accounting for spatial autocorrelation, had a higher index of occupancy in the surrounding eight sites. These relationships were overall consistent between the p < 1 and p = 1 model. Colonization between atlases was best explained by a constant (intercept) in the p < 1 model, indicating neither the inclusion of site northing (NORTH) nor neighbourhood occupancy (NEIGHBOURS) improved model predictions of colonization between the 1980 and 2000 Atlases. Past neighbourhood occupancy was weakly associated with a higher colonization probability in the p = 1 model. Sites with lower extinction probabilities had lower forest–non-forest edge length (EDGE) and were more isolated from other occupied sites.


<table>
<thead>
<tr>
<th>Level</th>
<th>Variable</th>
<th>Sites used in modelling</th>
<th>All sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site</td>
<td>NORTH (100 km)</td>
<td>47.8 ± 1.1</td>
<td>47.7 ± 1.1</td>
</tr>
<tr>
<td></td>
<td>ELEV (m)</td>
<td>416 ± 1.16</td>
<td>410 ± 1.66</td>
</tr>
<tr>
<td></td>
<td>FOREST (%)</td>
<td>76.6 ± 12.6</td>
<td>72.7 ± 13.7</td>
</tr>
<tr>
<td></td>
<td>EDGE (km)</td>
<td>132 ± 47</td>
<td>138 ± 47</td>
</tr>
<tr>
<td></td>
<td>ACOV (%)</td>
<td>75.9 ± 24.1</td>
<td>73.2 ± 25.8</td>
</tr>
<tr>
<td></td>
<td>NEIGHBOURS (%)</td>
<td>67.0 ± 11.5</td>
<td>65.0 ± 12.4</td>
</tr>
<tr>
<td>Block</td>
<td>forest (%)</td>
<td>77.5 ± 13.8</td>
<td>74.3 ± 14.1</td>
</tr>
<tr>
<td></td>
<td>effort (h)*</td>
<td>12.5 ± 2.1</td>
<td>17.6 ± 2.4</td>
</tr>
<tr>
<td></td>
<td>detect80 (%)</td>
<td>22.6 ± 4.18</td>
<td>23.9 ± 4.26</td>
</tr>
<tr>
<td></td>
<td>neigh3 (%)</td>
<td>14.2 ± 1.7</td>
<td>14.0 ± 1.75</td>
</tr>
</tbody>
</table>

Site-level (10 × 10 km) covariates predicting multiseason occupancy were the Universal Transverse Mercator northing (NORTH), elevation (ELEV), percentage forest cover (FOREST), the length of forest–non-forest edge (EDGE) and a 15 km radius (ACOV) and 50 km radius (NEIGHBOURS) spatial autocovariate of detected occupancy in the 1980–1985 Atlas. Block-level (5 × 5 km) or block-atlas level detection covariates were forest cover (forest), the total observer hours (effort), detection in the first atlas (detect80), within-atlas detections in the block neighbourhood (neigh3) and the year of sampling (6 years in each atlas). Samples were 504 sites and 1790 blocks in the sample used for modelling, and 811 sites and 6603 blocks (includes both atlases) for the entire state. All sites and blocks in the statewide set had at least 50% forest cover.

*Effort statistics were calculated from log-transformed values, which were subsequently back-transformed to their original scale.

Model fit and classification accuracy

Parametric bootstrap statistics from simulated data (p < 1 model: P = 0.23; p = 1 model: P = 0.50) indicated models provided adequate goodness-of-fit to detections. AUC for both p < 1 models (0.82) and p = 1 models (0.81) showed good performance in predicting sites with detections from those without detections in the 1980 Atlas. The model accounting for p < 1 performed 1.5% better in predicting 1980 detections than that assuming p = 1. AUC in the 2000 Atlas (p < 1 models = 0.71, p = 1 models = 0.61), indicating fair to poor performance, with the p < 1 model performing 16% better than the p = 1 model.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model accounting for (p &lt; 1)</th>
<th>Model assuming (p = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>UCI</td>
</tr>
<tr>
<td>(\psi)</td>
<td>Intercept</td>
<td>3.140</td>
</tr>
<tr>
<td></td>
<td>ELEV</td>
<td>0.196</td>
</tr>
<tr>
<td></td>
<td>ELEV(^2)</td>
<td>-0.676</td>
</tr>
<tr>
<td></td>
<td>FOREST</td>
<td>1.302</td>
</tr>
<tr>
<td></td>
<td>ACOV</td>
<td>1.097</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>Intercept</td>
<td>-0.310</td>
</tr>
<tr>
<td></td>
<td>NEIGHBOURS</td>
<td>-</td>
</tr>
<tr>
<td>(\varepsilon)</td>
<td>Intercept</td>
<td>-1.710</td>
</tr>
<tr>
<td></td>
<td>NEIGHBOURS</td>
<td>-0.947</td>
</tr>
<tr>
<td>(p)</td>
<td>Intercept</td>
<td>0.583</td>
</tr>
<tr>
<td></td>
<td>forest</td>
<td>0.404</td>
</tr>
<tr>
<td></td>
<td>effort</td>
<td>0.377</td>
</tr>
<tr>
<td></td>
<td>detect80</td>
<td>0.980</td>
</tr>
<tr>
<td></td>
<td>neigh3</td>
<td>0.217</td>
</tr>
<tr>
<td></td>
<td>Year 1980</td>
<td>0.646</td>
</tr>
<tr>
<td></td>
<td>Year 1981</td>
<td>0.374</td>
</tr>
<tr>
<td></td>
<td>Year 1982</td>
<td>0.652</td>
</tr>
<tr>
<td></td>
<td>Year 1983</td>
<td>0.360</td>
</tr>
<tr>
<td></td>
<td>Year 1985</td>
<td>0.759</td>
</tr>
<tr>
<td></td>
<td>Year 2000</td>
<td>0.552</td>
</tr>
<tr>
<td></td>
<td>Year 2001</td>
<td>1.468</td>
</tr>
<tr>
<td></td>
<td>Year 2002</td>
<td>1.913</td>
</tr>
<tr>
<td></td>
<td>Year 2003</td>
<td>0.867</td>
</tr>
<tr>
<td></td>
<td>Year 2004</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>Year 2005</td>
<td>1.021</td>
</tr>
</tbody>
</table>

Site-level covariates predicted initial occupancy \((\psi)\), colonization \((\gamma)\) and extinction \((\varepsilon)\), while block-level covariates (in the \(p < 1\) model) predicted detection \((p)\). All continuous covariates (described in Table 1) were scaled to a mean of 0 and standard deviation of 1. The effects of year (relative to 1984) are numbers preceded by ‘year’. Parameter estimates and upper and lower 95% confidence intervals are presented. Confidence intervals that do not include zero are in bold.

*The upper 95% confidence interval of EDGE was 0.0001.

Estimated statewide occupancy was higher for the \(p < 1\) model in both the 1980 Atlas (81% vs. 71%, \(\chi^2 = 7.92, P = 0.01,\) Fig. 2) and the 2000 Atlas (71% vs. 60%; \(\chi^2 = 8.39, P = 0.01)\) than for \(p = 1\) model. The estimated proportion of sites that remained absent in both atlases was marginally lower for \(p < 1\) and \(p = 1\) models (8% vs. 15%, \(\chi^2 = 3.24, P = 0.08)\), but the proportion of colonized sites was identical (14%). A higher proportion of sites were estimated to have had persistent occupancy in the \(p < 1\) vs. \(p = 1\) models (61% vs. 48%; \(\chi^2 = 4.88, P = 0.03)\), and a non-significantly lower proportion of sites was estimated to have experienced local extinction (17% vs. 23%; \(\chi^2 = 1.35, P = 0.25)\).

**DISCUSSION**

That an observer can fail to detect a species even when present is often overlooked by ecologists examining broad-scale...
distributional data and modelling species distributions, potentially leading to unreliable estimates and ecological conclusions (Tyre et al., 2003; MacKenzie et al., 2006; Kéry, 2010; Rota et al., 2011). Indeed, it was not long ago that a synthesis dedicated to the subject of species distributions (Scott et al., 2002) included only one chapter (Stauffer et al., 2002) in which the detectability of a species was modelled explicitly. However, despite the recent advances in modelling species distributions (Elith et al., 2006), the majority of these modelling approaches generally confound occurrence and detection (Kéry et al., 2010).

By implementing a space-for-time substitution in sampling, we accounted for variation in detection probabilities using occupancy models estimating the distribution and associated environmental constraints of a model species. By ignoring variation in detection probabilities, occupancy was underestimated by 14% in the 1980 Atlas and 19% in the 2000 Atlas. These differences contributed to a 28% higher estimated persistence and a 26% lower estimated extinction between atlases. Assuming similar patterns in available sites (≥ 50% forest cover) not included in our sample (total n = 811), these results suggest atlasers may have failed to detect the Canada Warbler in approximately 78 and 91 sites where it was likely present in the 1980 and 2000 Atlases, respectively. The effects of site-level covariates on occupancy and occupancy dynamics were similar in p < 1 vs. p = 1 models, with only slightly larger effects of elevation and forest on occupancy and forest–non-forest and neighbourhood occupancy on extinction. The influence of forest–non-forest as an index of effects presumably favourable to the Canada Warbler showed expected, though only marginal effects on extinction. Methods of quantifying forest age and structure (e.g. LiDAR; Vierling et al., 2008; Muller et al., 2010; Lesak et al., 2011) may facilitate better predictions of occupancy and occupancy dynamics in this species. While model improvements are no doubt possible, the more robust estimates of the p < 1 model may be particularly important in estimating past changes and predicting future changes in distributions of this and other species of conservation concern.

We effectively modelled patterns of detectability and occupancy using a parsimonious suite of environmental and sampling covariates that allowed us to explore the factors driving patterns and changes in occupancy. Predictably, we found that Canada Warblers were more likely to be detected in blocks that received greater amounts of sampling effort. Although not typically addressed, a priori knowledge of past detection in blocks or block neighbourhoods was also associated with a higher probability of detection. The importance of each covariate in modelling the detection process associated with atalasing, in particular observer effort, underscores the relevance of a modelling approach that does not assume species are detected perfectly.

Many conservation programs now appreciate the importance of accounting for imperfect detection when censusing and monitoring plant and animal populations. Kéry et al. (2010) provided an example of the potential pitfalls of not accounting for imperfect detection in modelling species distributions using volunteer-collected checklist data of Blue Hawks (Aeshna cyanea) [Müller 1764]; Odonata, Aeshnidae) along an elevational gradient in Switzerland. By comparing occupancy modelling with a conventional method that assumed species were detected perfectly, they found that ignoring the possibility that the species was not detected perfectly caused a noticeable underestimation of distribution and biased coefficients of the covariates shaping that distribution. Other recently developed approaches hold promise for generating more robust occupancy estimates. Notable are single-visit models of occupancy accounting for imperfect detection (Lele et al., 2012) and occupancy modelling approaches developed for analysing historical data sets with or without repeated visits (Tingley & Beissinger, 2009; Tingley et al., 2009).

While comparisons of our model accounting for imperfect detection with the model assuming Canada Warblers were detected perfectly illustrate the importance of accounting for imperfect detection, there are several important caveats to applying our approach to other atlases. First, the use of nested atlas blocks (e.g. 25 km²) as repeat surveys within a larger sampling site (e.g. 100 km²) decreases overall sample size. This is not a problem for some species, but for species with rare or restricted ranges, this reduction in sample size might decrease the predictive power of any analytical procedure and would be particularly problematic for atlases with very coarse sampling resolutions.

Our use of a single habitat-related sample-removal criterion (< 50% forest cover) was intended to ensure closure among within-site subsamples (blocks) by (1) emulating a spatial analogue to temporally clustered subsamples (Rota et al., 2009) and (2) reducing the potential variation in detection probabilities due to variation in the availability for detection (i.e. a ‘true presence’). The addition of temporally replicated sampling among a portion of space-for-time replicates has been recommended to account for the biases possibly when including space-for-time replicates in which a species is not available for detection (Kendall & White, 2009). However, as noted and shown in simulations by Guillera-Arroita (2011), this concern may not always be warranted, such as when the availability for detection varies randomly (MacKenzie et al., 2006). Additionally, these biases may be greatly reduced in the analyses of atlas data by removal of sites that are suspected a priori to be unavailable for detection. Researchers modelling the occupancy of understudied species should consider comparing estimates under multiple definitions of habitat availability (McClearn et al., 1998).

Careful consideration of potential closure violations is also important for atlases completed over multiple years. In addition to removing less forested and potentially non-available blocks, we included only those blocks with at least 90% of all species detected in one primary sampling year. While we allowed primary sampling years of blocks to vary within sites and, in turn, the sampling years of sites to vary within each atlas, the assumption of closure could be relaxed due to (1) our inclusion of sampling year as a covariate and (2) both

Diversity and Distributions, 19, 804–814, © 2013 John Wiley & Sons Ltd
small dispersal distances and high breeding site fidelity of this species (Hallworth et al., 2008; Reitsma et al., 2010). While our subsampling criteria left an adequate sample of sites, this reduction may, unfortunately, preclude occupancy modelling in other atlases.

Biological atlases, among other broad-scale surveys, are a form of citizen science that has become an increasingly important tool for ecological research (Devictor et al., 2010; Dickinson et al., 2010; Robertson et al., 2010). It may be assumed that the relatively coarse resolution of these types of surveys makes them less susceptible to issues of detectability because the sampling extent and resolution are relatively large and observations are often collected through repeated visits. In general, however, the assumption that larger sampling units and greater survey efforts lead to detectability close to 1 is relatively unexplored. In this study, we demonstrate that a space-for-time occupancy approach can quantify variation in detectability over relatively broad scales and how this potential bias can alter estimates of occupancy and occupancy dynamics. Given the global use of atlases for collecting biological observations across multiple scales of interest, more research is warranted on how to effectively account for imperfect detection in the analysis of atlas data.

ACKNOWLEDGEMENTS

We would like to thank the thousands of volunteers who participated in both New York State Breeding Bird Atlases. We also thank Kimberley Corwin and Kevin McGowan for supplying atlas databases and other information. This manuscript benefited from discussions with Michelle Peach-Lang, Thomas Albright, Darryl Mackenzie and James Nichols. We thank Richard Chandler, Marc Kéry, J. Andrew Royle and Ian Fiske for their assistance in using the unmarked package for program R. We thank Marc Kéry, Matthew Betts and three anonymous reviewers for their helpful comments on an earlier draft of this manuscript. This study received financial support from NASA.

REFERENCES


**SUPPORTING INFORMATION**

Additional Supporting Information may be found in the online version of this article:


**BIOSKETCH**

**Giancarlo Sadoti** is a doctoral student at the University of Nevada, Reno, examining the responses of birds to extreme weather events. **Benjamin Zuckerberg** is an assistant professor in the Department of Forest and Wildlife Ecology at the University of Wisconsin-Madison interested in the impacts of modern climate on wildlife populations. **Marta A. Jarzyna** is a doctoral student in the Department of Fisheries and Wildlife at Michigan State University interested in explaining macroecological patterns of species diversity, abundance and distributions. **William F. Porter** is the Boone and Crockett Chair of Wildlife Conservation in the Department of Fisheries and Wildlife at Michigan State University and leads the Quantitative Wildlife Laboratory. His research is directed toward understanding the relationships among habitat, movement behavior, and population dynamics as these shape key issue in wildlife conservation.

Author contributions: G.S., B.Z. and W.F.P. conceived the idea for the study; G.S., M.A.J. and B.Z. designed the study; G.S. analysed the data; G.S., B.Z., M.A.J. and W.F.P. wrote the manuscript.

Editor: Mark Robertson