

# A Flexible and Efficient Modeling Strategy for Trend Analysis of North American Bird Count Data

T. D. Meehan, N. L. Michel, H. Rue

June 20, 2018

## 1 Motivation

Volunteers with the Audubon Christmas Bird Count (CBC, Soykan et al. 2016) have been counting wintering birds across North America every year for the last 118 years. Population trends derived from CBC data, along with those derived from other large-scaled monitoring programs like the North American Breeding Bird Survey (BBS, Sauer and Link 2011), are important pieces of information for understanding the conservation needs of North American bird species (Partners in Flight 2016).

The current standard approach for generating trends from CBC data (Link et al. 2006, Soykan et al. 2016) is based on methods originally developed for BBS data (Link and Sauer 2002, Sauer and Link 2011). The general approach is to *(i)* assign counts in Canada and the US to roughly 150 spatial strata, which are intersections of US states, Canadian provinces, and Bird Conservation Regions (BCR, North American Bird Conservation Initiative, Figure 1A). Then, treating each stratum as independent, *(ii)* use a non-linear function to remove the effect of observer effort on counts, and *(iii)* model the residual as a function of count circle identifier, stratum, and year. Next, *(iv)* model parameters are used to derive a relative abundance index per stratum and year, and *(v)* those indices are used to compute percent change per stratum across different time intervals.

The pros of the current CBC approach are as follows. *(i)* By pooling count circles per stratum, this approach deals with the issue of count locations haphazardly becoming active or inactive over the time series. *(ii)* By pooling per stratum, a large enough sample of counts is attained to generate a reasonably robust count-effort correction function. *(iii)* The approach produces a relative abundance index per year and stratum, which can be summed across larger strata, such as states, provinces, or BCRs, and used to calculate temporal trends at larger spatial scales.

The cons of the current approach are as follows. *(i)* It is computationally intensive, especially for wide-ranging species such as the American Robin, as it uses Markov chain Monte Carlo (MCMC) to estimate model parameters for relative abundance, and then processes large MCMC chains to scale relative abundance to larger aggregate units and generate trend estimates. *(ii)* While trends can be scaled up to larger aerial units, they cannot be scaled down to smaller ones. The analytical strata is the finest level of resolution, which limits the extent to which trends can be attributed to specific factors occurring at finer spatial scales (Thogmartin et al. 2004, Bled et al. 2013). *(iii)* It does not take full account or full advantage of spatial relationships in counts. Modeling this structure would facilitate borrowing information across spatial boundaries, allowing more robust trend estimates in places where there is a shortage of data (Thogmartin et al. 2004, Bled et al. 2013). Indeed, borrowing information would allow trends to be estimated at spatial scales that are finer than the spatial strata currently used.

Work by Thogmartin et al. (2004), Bled et al. (2013), and Smith et al. (2015) offered spatially-explicit variations of the standard trend analysis approach. These works were focused on analysis of BBS data, but their approaches are easily related to analysis of CBC data. Their approaches contrasted with the standard approach, described above, in the following ways. *(i)* Instead of using the standard strata described above, Thogmartin et al. (2004) assigned count sites to irregular polygons, created by tessellation of BBS route locations. Bled et al. (2013) assigned routes to cells on a regular grid, with one-degree latitude and longitude spacing (similar to Figure 1B). *(ii)* All three studies utilized spatially-structured (intrinsic conditional autoregressive model, ICAR, Besag et al. 1991) random intercepts for relative abundance per polygon or grid cell. *(iii)* Thogmartin et al. (2006) utilized a fixed effect of time per BBS route, but that effect did not incorporate spatial



Figure 1: (A) Spatial strata from the standard CBC approach (Soykan et al. 2016) and (B) grid cells used for the current analysis, which are similar to those used by Bled et al. (2013).

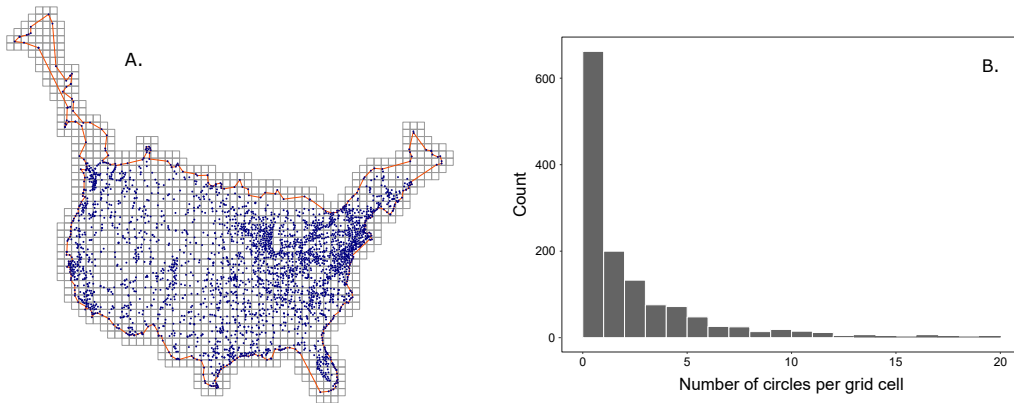


Figure 2: (A) CBC locations (blue points), along with the non-convex hull (red line), and grid cells (gray) used for this analysis of American Robin data. (B) The number of CBC locations with American Robin detections per grid cell.

structure. (iv) Bled et al. (2013) and Smith et al. (2015) did not include a fixed effect of time. Instead, they estimated relative abundances per year, and then trends were generated as a derived parameter during MCMC simulations.

Here, we present a different approach for calculating temporal trends in relative abundance, one that takes advantage of the considerable spatial structure in CBC data. This approach borrows components from previous ones, incorporates new components that prioritize robust trend estimation at finer spatial scales, and employs a relatively simple and computationally efficient process. Similar to Bled et al. (2013), we assigned CBC count sites to cells on a uniform grid that covered North America. Like Thogmartin et al. (2004), temporal trends were explicit components of the model. In contrast to previous work, effort and year effects were modeled as random slopes with spatial structure, following a spatially varying coefficient (SVC) approach (Congdon 2014). Finally, unlike prior studies using MCMC, we used integrated nested Laplace approximation (INLA) to estimate Bayesian posteriors for model parameters, which led to a considerable decrease in computing time.

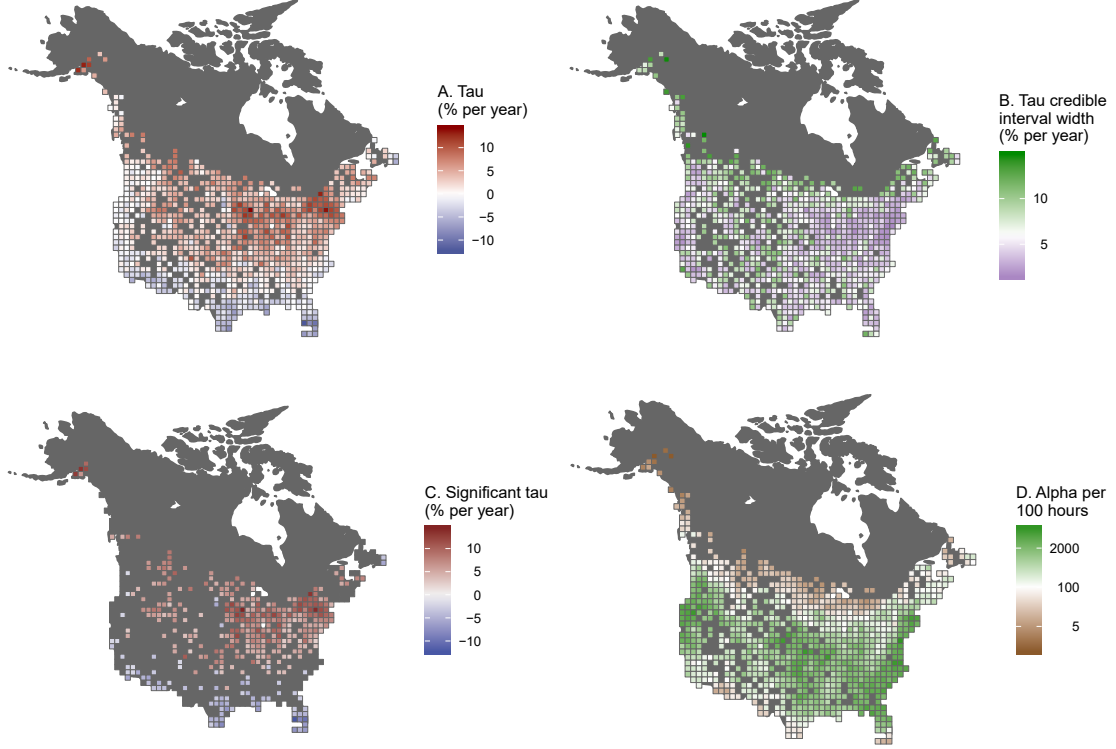


Figure 3: American Robin trends (A), along with the width of the 95% credible interval (B) per grid cell. Trends are  $\tau_i$  values that are exponentiated and converted to a percent. (C) The cells from (A) that are significantly different from zero. (D) The abundance index per grid cell,  $\alpha_i$ , is exponentiated and scaled to 100 party hours.

## 2 Model

We modeled Christmas Bird Counts,  $y_{i,k,t}$  for grid cell  $i$  encompassing count circle  $k$  during year  $t$ , as a random variable from a negative binomial distribution. Expected values for counts per grid cell,  $\mu_{i,t}$ , were assumed to be a function of spatially-structured grid-cell, count-effort, and year effects. The linear predictor for  $\mu_{i,t}$  took the form

$$\log(\mu_{i,t}) = \alpha_i + \epsilon_i \log(E_{i,k,t}) + \tau_i T_{i,k,t} + \kappa_k.$$

$\alpha_i$  was a cell-specific relative abundance term that comprised the sum of two components,  $\alpha_i = v_0 + v_i$ . The first component,  $v_0$ , was a global intercept averaged over all grid cells. The second component,  $v_i$ , was a spatially-structured random effect per cell, modeling deviations from the global intercept with ICAR structure, where values came from a normal distribution, with a conditional mean related to the average of adjacent cells, and with conditional variance proportional to the variance across adjacent cells and inversely proportional to the number of adjacent cells. Spatial structure was incorporated into  $\alpha_i$  to allow for information about relative abundance to be shared across neighboring cells.

$\epsilon_i$  was a cell-specific slope term for the effect of effort on counts.  $\epsilon_i$  was also a composite term, comprising the sum of two components,  $\epsilon_i = \psi_0 + \psi_i$ . The first component,  $\psi_0$ , was a global effect of effort derived over all grid cells. The second component,  $\psi_i$ , an SVC term, was a spatially-structured random slope per cell, modeling local variation in the effort effect with ICAR structure, where values came from a normal distribution, with a conditional mean related to the average of adjacent cells, and with conditional variance proportional to the variance across adjacent cells and inversely proportional to the number of adjacent cells. Spatial structure was incorporated into  $\epsilon_i$  to allow for information about the effort effect to be shared across neighboring cells. Effort was represented by  $E_{i,j,k}$ , the number of party hours expended during a count, where a party hour was the count effort of one party of unspecified size for one hour. Pairing log-transformed counts with

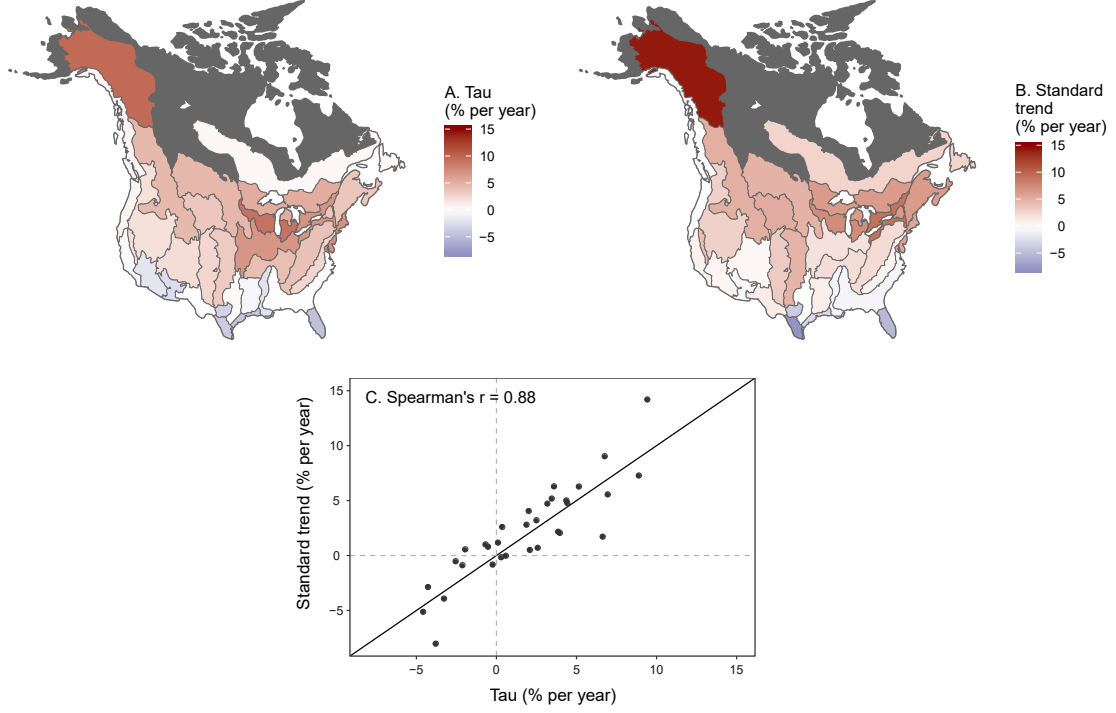


Figure 4: Maps of American Robin posterior median trends aggregated to the BCR level, produced using the SVC approach (A) and the standard approach (B). (C) Bivariate plot of the trends mapped in (A) and (B), with 1:1 line.

log-transformed effort in the linear predictor yielded a power function for effort correction, a flexible form that accommodated a decreasing, linear, or increasing impact of effort on expected counts (Link et al. 2006, Soykan et al. 2016).

$\tau_i$  was a cell-specific slope for the effect of year on counts. Similar to  $\epsilon_i$ ,  $\tau_i$  was a composite term, comprising the sum of two components,  $\tau_i = \delta_0 + \delta_i$ . The first component,  $\delta_0$ , was a global effect of year derived over all grid cells. The second component,  $\delta_i$ , an SVC term, was a spatially-structured random slope per cell, modeling local variation in the year effect with ICAR structure, where values came from a normal distribution, with a conditional mean related to the average of adjacent cells, and with conditional variance proportional to the variance across adjacent cells and inversely proportional to the number of adjacent cells. Spatial structure was incorporated into  $\tau_i$  to allow for information about the year effect to be shared across neighboring cells. Year, represented by  $T$ , was transformed before analysis such that  $\max(T) = 0$ , and each preceding year took an increasingly-negative integer value. Given the scaling of effort and year variables,  $\exp(\alpha_i)$  could be interpreted as cell-specific expected counts given one party hour of effort during the final year in a time series.

The final term in the model,  $\kappa_k$ , was a random effect that accounted for variation in relative abundance among circles, possibly due to differences in habitat conditions or observer experience.

### 3 Data

Data used here to demonstrate the SVC modeling approach was from the American Robin, from Christmas Bird Counts conducted across North America from 1966 through 2017. Before modeling the data, extreme outliers ( $> 3$  SD from the mean, after log transformation) in counts and effort were removed. After filtering, there were 78,140 counts from 3195 count circles for modeling.

Locations of the 3195 count circles (Figure 2A) were mapped using the North American Albers Equal Area Conic projection (EPSG 102008) and assigned to 880 cells on a grid divided along 100 km increments in latitude and longitude (Figure 2A). Grid cells formed a continuous lattice within a

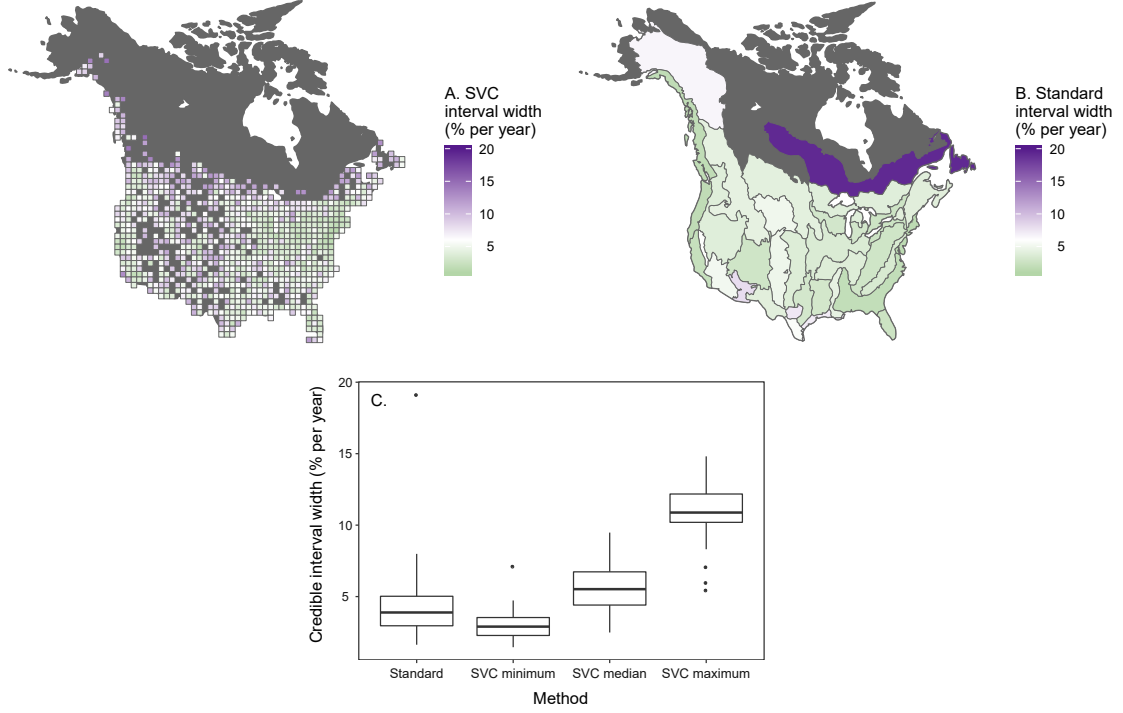


Figure 5: Maps of credible interval width for median trends, produced using the SVC approach (A) and the standard analysis (B). (C) Boxplots showing the credible interval widths across BCRs using the standard approach, along with minimum, median, and maximum credible interval widths across cells within BCRs using the SVC approach.

non-convex polygon created using circle locations. A continuous lattice was used to improve qualities of the neighborhood structure used in ICAR modeling. The number of count circles per grid cell varied from 0 to 20, but averaged 2.76 (Figure 2B). The number of neighbors for a given grid cell ranged from 1 to 8 and averaged 7.48.

## 4 Analysis

Hierarchical Bayesian model analysis was conducted using the R-INLA package for R. Priors for the global intercept and fixed effects were vague normal distributions with mean = 0 and precision = 0 and 0.001, respectively. Priors for exchangeable and ICAR random effects were penalized complexity priors with parameter values  $U_{pc} = 1$  and  $\alpha_{pc} = 0.01$  (Simpson et al. 2017). The model formula, written in R-INLA syntax, was specified as shown below.

```
formula <- count ~
  # alpha ~ epsilon_0 + epsilon_i
  1 + f(grid_id1, model="besag", graph=g, scale.model=TRUE, hyper=list(prec=
    list(prior="pc.prec", param=c(1, 0.01)))) +
  # epsilon ~ psi_0 + psi_i
  log_hrs + f(grid_id2, log_hrs, model="besag", graph=g, scale.model=TRUE,
    hyper=list(prec=list(prior="pc.prec", param=c(1, 0.01)))) +
  # tau ~ delta_0 + delta_i
  std_yr + f(grid_id3, std_yr, model="besag", graph=g, scale.model=TRUE, hyper=
    list(prec=list(prior="pc.prec", param=c(1, 0.01)))) +
  # kappa
  f(circle, model="iid", hyper=list(prec=list(prior="pc.prec", param=c(1, 0.01))))
```

The model was analyzed with a call to the `inla()` function and a summary of the result was

printed with a call to the `summary()` function. Note that parameter names in the summary have been modified to match those in the model description.

```
result <- inla(form1, family="nbinomial", data=modeling_data, control.compute=
  list(cpo=T, config=T), num.threads=3, verbose=T)
```

```
summary(result)
```

```
Time used: 16933.517 sec
```

Fixed effects:	mean	sd	0.025quant	0.5quant	0.975quant
upsilon_0	0.479	0.118	0.247	0.479	0.710
psi_0	0.798	0.027	0.745	0.798	0.852
delta_0	0.022	0.002	0.018	0.022	0.026

Model hyperparameters:	mean	sd	0.025quant	0.5quant	0.975quant
(1/overdispersion)	0.514	0.001	0.512	0.514	0.517
Precision, upsiilon_i ICAR	0.378	0.027	0.332	0.376	0.438
Precision, psi_i ICAR	14.329	2.335	10.242	14.160	19.402
Precision, delta_i ICAR	518.794	32.659	464.640	514.958	592.211
Precision, kappa_k exchangeable	0.909	0.035	0.849	0.906	0.986

Model analysis took approximately 5 hours. The analysis summary lists parameter estimates for global fixed effects as well as estimates for hyperparameters, such as the negative binomial dispersion parameter and the precision estimates associated with the random effects. The following results are noteworthy.

First, we see that the posterior median for the log-effort effect,  $\psi_0$ , was 0.798. This value represents a power law exponent for the relationship between effort and counts. A value less than 1 indicated that, on average, there was a diminishing return for additional count effort. Second, we get an estimate for  $\delta_0$  of 0.022, indicating that, on average, American Robins have increased by approximately 2.2% per year between 1966 and 2017, or 210% over the 52 year period.

The overdispersion parameter,  $1 / 0.514 = 1.946$ , was considerably larger than 1, highlighting overdispersion in Robbin counts. Precision estimates for the random effects showed that all were important for explaining variation in the data. The random effects were ranked  $v_i$ ,  $\kappa_k$ ,  $\psi_i$ , and  $\delta_i$ , in terms of the amount of variation explained.

Samples from the posterior distributions of latent effects were acquired using the `inla.sample.posterior()` function. A posterior distribution for  $\alpha_i$  was created by summing samples for  $v_0$  and  $v_i$ . A posterior distribution for  $\epsilon_i$  was created by summing samples for  $\psi_0$  and  $\psi_i$ . A posterior distribution for  $\tau_i$  was created by summing samples for  $\delta_0$  and  $\delta_i$ .

Characteristics of posteriors for  $\alpha_i$  and  $\tau_i$  are mapped in Figure 3. Figure 3 illustrates that, given a standard count effort of 100 party hours, American Robin counts are expected to vary dramatically across their range, from  $\sim 5$  to  $\sim 2000$ . The species is most abundant in regions with moderate climate and decreases in relatively warm and cold regions. Regarding trends in relative abundance, American Robins have generally decreased in the warmer parts of their range and increased in the colder parts of their range. These trends could signal a shift in the species winter range, possibly due to increased bird feeding and gradual increases in climate over the 52 year period. Figure 3B depicts our confidence in trend estimates across the species winter range, where it is higher (narrow credible intervals) in grid cells with a higher density of count circles and lower (wide credible intervals) in places where information is sparse. The credible intervals shown in Figure 3B can be used to determine where trends are significantly (probability  $\geq 0.95$ ) greater than or less than 0% change per year.

A common practice with the standard approach is to aggregate trends at the stratum level up to larger scales, such as the BCR. We were curious how aggregate trends using the SVC method compared to those from the standard method. We calculated the median of posterior medians across grid cells within each BCR, to compare aggregate median trends across the two methods. Figure 4 shows the posterior median of posterior medians across cells per BCR (A) along with the posterior median trend for each BCR from the standard analysis. Side-by-side comparison of these maps shows that aggregate trends are similar, regardless of method. When values from different methods are directly compared to one another, they give a correlation coefficient of 0.88.

We were also interested in learning how the precision of trend estimates compared across the two approaches. Figure 5 compares the credible interval widths per grid cell (A) with those for aggregate BCR estimates from the standard approach (B). When compared to the standard approach (Figure 5C, Standard), some grid cells within a BCR, ones in information rich regions, have trend estimates with relatively high precision (Figure 5, SVC minimum). Other grid cells, ones in information poor regions, have trend estimates with relatively low precision. On average, however, precisions of estimates are similar, regardless of method. These results indicate that the SVC approach is capable of generating trends at fine spatial scales, while also producing aggregate trends that are similar in magnitude and precision to those generated using the standard method.

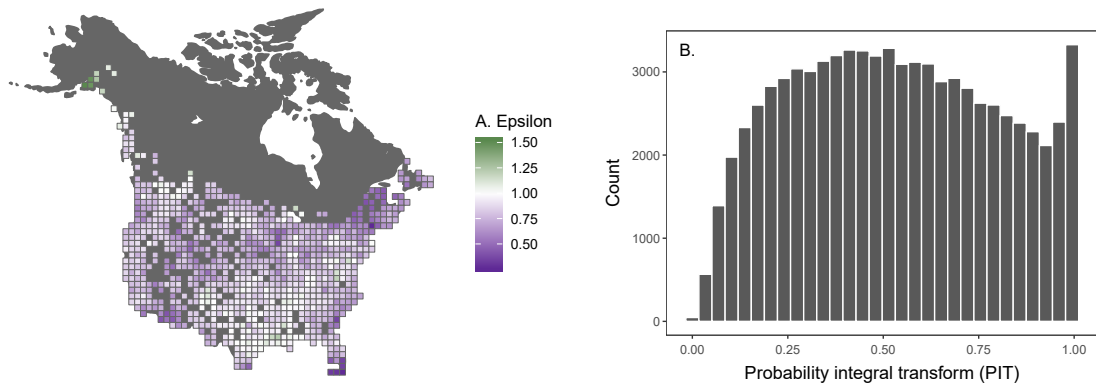


Figure 6: Supplemental figures showing spatial variation in  $\epsilon_i$  values (A) and a histogram of PIT values from the SVC model for American Robin counts (B). The approximately uniform distribution of PIT values suggests that the predictive distribution of the model is consistent with the data.