Modeling First Arrival of Migratory Birds using a Hierarchical Max-infinitely Divisible Process

Dhanushi A. Wijeyakulasuriya
Microsoft Corporation
1 Microsoft Way, Redmond, WA, 98052

Ephraim M. Hanks
Department of Statistics
Pennsylvania State University
University Park, PA, 16802

Benjamin A. Shaby
Department of Statistics
Colorado State University
Fort Collins, CO, 80523

April 30, 2024

Abstract

Humans have recorded the arrival dates of migratory birds for millennia, searching for trends and patterns. As the first arrival among individuals in a species is the realized tail of the probability distribution of arrivals, the appropriate statistical framework with which to analyze such events is extreme value theory. Here, for the first time, we apply formal extreme value techniques to the dynamics of bird migrations. We study the annual first arrivals of Magnolia Warblers using modern tools from the statistical field of extreme value analysis. Using observations from the eBird database, we model the spatial distribution of observed Magnolia Warbler arrivals as a max-infinitely divisible process, which allows us to spatially interpolate

observed annual arrivals in a probabilistically-coherent way, and to project arrival dynamics into the future by conditioning on climatic variables.

Keywords: Extreme value theory; Max-stable process; Spatial extremes

1 Introduction

Patterns in spring bird migrations are key indicators of ecosystem responses to climate pressures. For millenia, observers have used the annual first arrival of a given species to demarcate the migration (Lincoln, 1935). The date of first arrival is a statistical extreme value; that is, the first individual from among a large population to reach a given location is exactly the realized tail of the probability distribution of all arrivals. The statistical theory of extreme values (Beirlant et al., 2004; Davison et al., 2019) is thus ideally suited for modelling spring first arrivals, but this theory has never been used for modeling first arrivals. Ours is the first such study to deploy modern techniques from the statistics of extremes to model this phenomenon. In this analysis, we use state-of-the-art tools for spatial extremes to model and predict the first arrival of migratory birds. We show that using these methods allows for principled inference on the relationship between landscape or climatic variables and first arrival times, predictions at unobserved locations during past years, and predictions of future first arrivals under climate model projections.

We frame the problem of modeling the first arrival of migratory birds as a problem of modeling spatial extremes. We apply this method to study and predict the Spring arrival of Magnolia Warblers (*Setophaga magnolia*) in the Northeast portion of the United States from 2004–2019 using data from the eBird database (eBird, 2019), a citizen science website. By using hierarchical spatial extreme value models, we are able to obtain conditional predictions of first arrival dates at locations without eBird observations.

We chose to study the Magnolia Warbler based on recommendations from our ornithologist collaborators, who pointed out several advantageous characteristics of this species. First, it is a long-distance migrant whose summer (breeding) range is well-separated geographically from their winter (nonbreeding) range. The non-overlapping characteristic of their summer

and winter ranges eliminates ambiguity with respect to whether any observed individual has already migrated north or has yet to begin migration. Second, the eBird database contains a large number of observations of Magnolia Warblers, which helps in fitting a complicated statistical model. Finally, the Marginalia Warbler is relatively easy to identify, which helps to mitigate the problem of mis-identification which is inherent in citizen science data. The species is uncommon west of the Mississippi river despite regular vagrancy on the West Coast (Dunn and Hall, 2020). We focused our study on the northeast United States due to the prevalence of Magnolia Warbler sightings in this region.

Previous studies have investigated the mismatch between the arrival of migratory birds and other aspects of the onset of Spring (Koleček et al., 2020; Kullberg et al., 2015; Tøttrup et al., 2010; Jonzén et al., 2006). Most used linear regression or similar methods that regress a measure of first arrival time on covariates related to the onset of Spring (Møller et al., 2008; Mayor et al., 2017; Gunnarsson and Tómasson, 2011). Some methods incorporated multiple species using a random effect (Koleček et al., 2020; Kullberg et al., 2015). Ambrosini et al. (2014) used a binomial conditional auto-regressive mixed model, and is the only instance of which we are aware that employed a formal spatial model for first arrivals. There is little uniformity in the definition of first arrival times in existing literature. Palm et al. (2009), Gunnarsson and Tómasson (2011), and Koleček et al. (2020) calculated first arrival date for a given year to be the mean of first arrival dates recorded by observers averaging over all geographical locations. Another approach is to fit logistic or cumulative log log functions to estimate the first arrival time as the inflection point or specific percentiles (Mayor et al., 2017; Ambrosini et al., 2014). Zaifman et al. (2017) used a set of heuristics to filter out noisy data points from the eBird database to identify first arrival times. All of these approaches modeled first arrival times using standard, mean-focused statistical models, rather than modeling first arrival times as the extreme values that they are.

We construct our extreme value models by conceptualizing the geographical map of first arrivals as a spatial field of block minima. That is, at each point in space, a subset of the population of Magnolia Warblers visits, with each member of that population arriving at a particular time. In the parlance of extreme value statistics, this collection of arrival times is referred to as a "block", and the earliest arrival time is the block minimum. By singling out the first arrival time at each of many spatial locations, we arrive at a spatial field of block minima. This is exactly the structure of data that is the subject of the statistical study of spatial extremes. The enterprise of modeling spatial extremes has most often focused on extreme weather events like extreme precipitation, temperature, and wind (Reich and Shaby, 2012; Shaby and Reich, 2012; Huser and Wadsworth, 2019; Reich and Shaby, 2019). Extreme value methods have only been rarely used in ecological studies (Wijeyakulasuriya et al., 2019).

Widely-used models for spatial statistics based on Gaussian processes are appropriate for modeling spatial events in the bulk of a distribution, but they do not do well for spatial extremes because they have rapidly vanishing tail dependence which is usually not realistic for extremes (Bopp et al., 2021). Furthermore, they do not posses the max-infinite divisibility (max-id) property, which we will argue is necessary for any coherent model of block minima. Instead of using Gaussian processes, we model the timing of the Magnolia Warbler Spring migration using the hierarchical model of Bopp et al. (2021), which is max-id (and therefore appropriate for block minima like first arrival times), has flexible tail dependence properties that can be learned from data, and decomposes in such a way as to make computing tractable on large datasets.

The hierarchical modeling approach also makes it easy to incorporate covariates. We fit sixteen years of first arrival data from eBird, regressing model parameters on topographic, landcover, demographic, and climatological predictors. We then create predictive maps

of first observed spring arrival for the Magnolia Warbler in the northeast United States (Section 3). For an example of eBird data and the model's predicted map, see Figures 1(a) and 1(b). We then use climate model data from the fifth phase of the Coupled Model Intercomparison Project (CMIP5) for years 2151–2200 to demonstrate how these models can be used to make future predictive maps of first observed arrival (See Figures 10(1) and 10(b)).

1.1 Data Processing

We downloaded bird sightings data for the Magnolia Warbler from the citizen data website eBird (eBird, 2019). We focused our study on the northeast US and only considered years from 2004–2019 due to the relative lack of data before that period. The Magnolia Warbler winters in the neotropics. First spring migrants arrive in southern United States in early April. Since we were looking at Spring arrivals, we included sightings from March 20 to July 20 of each year. Any single observation of a Magnolia Warbler could be either an individual that is in transit to its summer breeding location or has already reached its final breeding location. We do not differentiate between these two cases.

First arrival data can be affected by observation effort and be very noisy (Lindén, 2011). Our approach to calculating first arrival time is similar to that of Zaifman et al. (2017), who used the same data source. In order to minimize the influence of potentially aberrant observations, we calculated first arrival time only in counties that had at least 12 sightings during the time period March 20 (roughly Spring equinox) through July 20 in a given year. We calculate the first arrival to be the number of days from March 20. There is a large proportion of missing data (80%) in the overall dataset. See Figure 1 (a) for the observed first arrival times for 2019. Counties colored in grey do not have first arrival observations in 2019.

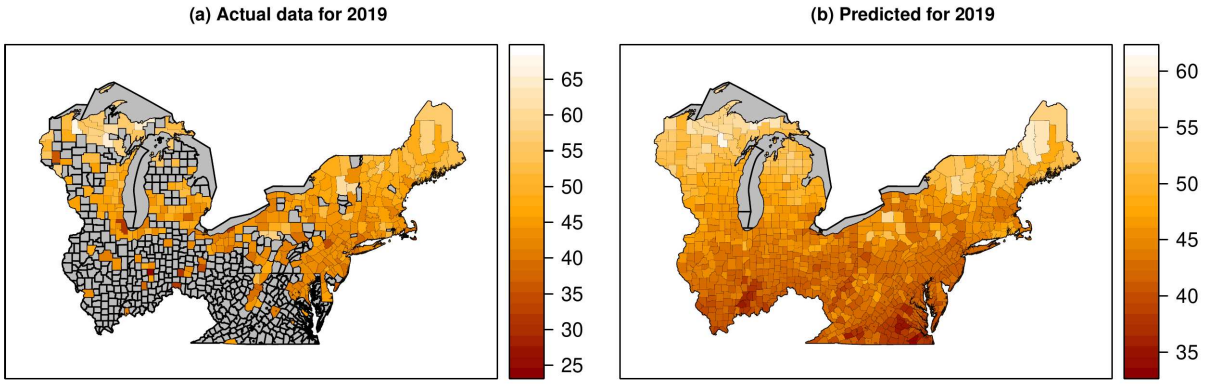

Figure 1: (a) First arrival data for 2019. Missing values are shown in grey. (b) Predicted First arrival for 2019.

Important biases are present in the eBird dataset, which should be considered when interpreting the results that we present. For example, observation effort (i.e. number of active bird watchers contributing to eBird) has increased dramatically over the years (see Figure S-2 in the Supplement). We attempt to minimize this effect to some degree by treating as missing any observations for counties in years where sightings are sparse. Another issue is preferential sampling, which arises because observers tend to go where they think they will find bird activity. This is a complicated problem that is beyond the scope of our study, but presents an interesting challenge for follow-up work. For these reasons, we take care to interpret our results as analysis of the process of *observed* first arrivals, rather than the process of first arrivals themselves, which is in some sense un-observable.

We included spatial and climate covariates to model variation in first arrival times. The spatial covariates are longitude, latitude, elevation (Hollister and Tarak Shah, 2017), forest cover (Blackard et al., 2008), proportion of water, and population density (U.S. Census Bureau, 2010). Longitude and latitude covariates were taken to be the centroid of each county. All other covariates are the spatial average over the county. The climate covariates

are temperature anomaly (Zhang et al., 2010) and North Atlantic Oscillation (NAO) (Hurrel, 2003) for the month of March of each year. For more details about the first arrival data and spatial and climate covariates, see Appendix S-1 in the Supplementary Material.

2 Hierarchical Max-Infinitely Divisible Spatial Process Model

Models for extreme values are defined for the right tail by convention, as the most common extremes of practical interest are maxima, not minima. Analysis of the earliest arrival concerns the left tail, however, so we will simply multiply the data (in units of days since March 20) by -1 and add a constant large enough to make all of the negated values positive, and proceed with the conventional extreme value terminology of referring to annual maxima, rather than minima. If we imagine choosing a single spatial location and consider the (negative) arrival date Z of each individual bird that passes through, then that variable Z has some cumulative distribution function (*cdf*) $G(z) = P(Z \leq z)$. Then, making the (unrealistic) working assumption that each individual's arrival is an independent and identically distributed copy of Z, the *cdf* of the maximum (negative) arrival from a population of size n is $H(z) = P(\max_{Z_1,\ldots,Z_n} \leq z) = G^n(z)$.

But rather than a single location, we are interested in the random vector $\mathbf{Z}$, the maximum (negative) arrival at all locations of interest, whose joint *cdf* is $H(\mathbf{z}) = P(\max_{\mathbf{Z}_1,\ldots,\mathbf{Z}_n} \leq \mathbf{z}) = G^n(\mathbf{z})$. Our goal is to model this distribution of spatial maxima, $H(\mathbf{z})$. The population size n is always going to be unknown, but more fundamentally, our model should remain valid even if the n changes from year to year. A coherent model for $H(z)$, then, would require that $H^{1/n}(\mathbf{z})$ be a valid joint distribution for any n. This is exactly the max-infinite divisibility property. Thus, we require any spatial model for (negative) first arrivals to

be max-id. Similar reasoning applies to random vectors that are not independent across individuals, like arrival dates of migratory birds.

To define the spatial process $Z_t(\mathbf{s})$ of (negative) first arrivals at year t, now written explicitly as a function of spatial location $\mathbf{s}$, we first define the process $Y_t(\mathbf{s})$ as a combination of basis functions,

$$Y_t(\mathbf{s}) = \left\{ \sum_{l=1}^{L} A_{tl} K_l(\mathbf{s})^{1/\alpha} \right\}^{\alpha}. \tag{1}$$

Here, α is a parameter that controls the smoothness of the process, and $A_{t1}, \ldots, A_{tL}$ are independent and identically distributed (iid) scaling coefficients, for $t = 1, \ldots, T$. The scaling coefficients have an exponentially-tilted positive stable distribution, with parameters α and θ that together control the strength of the spatial tail dependence of the resultant process (Bopp et al., 2021).

The functions $K_1(\mathbf{s}), \ldots K_L(\mathbf{s})$ in (1) form a collection of L spatial basis functions, which in combination form the surface of the spatial process of first arrivals (see Figure 2). We do not know *a priori* what shape of basis functions will result in the best-fitting combination, so we estimate the shape of these functions by assigning them prior distributions based on Gaussian processes. The construction requires that basis functions be positive and satisfy a sum-to-one constraint, so we specify the priors by transforming independent mean-zero stationary Gaussian processes $\tilde{K}_1(\mathbf{s}), \ldots, \tilde{K}_{L-1}(\mathbf{s})$ as $K_l(\mathbf{s}) = \exp\{\tilde{K}_l(\mathbf{s})\} / \sum_{i=1}^{L} \exp\{\tilde{K}_i(\mathbf{s})\}$, $l = 1, \ldots, L$, with $\tilde{K}_L(\mathbf{s}) \equiv 0$ to complete the specification. In this way, the shape of the basis functions can be learned from the data.

The form of the basis combination (1) resembles a spatial factor model, where traditionally α is taken to be 1 and $A_{t1}, \ldots, A_{tL}$ are iid Gaussian. However, the particular L^p norm (with $p = 1/\alpha$) construction of (1) and the particular tilted stable distribution of $A_{t1}, \ldots, A_{tL}$ are the keys to obtaining the desired max-id property (Bopp et al., 2021).

To complete the model, we introduce an everywhere multiplicative "nugget" effect $\epsilon(\mathbf{s})$ with iid Fréchet$(1/\alpha)$ marginal distributions. The final spatial max-id model for first arrivals is thus

$$Z_t(\mathbf{s}) = \epsilon_t(\mathbf{s}) Y_t(\mathbf{s}). \tag{2}$$

The nugget effect $\epsilon_t(\mathbf{s})$ represents small-scale variation. In our case, small-scall variation is particularly relevant because it can capture local habitat or resource variation that is not present in the covariates. For example, patches of intact forest land or preferred food sources, or even a neighborhood cat, are too small-scale to be captured in county-level data, but can be captured in the model by $\epsilon_t(\mathbf{s})$.

Univariate extreme value theory says that the only possible limit of appropriately rescaled block maxima is the Generalized Extreme Value (GEV) distribution. This classical result suggests that marginally (i.e. at any location), the distribution of the first arrival will be well-approximated by the GEV distribution as the population grows. The classical convergence result may not apply precisely to the present case, as issues have been identified with data that are discrete and bounded (Hitz et al., 2017), like the (negated) first arrival dates that we analyze here. A simple approach to reconcile the discrete data with the continuous limiting distribution is to consider it as interval censored. This works well, but in our experience has negligible impact on the resultant analysis (Bopp and Shaby, 2017). We therefore regard the GEV as an appropriate model for the first arrivals. The Bopp et al. (2021) model includes a marginal transformation to GEV, from the distribution implicitly defined by (2), inside the model hierarchy. This allows flexible modeling of the marginal surfaces, including dependence on covariates, as well as uncertainty propagation between the marginal and joint components of the model.

The GEV(μ, σ, ξ) is a three-parameter distribution with location parameter $\mu \in \mathbb{R}$, scale parameter $\sigma > 0$, and shape parameter $\xi \in \mathbb{R}$.

We define $\tilde{Z}_t(\mathbf{s}) = \text{GEV}^{-1}[G_{\mathbf{s}}\{Z_t(\mathbf{s})\}; \mu_t(\mathbf{s}), \sigma_t(\mathbf{s}), \xi_t(\mathbf{s})]$ as the (negative) first arrival date on the original observation scale, where $GEV^{-1}\{\cdot; \mu_t(\mathbf{s}), \sigma_t(\mathbf{s}), \xi_t(\mathbf{s})\}$ is the quantile function of a GEV distribution with parameters $\mu_t(\mathbf{s}), \sigma_t(\mathbf{s})$ and $\xi_t(\mathbf{s})$. $G_{\mathbf{s}}(z)$ is the marginal distribution function of $Z_t(\mathbf{s})$, which is implicitly defined by the construction (1) and (2). Critically, the max-id property of the model is preserved when using this transformation. Therefore, $\tilde{Z}_t(\mathbf{s})$ is the observed (negative) first arrival date, with $\text{GEV}(\mu_t(\mathbf{s}), \sigma_t(\mathbf{s}), \xi_t(\mathbf{s}))$ marginal distribution, whereas $Z_t(\mathbf{s})$ is the same quantity, under the transformation to the marginal distribution induced by (1) and (2).

We assume the first arrival process $\tilde{Z}_t(\mathbf{s})$ is independent across years, given a collection of covariates that may vary in time. Let $\tilde{Z}_t(\mathbf{s})$ to be the process observed at location $\mathbf{s}$ at year t. Marginal GEV parameters vary with time via the climate covariates as

$$\begin{aligned} \mu_t(s) &= \mathbf{X}_{t\mu}(\mathbf{s})\boldsymbol{\beta}_\mu + W_\mu(\mathbf{s}) \\ \log \sigma_t(\mathbf{s}) &= \mathbf{X}_{t\sigma}(\mathbf{s})\boldsymbol{\beta}_\sigma + W_\sigma(\mathbf{s}) \\ \xi_t(\mathbf{s}) &= \xi, \end{aligned} \tag{3}$$

where $\mathbf{X}_{t\mu}(\mathbf{s})$ and $\mathbf{X}_{t\sigma}(\mathbf{s})$ are matrices of climatological and demographic covariates, some of which vary in space or time, $\boldsymbol{\beta}_\mu$ and $\boldsymbol{\beta}_\sigma$ are vectors of regression coefficients, and $W_\mu(\mathbf{s})$ and $W_\sigma(\mathbf{s})$ are zeromean Gaussian processes. The spatially varying basis functions are common across time, whereas the scaling coefficients of the basis functions, $A_{tl}, l = 1, \ldots, L$ and $t = 1, \ldots, T$, vary with time. The data did not exhibit evidence of temporal non-stationarity in the spatial dependence parameters α or θ.

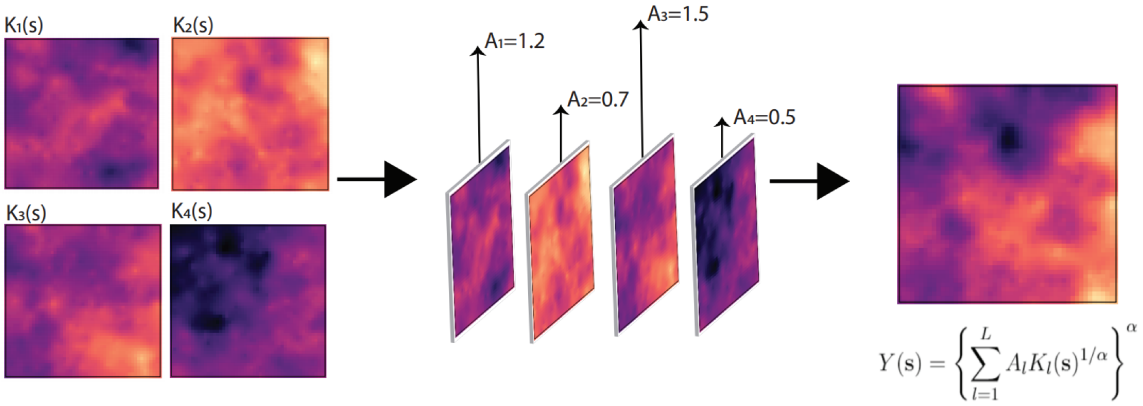

Figure 2: Illustration of the construction of $Y(s)$ using $K_s(s)$ and A_l

2.1 Model Fitting

We transform our first arrival dates, of which the extreme arrivals are minima (earliest day of the year) to be maxima by multiplying by -1 and adding a large enough constant such that all values are positive. We do all steps of the analysis using the negated data. When plotting predictions, we transform them back to calendar dates for ease of interpretation.

We follow common practice by fixing $\xi_t(\mathbf{s})$ to be constant in space and time, with prior $\xi \sim N(0, 100)$, as it is notoriously difficult to discern spatial variation in this parameter (Cooley et al., 2007; Bopp et al., 2021). We assume a Gaussian process prior for $\mu_t(\mathbf{s})$ that is constant across time, thereby allowing the marginal location parameter of first arrival date to vary flexibly across space. We use spatial and climate covariates to model the mean function of the Gaussian process. We selected the set of covariates with the lowest AIC and BIC scores. We use goodness of fit measures to select the number of basis functions L, with $6, 8, 10, 12$ and 14 as candidate values for L.

3 Results

3.1 Model Comparison

We used out of sample predictive log scores to select the number of basis functions L and to decide between modeling $\log \sigma_t(\mathbf{s})$ as a Gaussian process or as a fixed linear model. For this model selection, we used a subset of 114 counties where at least 10 years of data were present. We then randomly sampled 12 counties as the out of sample set. We fit the model on the rest of the data and evaluated the log likelihood, given the MCMC samples of the model parameters, of the out of sample data set, yielding a log score for each set of posterior samples. We used a 95% trimmed mean to calculate the average log score for each candidate model. The results are given in Table S-1 in Appendix S-4 in the Supplementary Material. The best model under this approach is the model with 8 basis functions and $\log \sigma(\mathbf{s})$ as a fixed linear model.

3.2 Final Fitted Model

We then fit this model to (negative) first arrival dates from all 869 counties and all years. We used draws from posterior predictive distributions to predict first arrival times at the counties without first arrival observations.

Posterior means and 95% credible intervals for α, θ and ξ are 0.3340 (0.3057, 0.3604), 0.00018 (0.000028, 0.00057) and -0.4095 (-0.4435, -0.3784). Posterior means and 95% credible intervals for the location parameter and scale parameter coefficients are given in Figures 3 and 4, respectively. These results correspond to negated first arrival data (i.e. maxima), so that larger values for the location parameter correspond to earlier first arrivals. Latitude has the most negative coefficient from the location parameter covariates, indicating

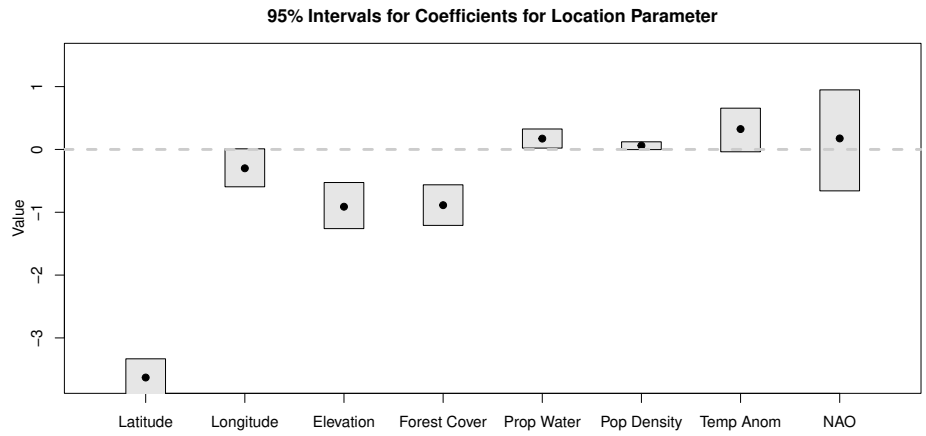

Figure 3: Coefficient estimates and 95% credible intervals for covariates in the mean function of the location parameter. Transformed (i.e. negated) first arrival data is used here. Larger values for the location parameter correspond to earlier first arrival.

that counties at lower latitudes have earlier arrivals, as expected. Elevation also has a negative coefficient, indicating that regions at higher elevation have later arrivals on average. Forest cover has a negative coefficient. Population density has a slightly positive coefficient. This effect could be a proxy for observation effort, wherein more densely populated regions have more people available to observe the earliest warbler arrivals. Temperature anomaly and NAO both have positive point estimates. Higher values for NAO and temperature anomaly corresponds to warmer weather, which in turns leads to earlier arrival. This too is consistent with what we expect.

Interpreting the coefficients for the scale parameter covariates is not as straightforward. We can only say that covariates with higher coefficients lead to more variability in first arrivals. For example, regions at higher elevation or more forest cover have higher variability in the first arrival dates.

Posterior means of the eight random basis functions are plotted in Figure 5. They are

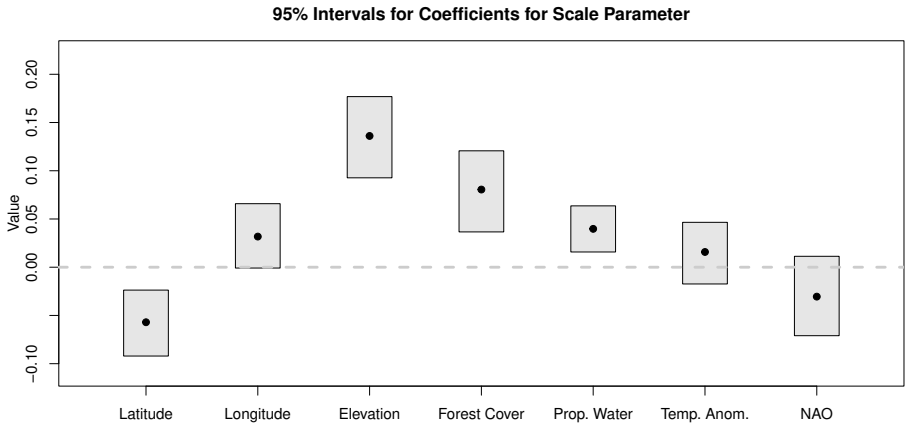

Figure 4: Coefficient estimates and 95% credible intervals for covariates in the mean function for the scale parameter

ordered by the variance of their corresponding coefficients, in a manner analogous to the ordering of principle components (also known as empirical orthogonal functions). The first six basis functions account for over 85% of the variability. These spatial patterns can be interpreted as locations where early or late arrivals tend occur together in the same year, perhaps because they represent migration corridors.

We also calculated the median first arrival date using posterior sample values for $\mu_t(\mathbf{s})$, $\sigma_t(\mathbf{s})$ and ξ, using the formula $z_{0.5}(\mathbf{s}) = \hat{\mu}_t(\mathbf{s}) + \frac{\hat{\sigma}_t(\mathbf{s})}{\hat{\xi}} \left((-\log(0.5))^{-\hat{\xi}} - 1 \right)$. Here, we averaged over temperature anomaly and NAO and then averaged over the posterior samples and have plotted median first arrival in Figure 6. Dark red corresponds to earlier arrival, while light yellow corresponds to later arrival. As evident in Figure 6, the earliest median first arrival of Magnolia Warblers occurred in Illinois and Virginia. Higher elevation regions of West Virginia and Pennsylvania had later arrivals. The states in the extreme northeast like Maine, New Hampshire, and Vermont, as well as the northern regions of Wisconsin and Michigan, had latest median first arrival.

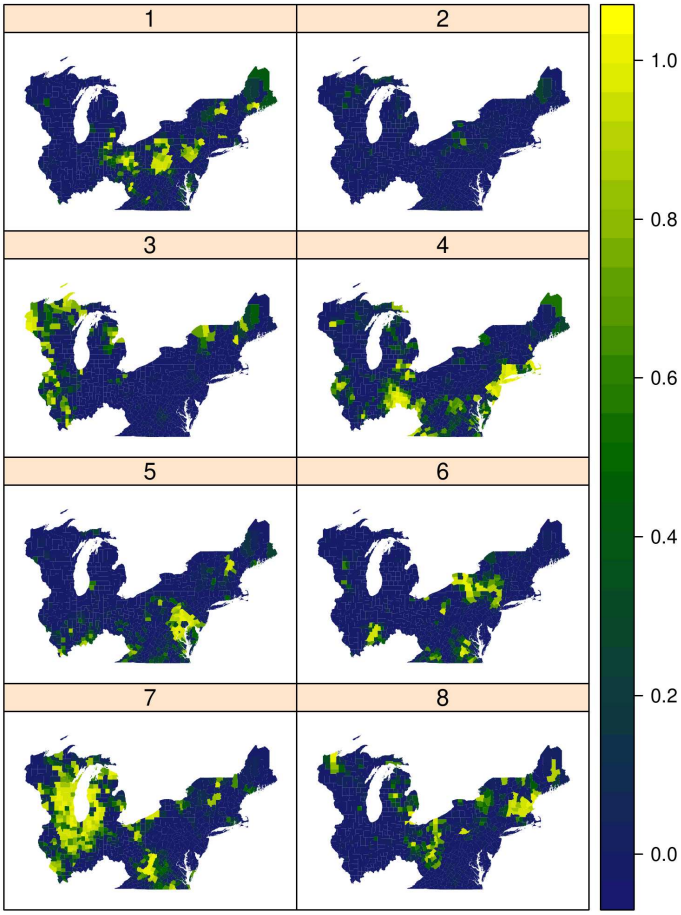

Figure 5: Posterior estimates of the spatial basis functions, ordered by variance of the corresponding random basis coefficients, from largest to smallest.

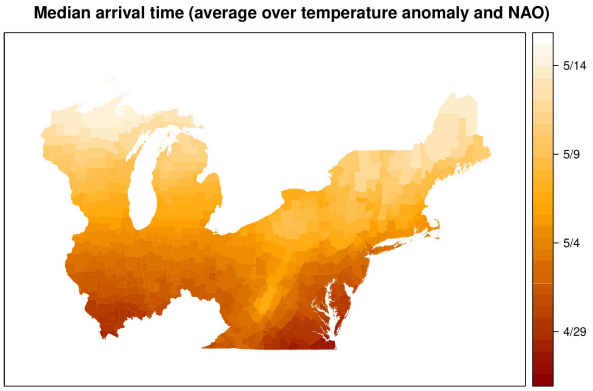

Figure 6: Estimated median first arrival of Magnolia Warblers for 2016–2019, averaging over temperature anomalies and NAO, i.e. predicting using 0 as covariate values for both temperature anomaly and NAO.

3.2.1 Model Checks

We examined several diagnostic plots to assess the adequacy of both the marginal and spatial components of the fit. Figure 7 shows a *qq*-plot of the observed data against a standard Gumbel. The observations at each location and time were transformed to standard Gumbel using the posterior samples of the marginal parameters from the marginal model (3), and then averaged across MCMC iterations. Figure 7 shows good agreement between the model and observations for most of the first arrivals. There is noticeable departure from $1:1$ line in the upper tail, but there is also a huge amount of variation in this region. Overall, the marginal fit of the model seems adequate.

Figure 8 shows empirical and model-based estimates of the coefficient $\chi_h(u)$ for points separated by $h = 50$km and $h = 500$km. The coefficient $\chi_h(u) = P[F_1\{Y(\mathbf{s}_1)\} > u \,|\, F_2\{Y(\mathbf{s}_2)\} > u]$, for points $\mathbf{s}_1$ and $\mathbf{s}_2$ separated by a distance h, where F_1 and F_2 are the marginal *cdf*s of $Y(\mathbf{s}_1)$ and $Y(\mathbf{s}_2)$, respectively, is a common measure of the strength

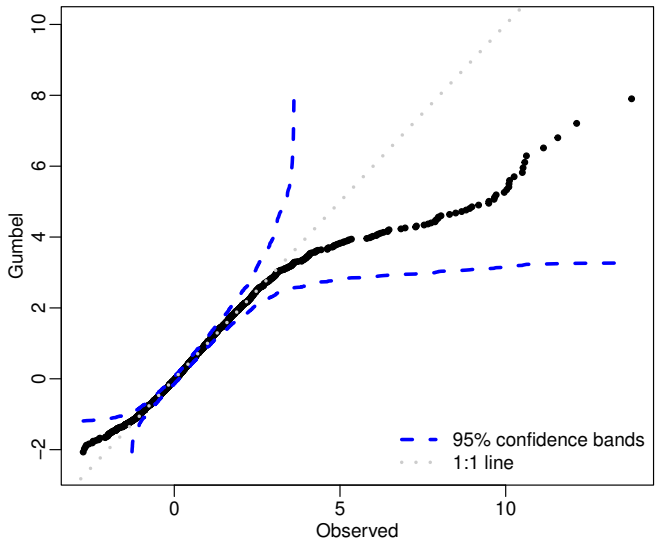

Figure 7: Quantiles of the observed data plotted against quantiles of the standard Gumbel distribution. The observations were transformed to standard Gumbel using posterior samples of the marginal parameters. The region enclosed by the blue dotted curves is the pointwise 95% confidence interval.

of tail dependence. Empirical estimates of this conditional probability are subject to a great deal of uncertainty at high quantiles u, especially in this case because so much of the data is missing. However, despite the very wide confidence intervals, Figure 8 shows good agreement between function $\chi_h(u)$ estimated empirically and $\chi_h(u)$ calculated from the fitted model, at both short and long separation distances. We conclude that the model fits the tail dependence in the data well.

3.3 Predictive Maps

Figure 9 gives the mean posterior predictive first arrival for 2019, as well as the difference between 2019 and 2016, 2017, and 2018. In the difference plots, the blue color denotes

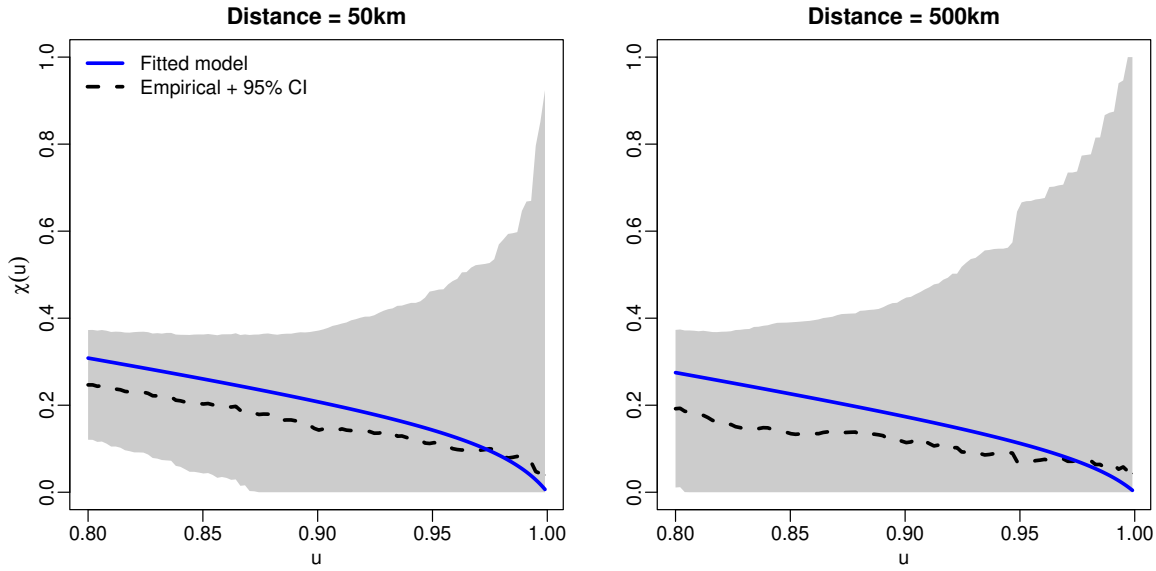

Figure 8: The coefficient $\chi_h(u)$ calculated at separation distances $h = 50$km (left) and $h = 500$km (right). The blue solid curves show $\chi_h(u)$ calculated from the fitted model, while the black dotted curves show empirical esimates of $\chi_h(u)$, with the grey regions denoting pointwise 95% confidence intervals for the empirical estimates.

later arrival while the red color denotes earlier arrival. In 2018 a majority of the counties had earlier arrival compared to 2019 where as in 2017 most counties had later arrival compared to 2019. In 2016 there is a mix of red and blue, with counties in the western region having earlier arrivals while Virginia and parts of the Eastern seaboard having later arrivals. Compared to the median maps, these posterior predictive maps show more small-scale spatial variability, as we are predicting the actual date of first arrival at a given year, rather than an average date. However we see the expected trend of first arrival times being later at higher latitudes compared to lower latitudes. See Appendix S-6 Figure S-7 in the Supplementary Material for standard deviations of the posterior predictive first arrival times for 2019.

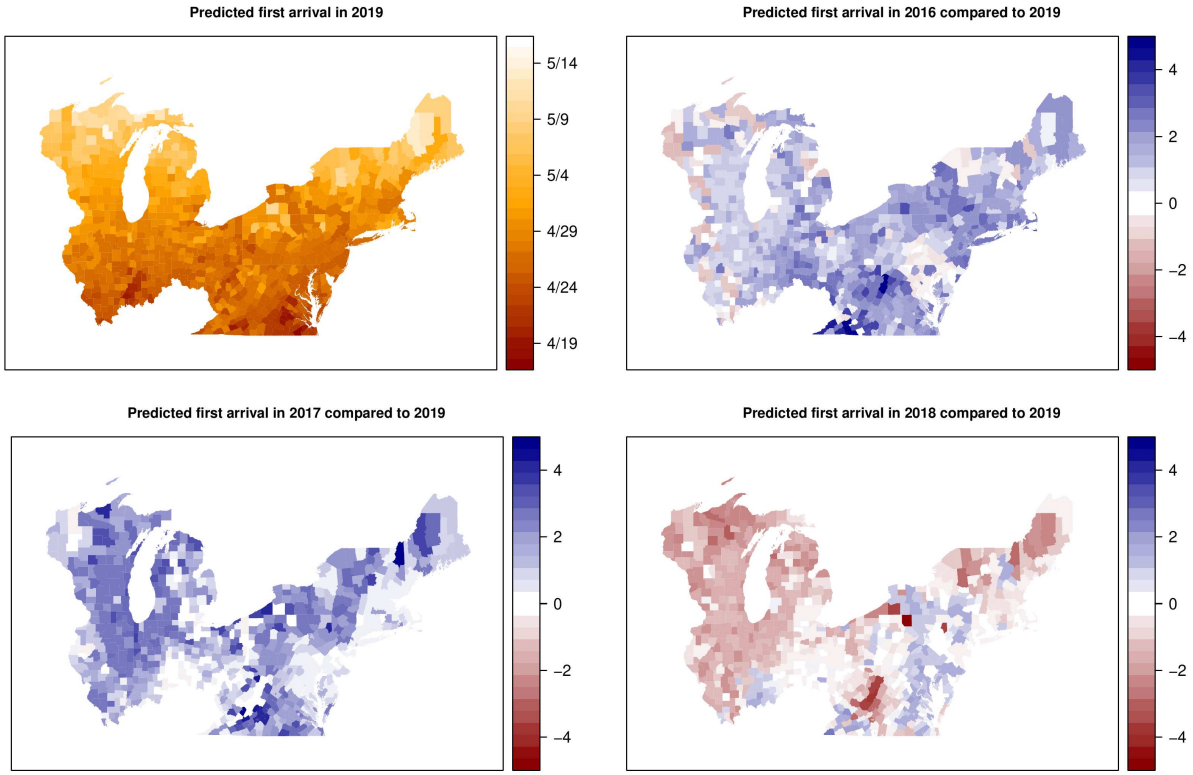

Figure 9: Posterior mean predictive plot for first arrival in years 2016–2019.

We obtained monthly climate model output from the CMIP5 data portal (https://esgf-node.llnl.gov/search/cmip5/). The spatial resolution of the dataset is 0.5×0.5 degrees. We use the climate model GISS-E2-H from NASA's Goddard Institute for Space Studies and picked Representative Concentration Pathway (RCP) scenario 'high' (RCP 8.5). RCP8.5 corresponds to high greenhouse gas emissions and is the upper bound of the available RCPs. It is commonly used as a baseline scenario that does not account for any specific climate mitigation strategies (Riahi et al., 2011). We used ensemble member `r1i1p1` in this study and extracted data for the time period 2151–2200.

Figure 10 (a) gives the difference between the average projected first arrival dates for 2151–2200 and the first arrival date for 2019 (the last year in our study with actual data). A blue hue denotes later arrival compared to 2019 while, a red hue denotes earlier arrival compared to 2019. As evident in Figure 10 (a), over 80% of counties in our region of interest are projected to have earlier arrivals compared to 2019. Some counties in the Appalachian region as well as some with higher forest cover are projected to have slightly later arrivals, as indicated by the light blue hue. Figure 10 (b) gives the standard deviations of the posterior predictive samples. Higher elevation and higher forest cover regions West Virginia showed high standard deviation values. For difference in first arrival for individual years and their standard deviations, see Figure S-10 and Figure S-11 respectively in Appendix S-6 in the Supplementary Material. These predictive maps show a general trend of earlier Magnolia Warbler arrivals under the projected future climate, relative to 2019.

4 Discussion

In this study we frame the problem of modeling first arrival of migratory birds as a spatial extremes problem, recognising that first arrival is the realization of the tail of the probability

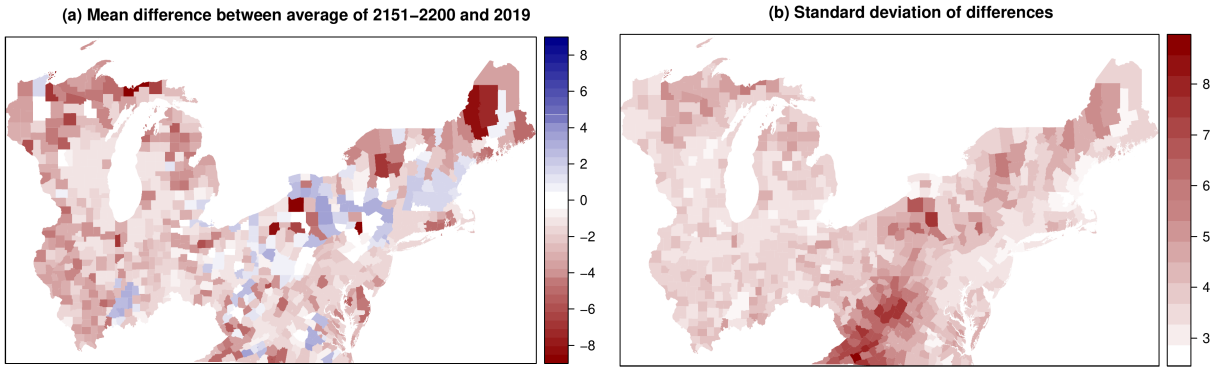

Figure 10: (a) Difference in predicted first arrival using base year 2019 for CMIP data averaged over years 2151–2200. (b) Standard deviations of the predicted differences

distribution of arrivals. We adapt the hierarchical max-infinitely divisible process model of Bopp et al. (2021) to model first arrival of the long distance migrant the Magnolia Warbler, obtaining maps of first spring arrival for the period 2004–2019. We use the posterior predictive distribution from the model to interpolate first arrival dates to counties with missing data in a probabilistically coherent way.

Furthermore, we demonstrated how this method can be used in conjunction with climate model output to produce predictive maps of spring first arrival for future years under projected future climate. Based on the CMIP5 RCP8.5 climate model output we used, we found that first arrival of Magnolia Warblers will be earlier in over 80% of the counties in the region of interest. We also provided uncertainty estimates for these predictions. Running the model on output from a suite of climate models forced with different emissions scenarios would give a richer picture of the possible trajectories of first arrivals. However, our goal here is not so much to produce a reliable forecast—rather, we demonstrated that we can combine a spatial extremes model with a climate projection to get a plausible picture of the future, conditional on the quality of the climate projection.

We found that latitude, elevation, and forest cover had negative coefficients when modeling the location parameter of the marginal GEV, indicating later arrival for counties with higher values for these covariates. This is an expected result for latitude but unexpected for forest cover. The forest cover covariate did not distinguish between different types of forests, which could be useful to determine the relationship between forest cover and first arrival more precisely. Both forest cover and elevation had positive coefficients for the scale parameter of the GEV, indicating greater variability for larger values of these variables.

As expected, temperature anomaly had a positive coefficient for the location parameter, indicating that warmer than average March temperatures result in earlier arrivals of the spring migrants in April. Although the credible interval for NAO for the location parameter overlaps zero, it can be interpreted as having largely positive values, which signals that higher NAO values correspond to earlier arrivals.

We fitted the model at the county level, aggregating over observations made within each county. It was necessary to do aggregation at some level since observations are not made at the same locations every year and we need fixed spatial locations to fit the model proposed by Bopp et al. (2021). We used counties as our observation unit since it is easier to obtain values of covariates aggregated at this level. However, not all counties are similarly sized, which could lead to our heuristics being very stringent on smaller counties. Another approach to do this could be to use an evenly spaced grid over the region of interest.

This study provides a rigorous basis for studying the first arrival of migratory birds across space and time. It enables prediction of first arrival at unobserved locations while also providing powerful tools to understand the ecosystem response to different climate scenarios.

SUPPLEMENTARY MATERIAL

Supplementary Material: Additional plots, model diagnostics, etc. (.pdf file)

R-package stablemixEco for fitting the model: A modified version of the package stablemix. To improve mixing, we have changed the way the spatially-varying fields are updated (see the Supplementary Materials for details). (GNU zipped tar file)

First arrival data: Data set used for fitting the model. (.RData file)

Climate projection covariates: Data set used for to project first arrival dates under a climate model projection. (.RData file)

ACKNOWLEDGMENTS

We acknowledge input by Dr. Viviana Ruiz Gutierrez and Dr. Daniel Fink of the Cornell Lab of Ornithology. We would also like to thank the two anonymous referees for their helpful comments and suggestions. The authors gratefully acknowledge support from the US National Science Fountation under NSF DMS-2001433, NSF DMS-2308680, and NSF DMS-2015273. Computations for this research were performed on the Pennsylvania State University's Institute for CyberScience Advanced CyberInfrastructure (ICS-ACI). This content is solely the responsibility of the authors and does not necessarily represent the views of the Institute for CyberScience. This work has no connection to Microsoft Corporation and was conducted by D.A.W. during her doctoral studies, prior to joining Microsoft Corporation.

Declarations

Conflict of Interest Statement The authors declare that there are no conflicts of interest.

References

Ambrosini, R., Borgoni, R., Rubolini, D., Sicurella, B., Fiedler, W., Bairlein, F., Baillie, S. R., Robinson, R. A., Clark, J. A., Spina, F., and Saino, N. (2014), "Modelling the Progression of Bird Migration with Conditional Autoregressive Models Applied to Ringing Data," *PLOS ONE*, 9, 1–10.

Beirlant, J., Goegebeur, Y., Teugels, J., and Segers, J. (2004), *Statistics of extremes*, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Chichester, theory and applications, With contributions from Daniel De Waal and Chris Ferro.

Blackard, J., Finco, M., Helmer, E., Holden, G., Hoppus, M., Jacobs, D., Lister, A., Moisen, G., Nelson, M., Riemann, R., Ruefenacht, B., Salajanu, D., Weyermann, D., Winterberger, K., Brandeis, T., Czaplewski, R., McRoberts, R., Patterson, P., and Tymcio, R. (2008), "Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information," .

Bopp, G. P. and Shaby, B. A. (2017), "An exponential-gamma mixture model for extreme Santa Ana winds," *Environmetrics*, 28, e2476, 14.

Bopp, G. P., Shaby, B. A., and Huser, R. (2021), "A Hierarchical Max-Infinitely Divisible Spatial Model for Extreme Precipitation," *J. Amer. Statist. Assoc.*, 116, 93–106.

Cooley, D., Nychka, D., and Naveau, P. (2007), "Bayesian spatial modeling of extreme precipitation return levels," *J. Amer. Statist. Assoc.*, 102, 824–840.

Davison, A., Huser, R., and Thibaud, E. (2019), "Spatial extremes," in *Handbook of environmental and ecological statistics*, CRC Press, Boca Raton, FL, Chapman & Hall/CRC Handb. Mod. Stat. Methods, pp. 711–744.

Dunn, E. H. and Hall, G. A. (2020), "Magnolia Warbler (*Setophaga magnolia*)," .

eBird (2019), "eBird Basic Dataset," http://www.ebird.org, accessed: 2019-12-25.

Gunnarsson, T. G. and Tómasson, G. (2011), "Flexibility in spring arrival of migratory birds at northern latitudes under rapid temperature changes," *Bird Study*, 58, 1–12.

Hitz, A., Davis, R., and Samorodnitsky, G. (2017), "Discrete extremes," *arXiv preprint arXiv:1707.05033*.

Hollister, J. and Tarak Shah (2017), *elevatr: Access Elevation Data from Various APIs*, r package version 0.1.3, doi:10.5281/zenodo.400259.

Hurrel (2003), "NAO Index Data provided by the Climate Analysis Section," Accessed on 04.03.2020.

Huser, R. and Wadsworth, J. L. (2019), "Modeling spatial processes with unknown extremal dependence class," *J. Amer. Statist. Assoc.*, 114, 434–444.

Jonzén, N., Lindén, A., Ergon, T., Knudsen, E., Vik, J. O., Rubolini, D., Piacentini, D., Brinch, C., Spina, F., Karlsson, L., Stervander, M., Andersson, A., Waldenström, J., Lehikoinen, A., Edvardsen, E., Solvang, R., and Stenseth, N. C. (2006), "Rapid Advance of Spring Arrival Dates in Long-Distance Migratory Birds," *Science (New York, N.Y.)*, 312, 1959–61.

Koleček, J., Adamík, P., and Reif, J. (2020), "Shifts in migration phenology under climate change: temperature vs. abundance effects in birds," *Climatic Change*, 159, 177–194.

Kullberg, C., Fransson, T., Hedlund, J., Jonzén, N., Langvall, O., Nilsson, J., and Bolmgren, K. (2015), "Change in spring arrival of migratory birds under an era of climate change, Swedish data from the last 140 years," *Ambio*, 44 Suppl 1, S69–S77.

Lincoln, F. (1935), *Migration of Birds, Circular 16*, U.S. Fish and Wildlife Service.

Lindén, A. (2011), "Using first arrival dates to infer bird migration phenology," *Boreal Environment Research*, 16, 49–60.

Mayor, S. J., Guralnick, R. P., Tingley, M. W., Otegui, J., Withey, J. C., Elmendorf, S. C., Andrew, M. E., Leyk, S., Pearse, I. S., and Schneider, D. C. (2017), "Increasing phenological asynchrony between spring green-up and arrival of migratory birds," *Scientific Reports*, 7, 1902.

Møller, A. P., Rubolini, D., and Lehikoinen, E. (2008), "Populations of migratory bird species that did not show a phenological response to climate change are declining," *Proceedings of the National Academy of Sciences*, 105, 16195–16200.

Palm, V., Leito, A., Truu, J., and Tomingas, O. (2009), "The spring timing of arrival of migratory birds: Dependence on climate variables and migration route," *Ornis Fennica*, 86.

Reich, B. J. and Shaby, B. A. (2012), "A hierarchical max-stable spatial model for extreme precipitation," *Ann. Appl. Stat.*, 6, 1430–1451.

— (2019), "A spatial Markov model for climate extremes," *J. Comput. Graph. Statist.*, 28, 117–126.

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., and Rafaj, P. (2011), "RCP 8.5—A scenario of comparatively high greenhouse gas emissions," *Climatic Change*, 109, 33.

Shaby, B. A. and Reich, B. J. (2012), "Bayesian spatial extreme value analysis to assess

the changing risk of concurrent high temperatures across large portions of European cropland," *Environmetrics*, 23, 638–648.

Tøttrup, A. P., Rainio, K., Coppack, T., Lehikoinen, E., Rahbek, C., and Thorup, K. (2010), "Local Temperature Fine-Tunes the Timing of Spring Migration in Birds," *Integrative and Comparative Biology*, 50, 293–304.

U.S. Census Bureau (2010), "Population, Housing Units, Area, and Density: 2010—United States—County by State; and for Perto Rico," .

Wijeyakulasuriya, D. A., Hanks, E. M., Shaby, B. A., and Cross, P. C. (2019), "Extreme value-based methods for modeling elk yearly movements," *Journal of Agricultural, Biological and Environmental Statistics*, 24, 73–91.

Zaifman, J., Shan, D., Ay, A., and Jimenez, A. G. (2017), "Shifts in Bird Migration Timing in North American Long-Distance and Short-Distance Migrants are Associated with Climate Change," *International Journal of Zoology*, 2017, 6025646.

Zhang, H. B. H., Lawrimore, J., Menne, M., and Smith, T. M. (2010), "NOAA Global Surface Temperature Dataset (NOAAGlobalTemp)," .