Modeling Bird Migration by Disaggregating Population Level Observations

Miguel Fuentes, Benjamin Van Doren, Daniel Sheldon

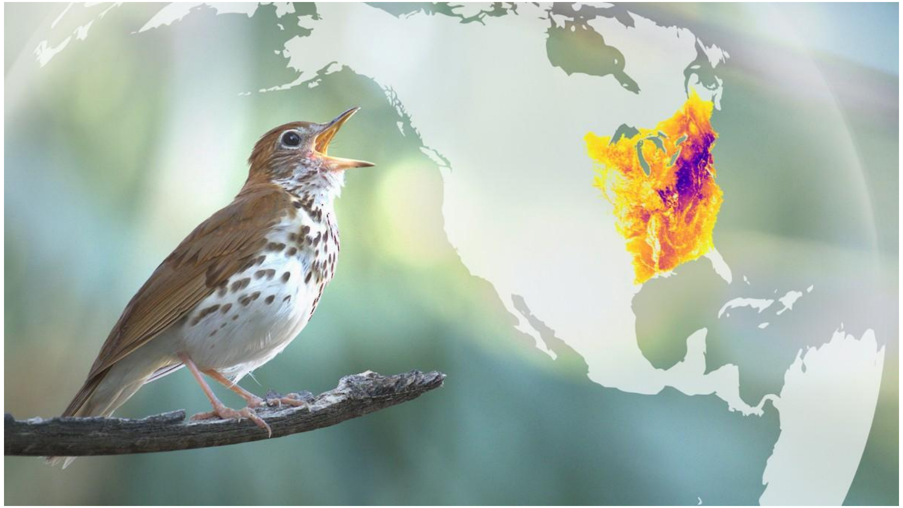
The Problem

Climate Change is Affecting Bird Migration

In recent decades, the migratory timing of migratory birds has shifted on a continental scale and climate change is an important factor in this change.

In order to fully understand these changes, we need effective methods for modeling bird migration.

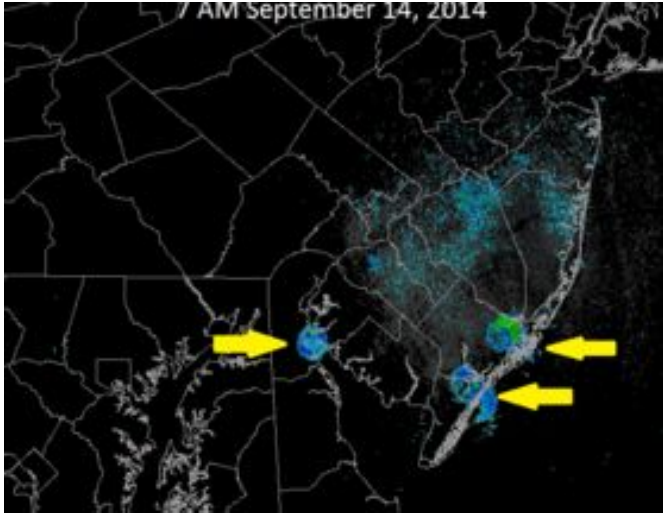
Some of the most common data sources used to build models are individual bird tracks and weather radar.



The Popular Data are Difficult to Use

Individual track data are very expensive and time consuming to use. Birds are often captured from one particular location so the data are usually not representative of the entire population.

Weather radar is unable to differentiate species of bird, so it is only able to detect broad trends in overall bird migration



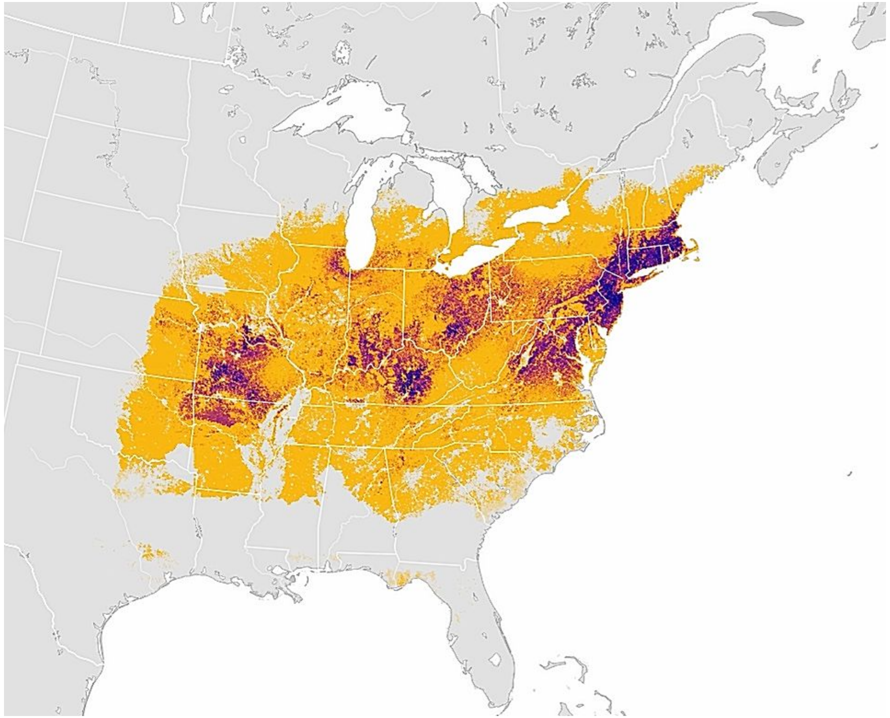
Our Solution

Use Population Level Data

The eBird project uses data from recreational birders to create population models for many bird species and makes those models widely available.

The population models come in the form of weekly abundance maps.

We process this data to create weekly "ground truth" distributions p* which our migration model will try to match.



The Model - Parameters

Our model is Markovian, and it assumes that the position of the bird in one time step only depends on the position in the previous timestep. We specify the model with these parameters.

$$\text{Initial Parameters:} \qquad Z^{(1)} \in \mathbb{R}^{n},$$

$$\text{Transition Parameters:} \qquad Z^{(t,t+1)} \in \mathbb{R}^{n \times n}.$$

Where n is the number of map cells and t ranges over all the weeks in the ground truth distribution.

The Model - Distributions

Those parameters specify an initial distribution with this form (sigma indicated the softmax function) :

$$p_Z(X_1 = i) = \sigma(Z^{(1)})_i$$

And transition distributions of this form:

$$p_Z(X_{t+1} = j | X_t = i) = \sigma(Z_i^{(t,t+1)})_j$$

The Model - Optimization

To learn the parameters we minimize a loss function with 3 parts

$$\mathcal{L}(Z) = \theta_s \mathcal{S} + \theta_b \mathcal{B} - \theta_h \mathcal{H}$$

- S - The mean squared error between the model distribution and ground truth
- B - The expected distance a bird travels in a week under the model
- H - The joint entropy of the model's learned distribution

Results

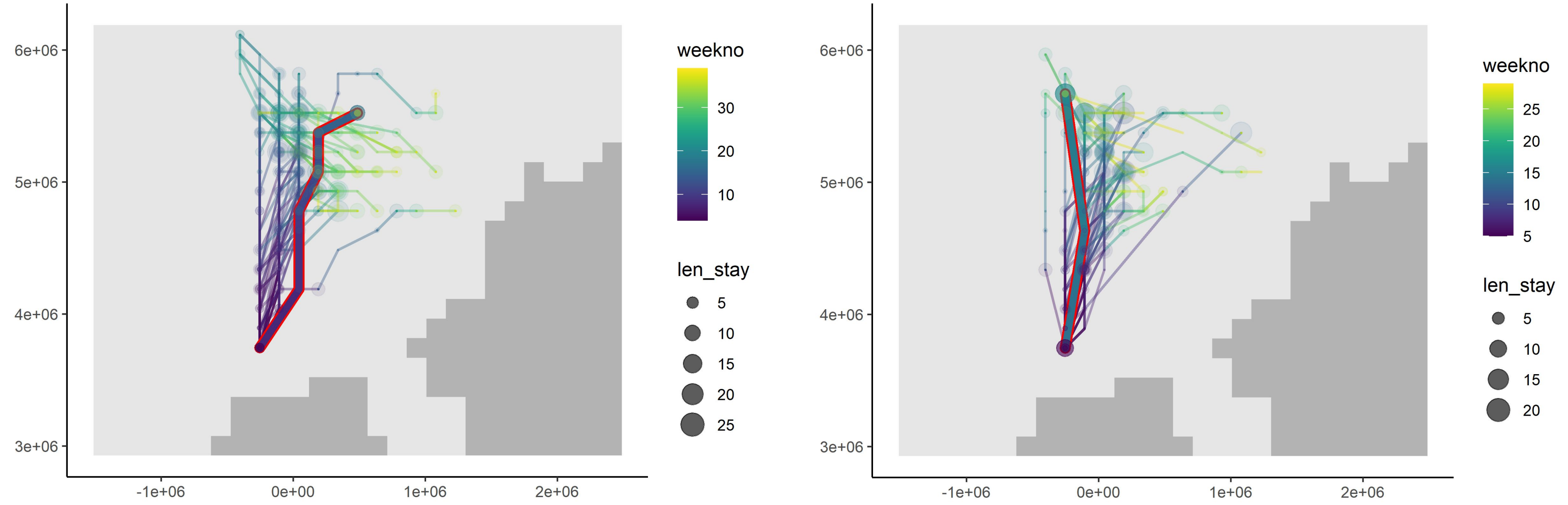
The American Woodcock

We applied our model to the abundance data for the American Woodcock. We chose this species because Moore et al. recently released individual American Woodcock tracks via the Movebank data repository (Kranstauber et al).

The availability of individual tracks simplifies the process of evaluating the model. We tuned the weights and chose the ones which assigned the highest log-likelihood to the real tracks.

Comparison To Real Tracks

Here, real tracks (outlined in red) can be seen next to several synthetic tracks which begin in the same location.



Discussion

Discussion

The model does a good job of positional forecasting and the real trajectories seem to lie within the “spread” of the samples. This shows that the model is not learning trajectories which are too restrictive.

There are some ways in which the model's trajectories differ from real ones. The model's trajectories move more often than true trajectories Also, some of the year long sampled trajectories start and end in different locations, true trajectories end very close to where they start.

By using the learned distribution from our model it may be possible to infer behavioral changes due to climate change with the use of fewer real tracks.

Questions?

Contact Us

Miguel Fuentes
(mmfuentes@umass.edu)

Benjamin Van Doren
(vandoren@cornell.edu)

Daniel Sheldon
(sheldon@cs.umass.edu)