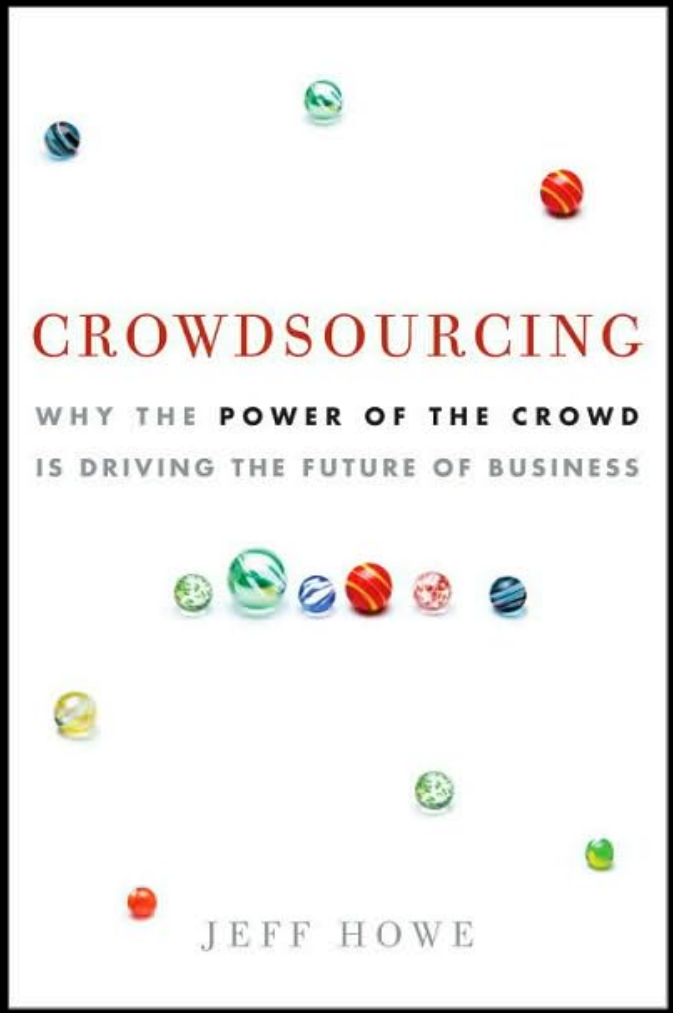
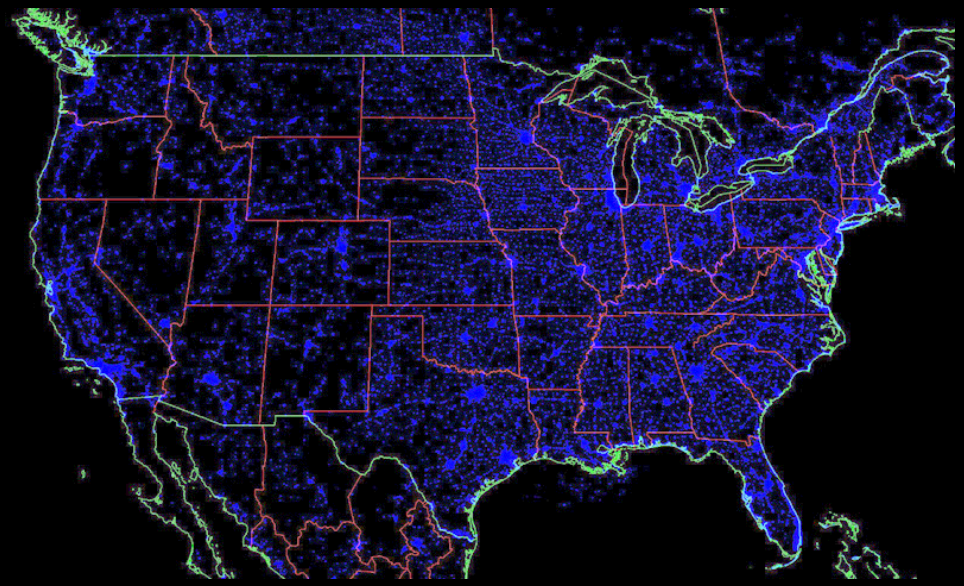

TheCornellLab of Ornithology

Ornithology has always engaged amateur participation

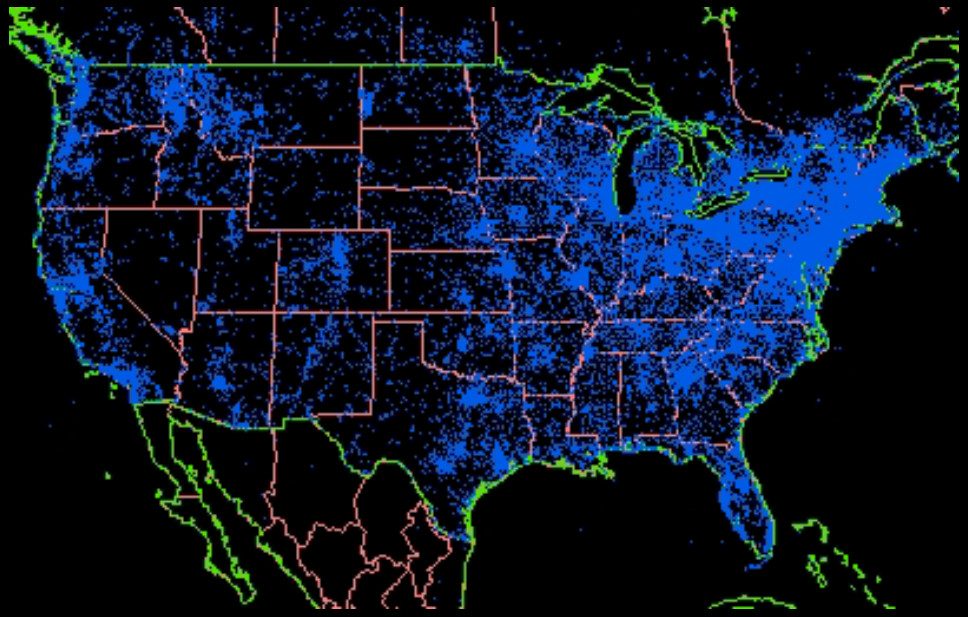
> If stock photography is the first industry to be transformed by crowdsourcing, then ornithology is the first academic discipline to undergo the same process.



Lights at Night



Citizen Science Submissions



Puerto Rico eBird

SOPI SOCIEDAD ORNITOLÓGICA PUERTORRIQUEÑA INC.

Inicio | Acerca de eBird | Envía tus observaciones | Ver y explorar datos | Mi eBird

English Français

Bienvenidos a eBird Puerto Rico

¡Donde tus observaciones hacen la diferencia! Registrarse como nuevo usuario

Noticias sobre observación de aves

Aves en larga jornada

En este mes hay que estar atento. El mes de octubre es quizás el mes donde más observaciones "raras" se reportan en Puerto Rico. Estas observaciones raras corresponden a especies que no están presentes en la isla todo el año pues son especies migrantes.

Juvenil de Gallito Amarillo en la Hacienda la Esperanza

En la mañana del pasado 15 de agosto de 2007, en la Reserva Natural Hacienda La Esperanza, al norte del municipio de Manatí, una compañera de trabajo me comentó que había observado unos Gallitos Negros (*Laterallus jamaicensis*) caminando buscando comida entre la vegetación del humedal. Para mi sorpresa no fueron Gallitos Negros lo que encontré sino un juvenil de Gallito Amarillo (*Porzana flaviventer*).

cambios en patrones de migración. Una de estas especies lo es el Vireo Ojiblanco (*Vireo griseus*).

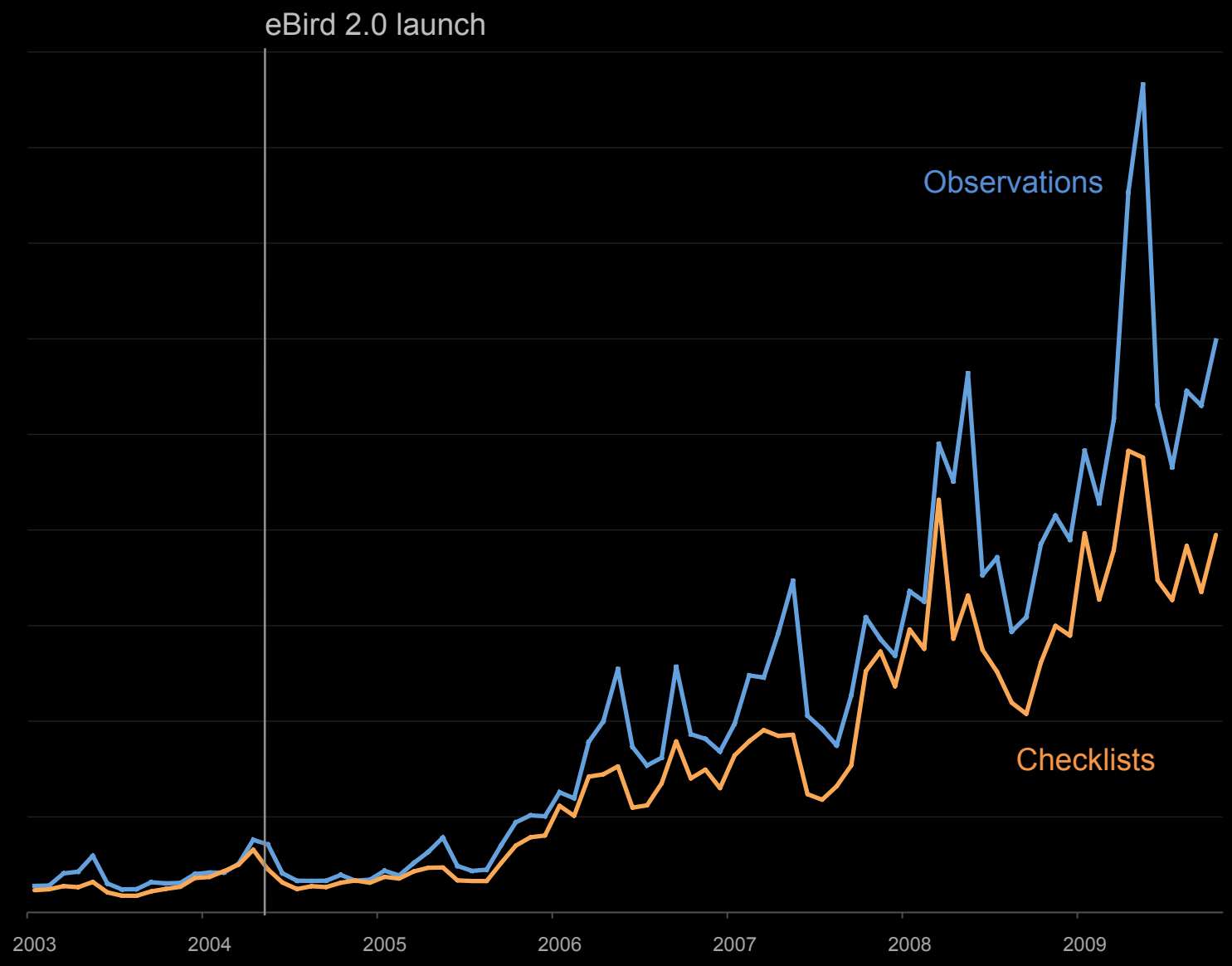
NSF

Wolf Creek Charitable Foundation

NSF

Wolf Creek Charitable Foundation

Growth in eBird Observations and Checklists



eBird

Home | About eBird | Submit Observations | View and Explore Data | My eBird

Hello Steve Kelling (stk091655) | Preferences | Sign Out

Translate to: English | Español | Français

« back

Step: [1] [2] **[3]** [4]

Step 3: What did you see or hear?

aa NY TOM BCF

Snow Goose		Cackling Goose		Canada Goose	
Wood Duck		Gadwall		American Wigeon	
American Black Duck		American Black Duck x Mallard (hybrid)		Mallard	
Blue-winged Teal		Northern Shoveler		Northern Pintail	
Green-winged Teal		Green-winged Teal (American)		Canvasback	
Redhead		Ring-necked Duck		Greater Scaup	
Lesser Scaup		Surf Scoter		White-winged Scoter	
Black Scoter		Long-tailed Duck		Bufflehead	
Common Goldeneye		Hooded Merganser		Common Merganser	
Red-breasted Merganser		Ruddy Duck			

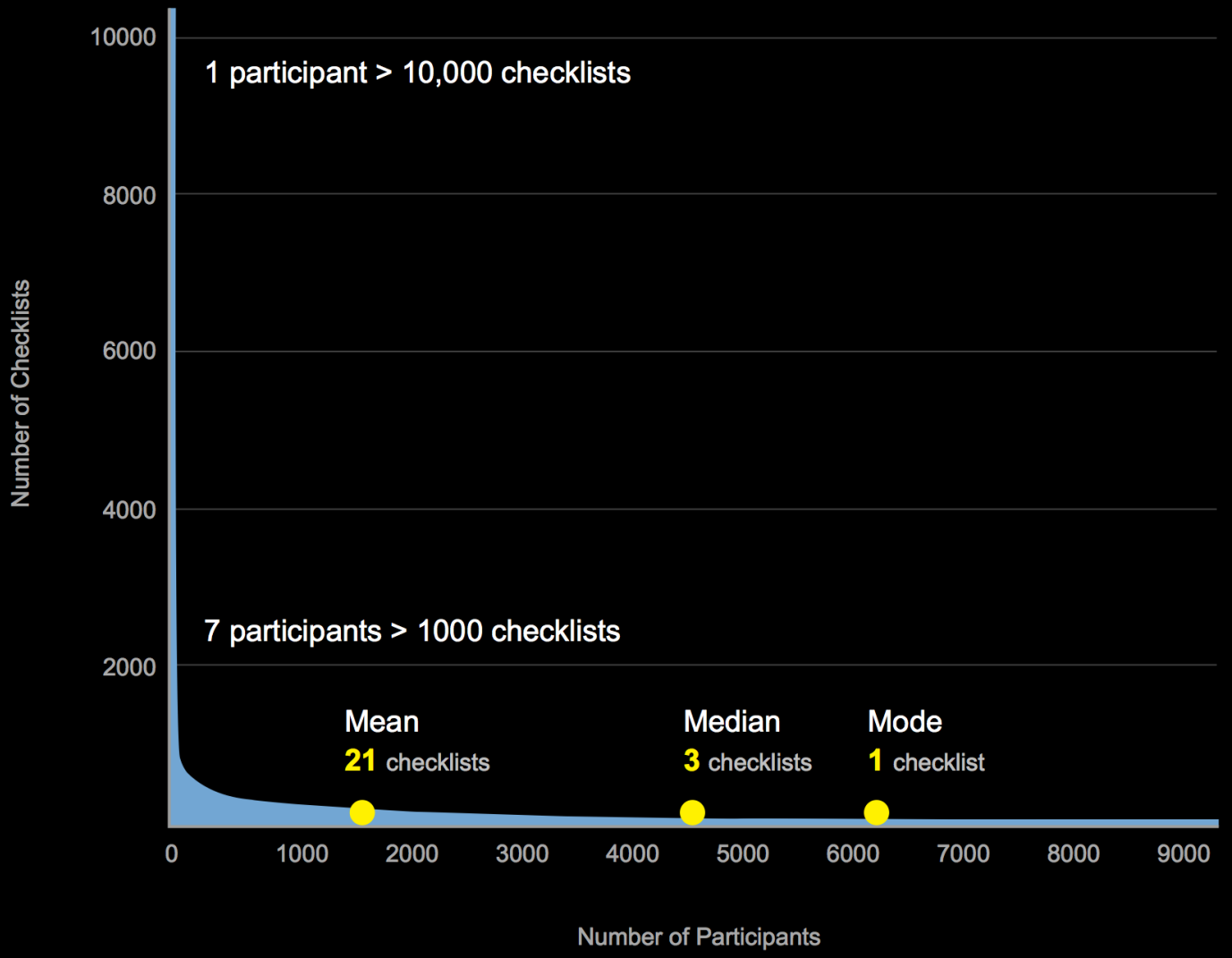
Grouse, Quail, and Allies Top

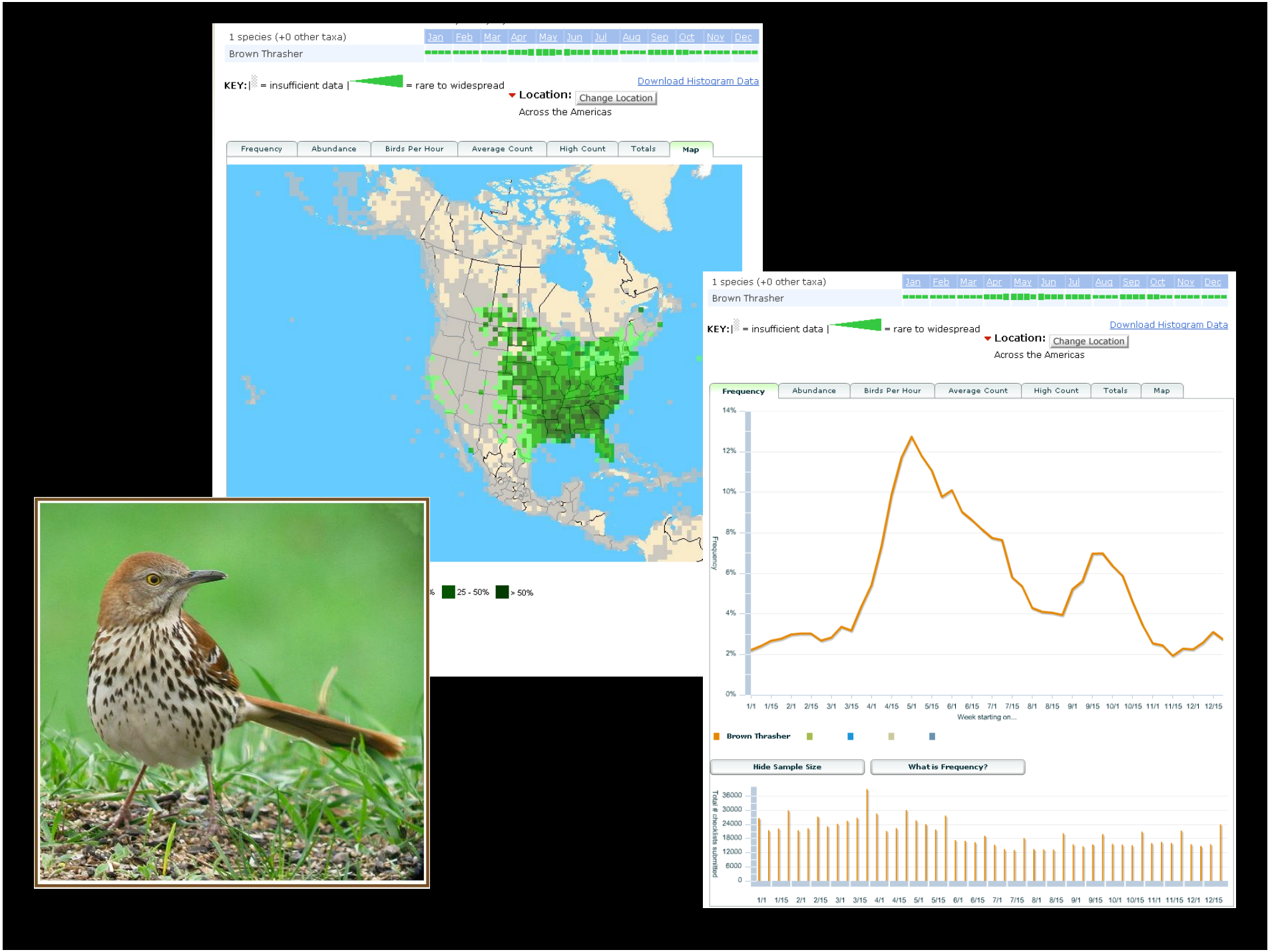
Ring-necked Pheasant		Ruffed Grouse		Wild Turkey	

Loons and Grebes Top

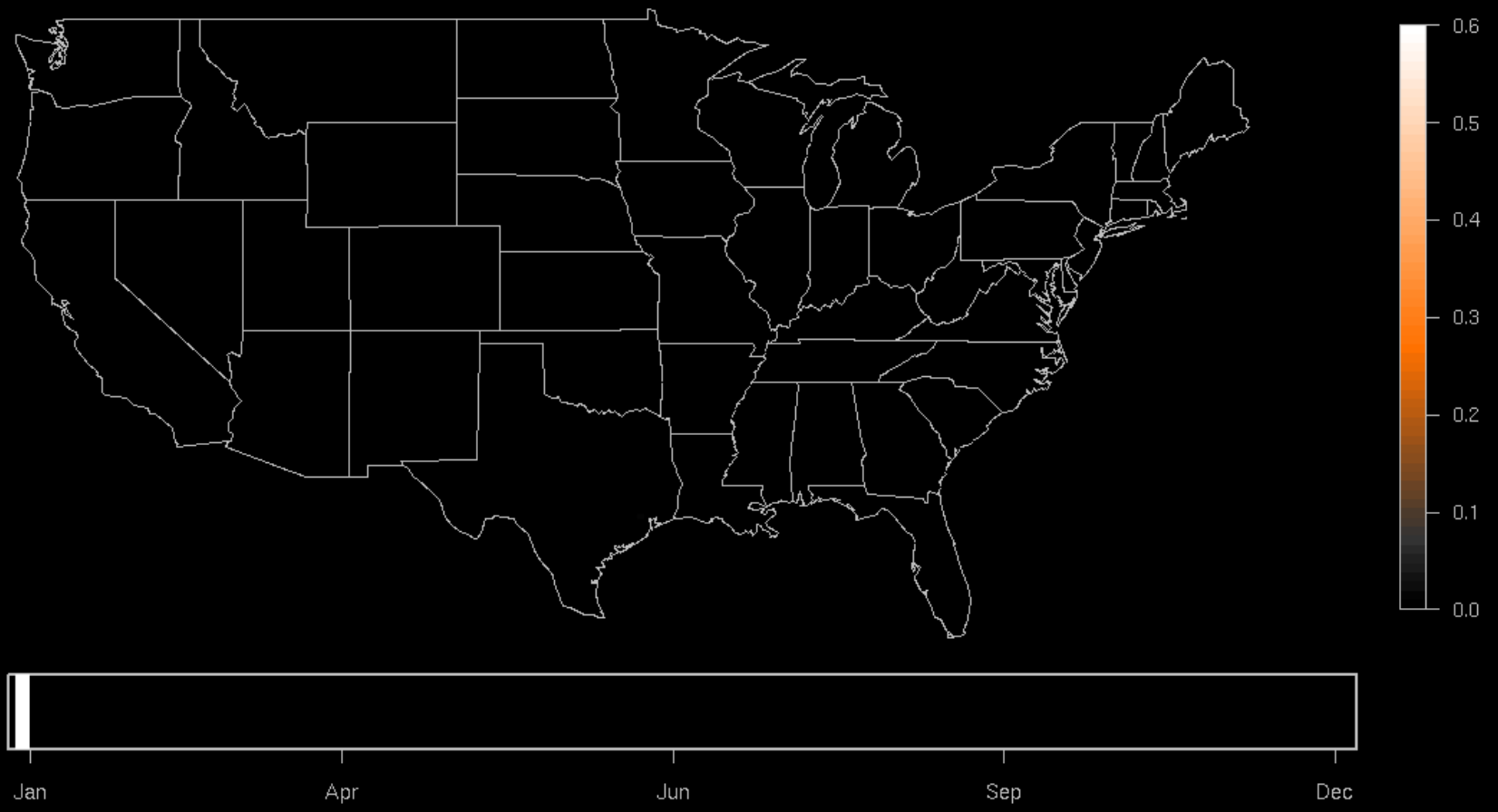
Red-throated Loon		Common Loon		Pied-billed Grebe	
Horned Grebe		Red-necked Grebe			

2008 Checklist Submissions Follow a Power Law Distribution

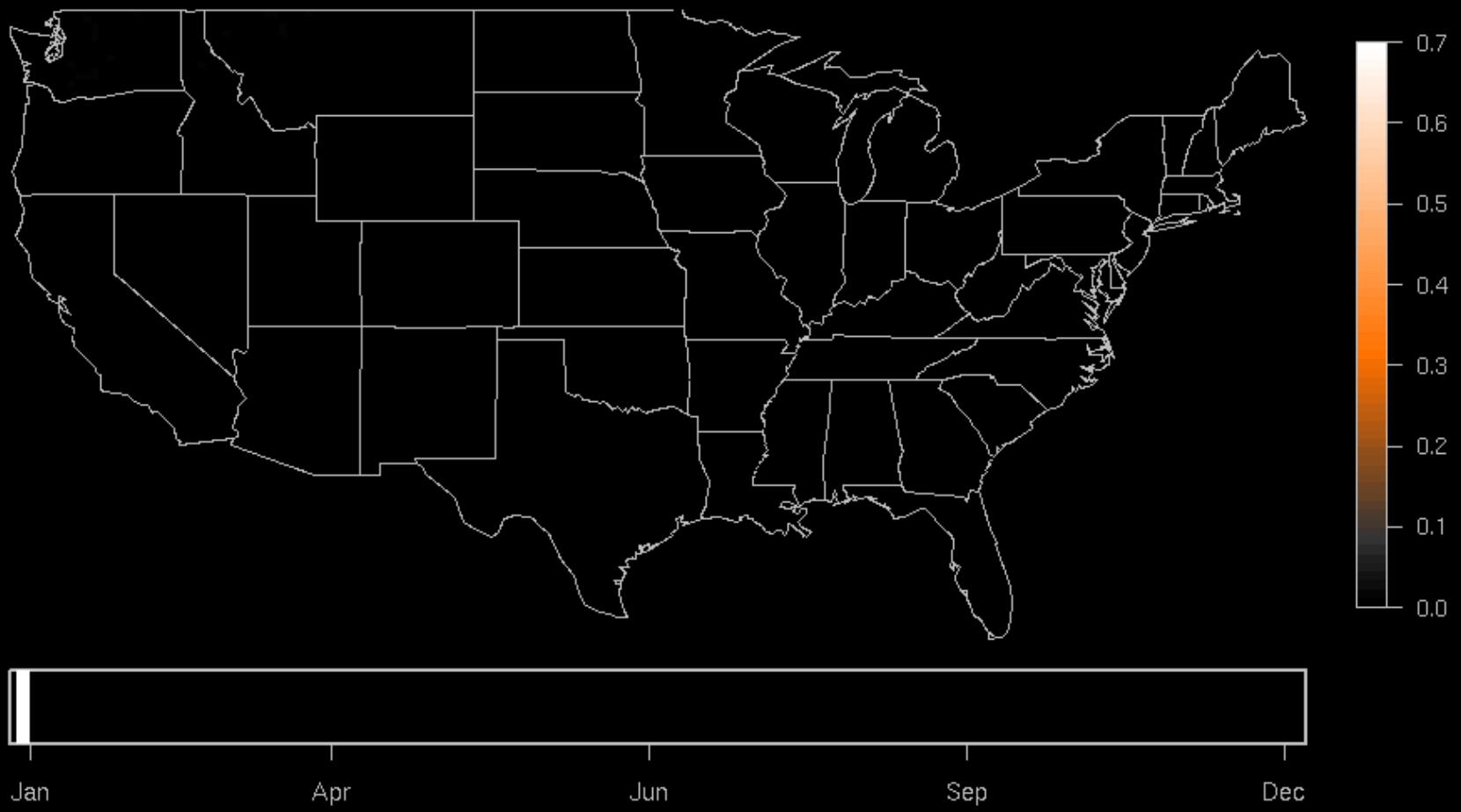



1 species (+0 other taxa)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Brown Thrasher
KEY: = insufficient data | = rare to widespread
Download Histogram Data
Location: Change Location
Across the Americas
Frequency Abundance Birds Per Hour Average Count High Count Totals Map
25 - 50% > 50%
1 species (+0 other taxa)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Brown Thrasher
KEY: = insufficient data | = rare to widespread
Download Histogram Data
Location: Change Location
Across the Americas
Frequency Abundance Birds Per Hour Average Count High Count Totals Map
14%
12%
10%
8%
6%
4%
2%
0%
Frequency
1/1 1/15 2/1 2/15 3/1 3/15 4/1 4/15 5/1 5/15 6/1 6/15 7/1 7/15 8/1 8/15 9/1 9/15 10/1 10/15 11/1 11/15 12/1 12/15
Week starting on...
Brown Thrasher
Hide Sample Size
What is Frequency?
36000
30000
24000
18000
12000
6000
0
Total # checklists submitted
1/1 1/15 2/1 2/15 3/1 3/15 4/1 4/15 5/1 5/15 6/1 6/15 7/1 7/15 8/1 8/15 9/1 9/15 10/1 10/15 11/1 11/15 12/1 12/15

Chimney Swift | Full Year



Animation - Western Wood Peewee



EVA Workgroup Overview:

Scientific Exploration, Visualization, and Analysis

DataONE The Observation Network for Earth

Goal:

- To ensure the preservation and access to multi-scale, multi-discipline, and multi-national data.
- Provide an open, persistent, robust and secure access to well-described and easily discovered data.

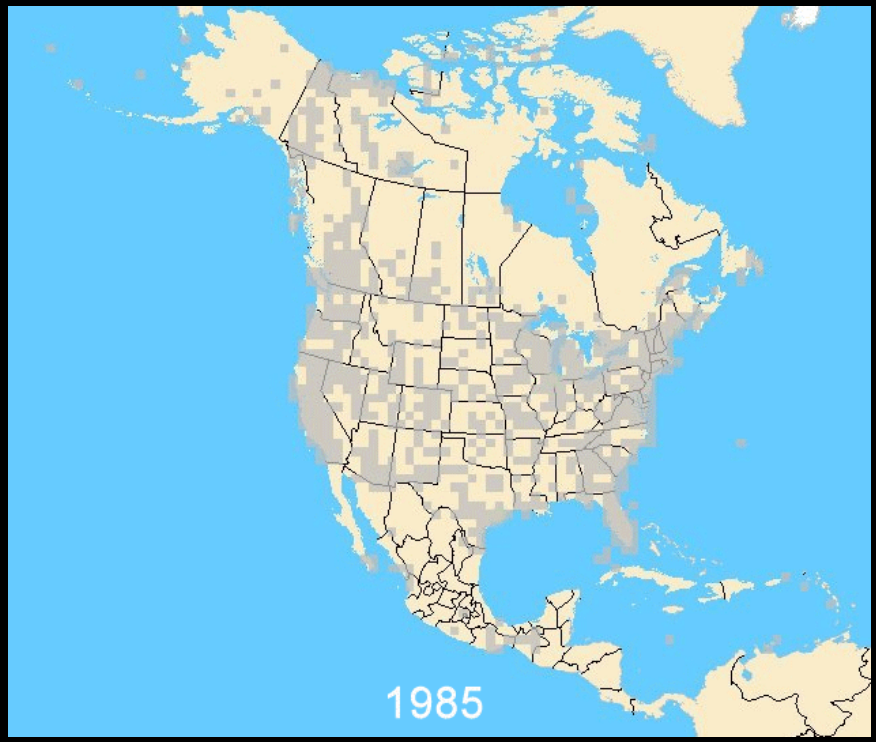
Transcend domain boundaries

- Biological data: from the genome to the ecosystem;
- Environmental data: from atmospheric, ecological, hydrological, and oceanographic sources;

➔ *Not Just BIG DATA*

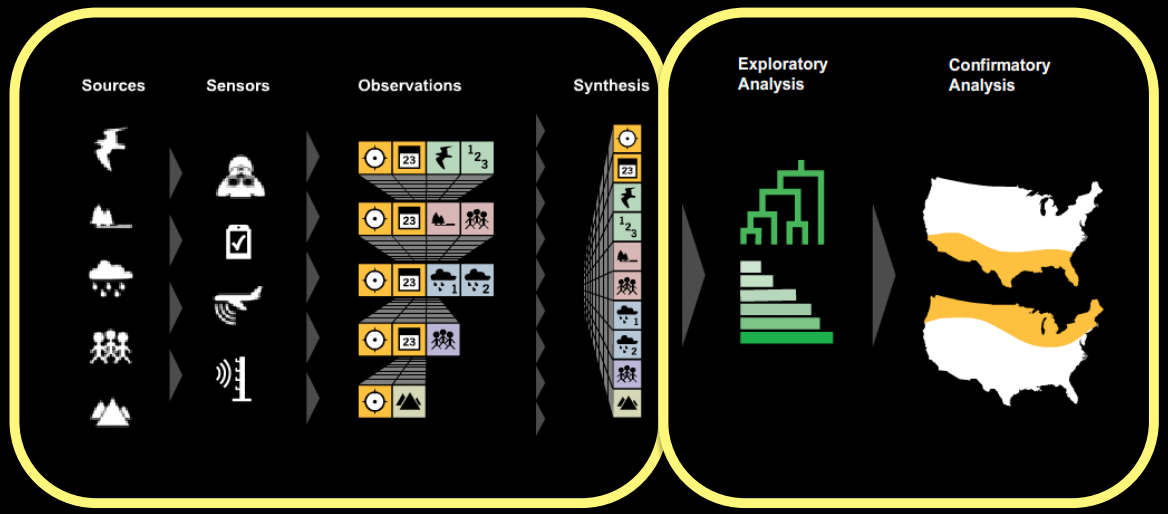
Migration Phenology

Invasive Species Monitoring



The Data Intensive Science

Data Collection Organization Validation Preservation, & Access



Data Exploration, Visualization, And Analysis

The full data “life cycle” from data curation to scientific learning

Engage scientists, land-managers, policy makers, students, educators, and the public

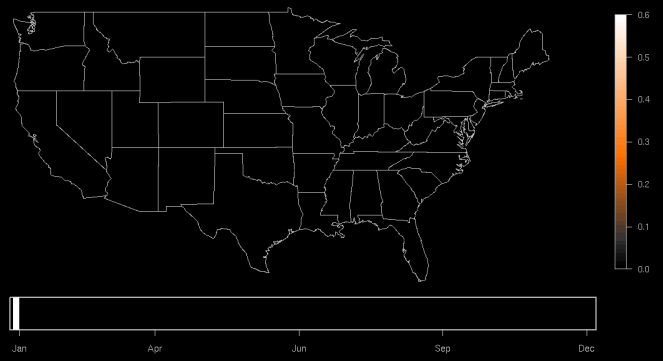
EVA Workgroup Goals

Identify and develop exemplar problems for which we will create complete workflows

With exemplar problems

- Conduct publishable scientific research
- Test and guide development of DataONE cyberinfrastructure
- Engage communities (outreach & broader impacts)

EVA Exemplars



Spatiotemporal analyses of

1. bird migrations and
2. the expansion of invasive species

We need to understand how species distributions change and evolve through time and what are the drivers associated with these changes

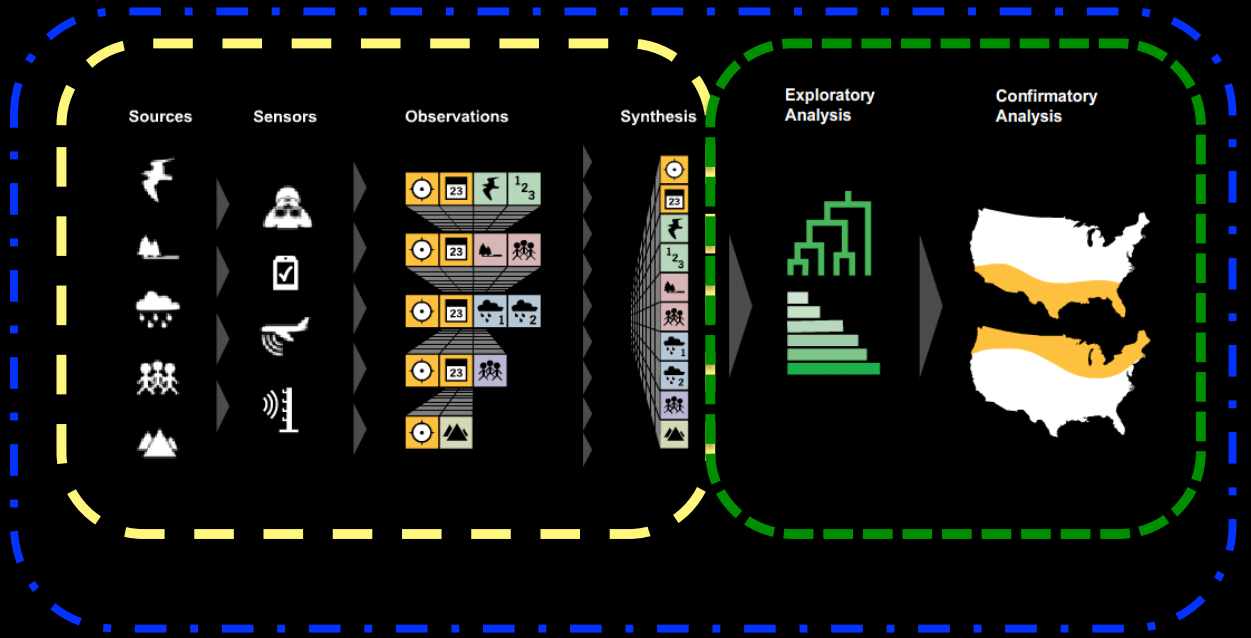
- Timing, Direction, Speed
- Forecast and control distributional changes
- Inform conservation & land use policy

EVA Participants

- Participants:
- Steve Kelling (co-chair), Cornell Lab of Ornithology
- Bob Cook (co-chair), Oak Ridge National Lab
- Suresh SanthanaVannan, Oak Ridge National Lab
- Patrick O'Leary (U of Idaho)
- Jeff Morisette, USGS
- Claudio Silva, University of Utah
- Tom Dietterich, Oregon State
- Daniel Fink, Cornell Lab of Ornithology
- Kevin Webb, Cornell Lab of Ornithology
- Bill Michener, University of New Mexico
- Alyssa Rosemartin National Phenology Network
- Damian Gessler iPlant

The 3 EVA Subgroups

1 Data Subgroup:
Collection
Organization
Validation



2 Model Subgroup:
Exploratory & Confirmatory Models & Visualizations

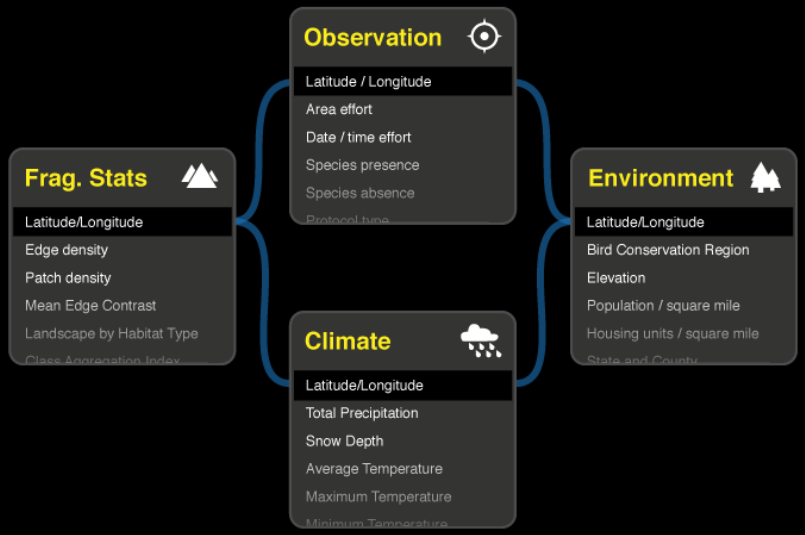
3 Workflow Subgroup:
Workflow tools to facilitate Data Intensive Analysis

Subgroup Tasks: 6-9 Months

- Data Subgroup
  - Acquire & Process additional data layers (Greening, PRISM, forest biomass, etc)
- Modeling Subgroup
  - Develop inferential methods to describe how spatiotemporal drivers (e.g. greening) are associated with change in distributions
  - Develop models to facilitate study of ST processes
- Workflow Subgroup
  - Use VisTrails to manage & develop workflows with SAHM
  - Use VisTrails to manage & develop workflows with STEM
  -

The challenges:

Data synthesis

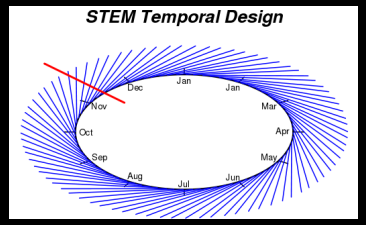


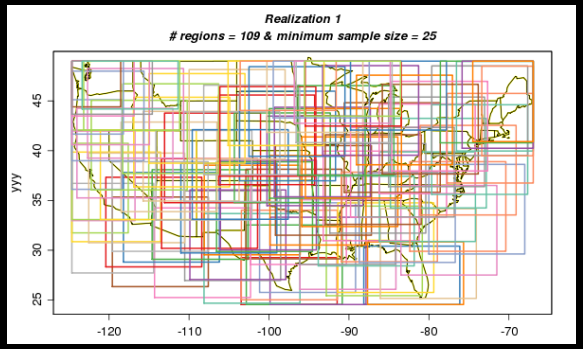
Suresh SanthanaVannan
Kevin Webb

The challenges:

Modeling

$$\frac{1}{n(s,t)}\sum_{i=1}^{m} f_i(X,s,t)I(s,t \in \theta_i)$$



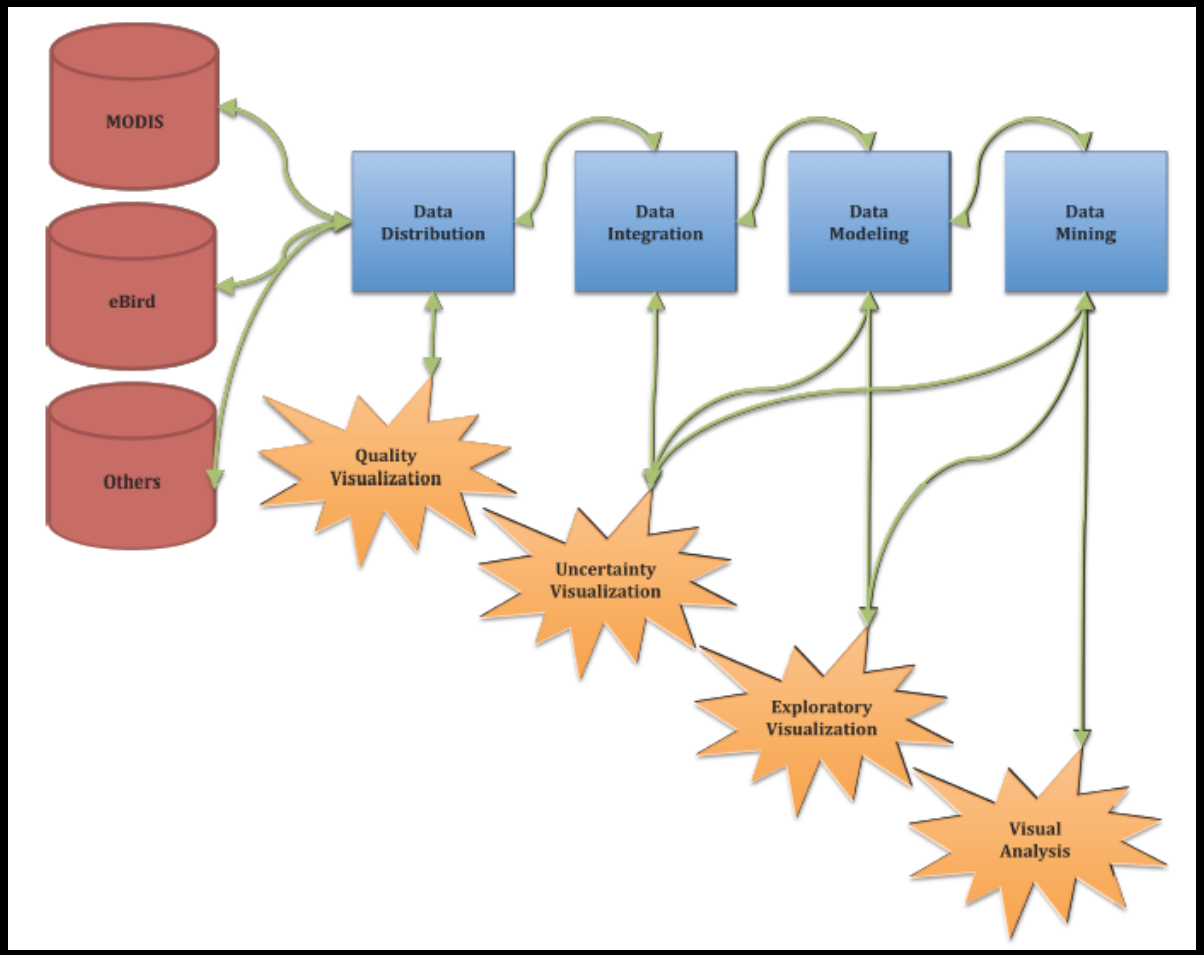


Daniel Fink
Tom Dietterich
Jeff Morisette
Damian Gessler

model latent variables
(i.e., variables which are not directly observed but inferred
from model outputs based on variables that were observed).

The challenges:

Workflow



The products:

Reference Dataset

Vis Trails Workflow

Migration Atlas of North America (Web)

Data Intensive Science Process Paper (Peer review)

Invasive Species Modeling (Peer review)

The Timeline:

	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Data organization:	x	x	X	x				X	
Data modeling:	x	x	X	x	x	x	x	X	
Workflow:			x	x	x	x	x	X	
Visualization:					x	x	x	X	x
Migration Atlas:									x
Process Paper:									x
Invasive Species Paper:									x

Modeling Species Distribution Dynamics with Spatiotemporal Exploratory Models: *Inter-annual Bird Migrations*

Daniel Fink
df36@cornell.edu
Steve Kelling, Kevin Webb, and
the AKN Research Team

The Cornell Lab of Ornithology

NSF

LEON LEVY FOUNDATION
The Legacy Of Leon Levy

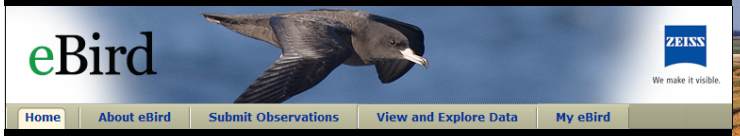
Goal: Modeling Species Distribution Dynamics

We need to understand how species distributions change and evolve through time

- Phenology – annual cycle dynamics
- Conservation - migration corridors & stopover sites
- Spread of invasive species, diseases, dispersal, etc.
- Anthropogenic changes to environment
- ➔ *Large spatial & temporal extent*
- ➔ *Fine spatial & temporal scale*

Exploratory analysis:

1. Little a priori knowledge
2. Facilitate Discovery & Description



eBird Data

- Traveling Count
- Lat - Lon
- Year (2004 – 2008)
- Julian Date (1-365)

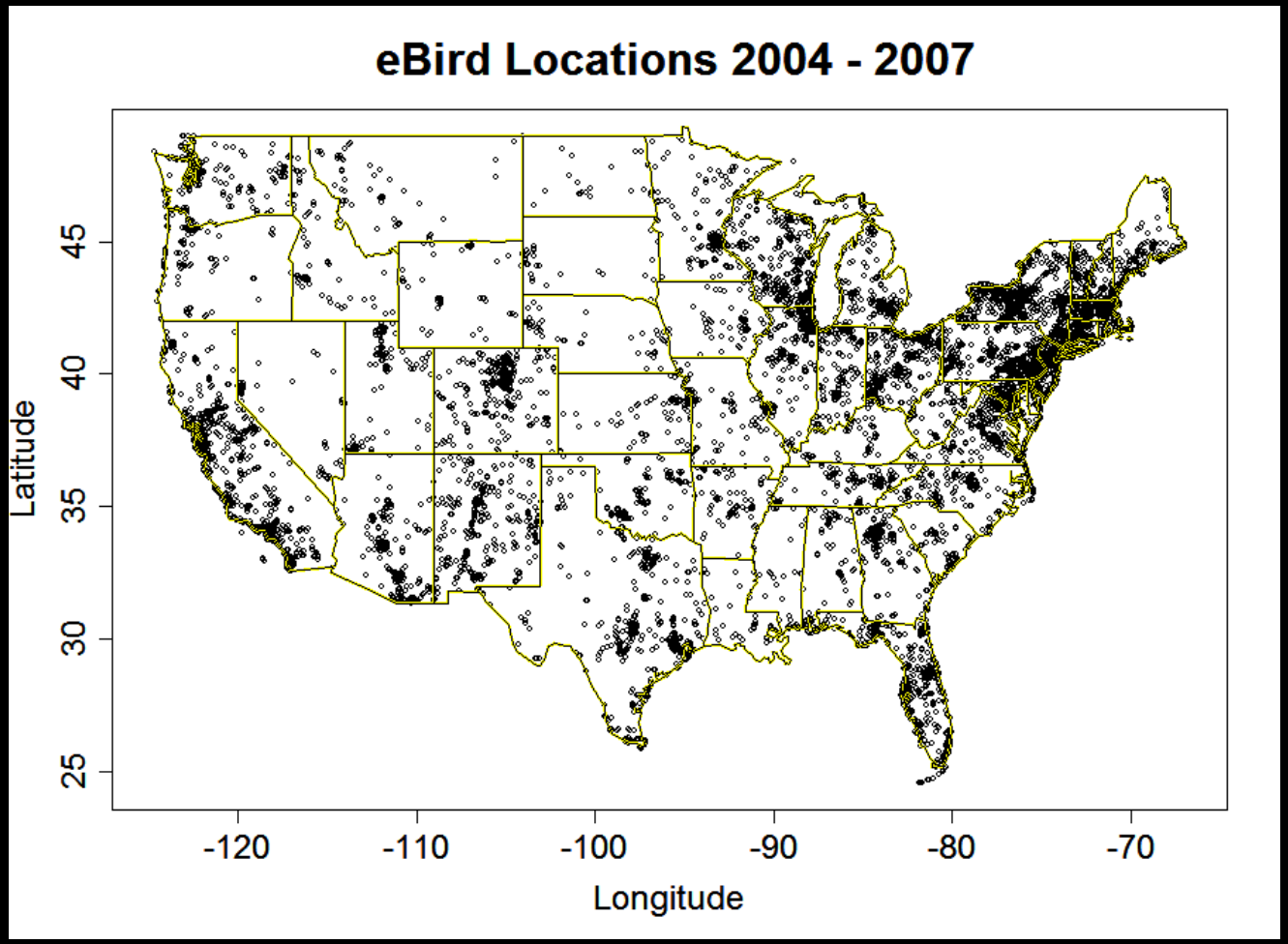
- Observation Effort
- Observation Time
- Number of Observers

➔ Presence/absence
  <= 8 km long
  <= 3 hours
  ~150,000 observations
  ~30,000 locations

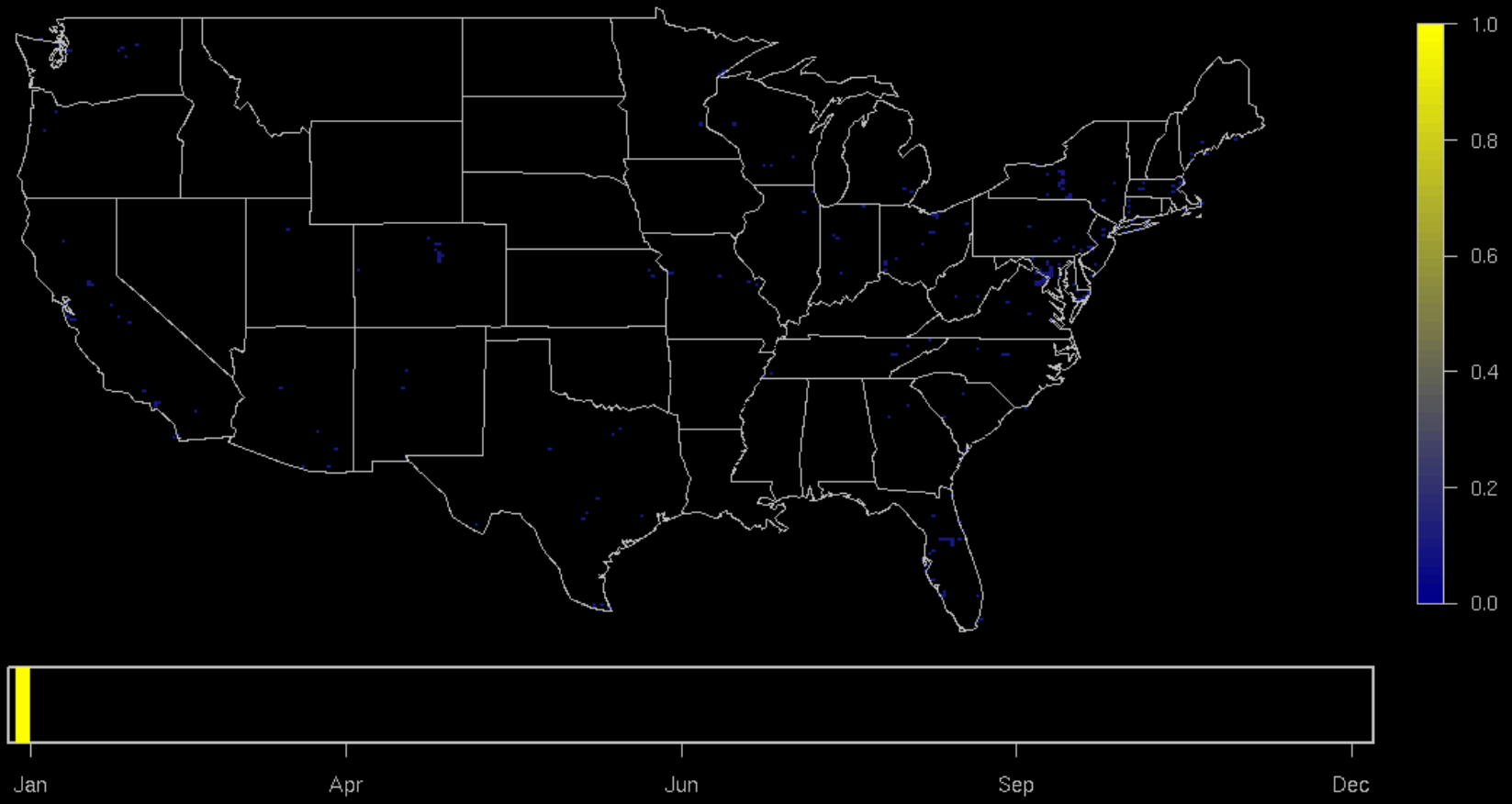
National Landcover Data (NLCD)

- 2001
- 16 Classes
- 1.5 km pixel
- Spatial Composition (% coverage)
- Spatial Configuration (Fragstats)

- Elevation
- Climate (30 yr average)
- Housing Density


eBird Locations 2004 - 2007
Latitude
25
30
35
40
45
-120
-110
-100
-90
-80
-70
Longitude

eBird data | 2008 traveling count protocol



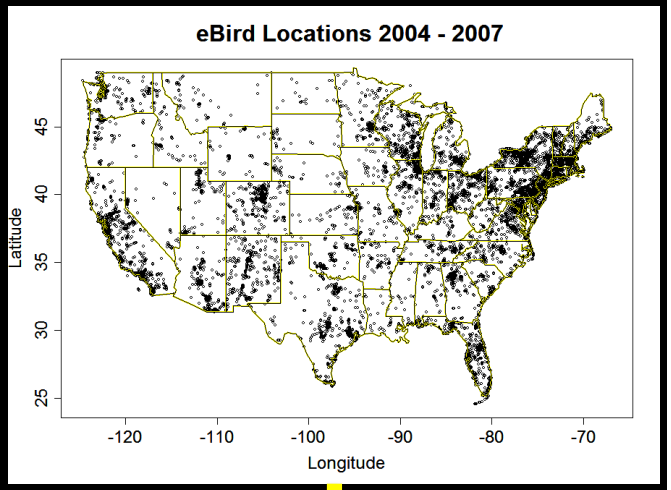
Species Distribution Modeling

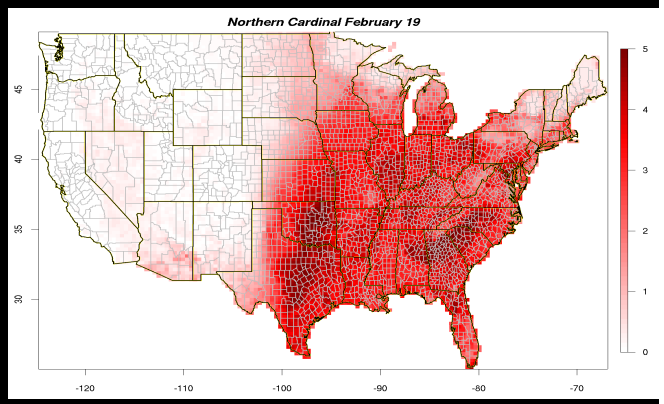
Goal: broad-scale & fine resolution

- Observational data are sparsely distributed in space and time
- Interpolation is essential

Modeling also buys us

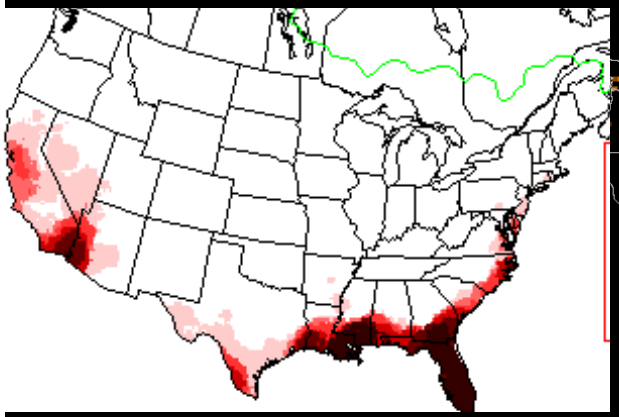
- Control bias
- Quantify uncertainty
- Framework for predictive experimentation



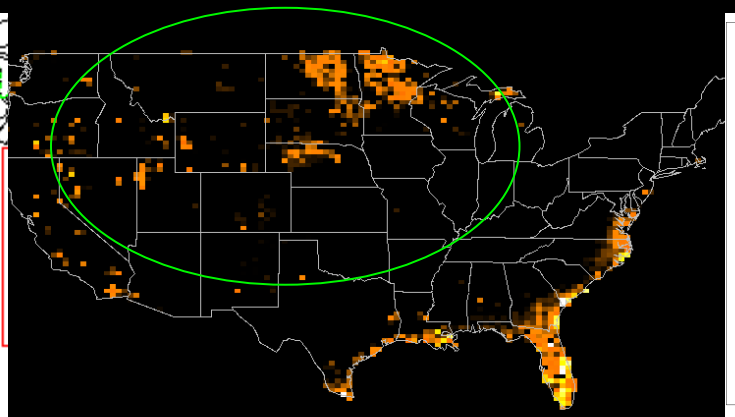


(Tachycineta bicolor)

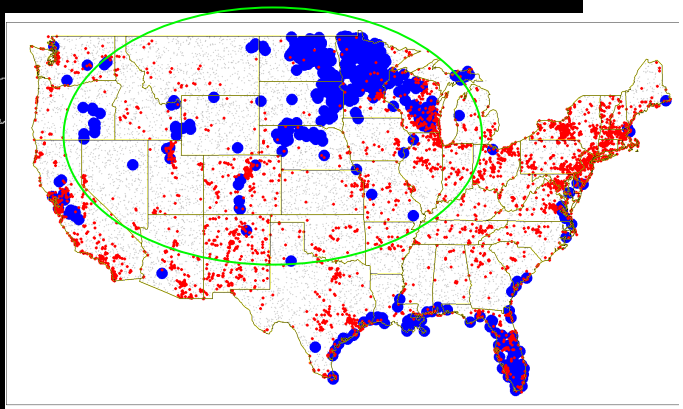
Tree Swallow Winter Distribution Analysis Bagged Decision Tree



Christmas Bird Count



eBird BDT



Wetland Coverage > 5%

- *Nonparametric model with global support may aggregate data in ways that are ecologically impossible*
- *SDM shares habitat information across regions and seasons where Tree Swallows do not coexist.*

The Multi-scale Challenge

Goal: Analysis at broad-scale with fine resolution

Challenge: spatiotemporal patterning at multiple scales

- Local-scale Homogeneity
  - Fine-scale spatial and temporal resource patterns
  - Local-scale dispersal
- Large-scale Heterogeneity
  - Regional & seasonal variation in species' habitat utilization
  - Source-sink dynamics & Allee effects
  - La Nina & North Atlantic Oscillation

SpatioTemporal Exploratory Model (STEM)

Current nonparametric SDM's are very good for local-scale modeling by relating environmental predictors (X) to observed occurrences (y)

$$y = f(X)$$

SpatioTemporal Exploratory Model (STEM)

Current nonparametric SDM's are very good for local-scale modeling by relating environmental predictors (X) to observed occurrences (y)

$$y = f(X)$$

Multi-scale strategy: differentiate between local and global-scale ST structure.

1. Make explicit time (t) and location (s)

$$f(X, s, t)$$

SpatioTemporal Exploratory Model (STEM)

Current nonparametric SDM's are very good for local-scale modeling by relating environmental predictors (X) to observed occurrences (y)

$$y = f(X)$$

Multi-scale strategy: differentiate between local and global-scale ST structure.

1. Make explicit time (t) and location (s)
2. "Regionalize" by restricting support

Restricted Support Set (Q)

$$f(X,s,t)I(s,t \in \theta)$$

SpatioTemporal Exploratory Model (STEM)

Current nonparametric SDM's are very good for local-scale modeling by relating environmental predictors (X) to observed occurrences (y)

$$y = f(X)$$

Multi-scale strategy: differentiate between local and global-scale ST structure.

1. Make explicit time (t) and location (s)
2. "Regionalize" by restricting support
3. Predictions at time (t) and location (s) are made by averaging across a set of local models containing that time and location

Restricted Support Set (q)

$$f(X,s,t)I(s,t \in \theta)$$

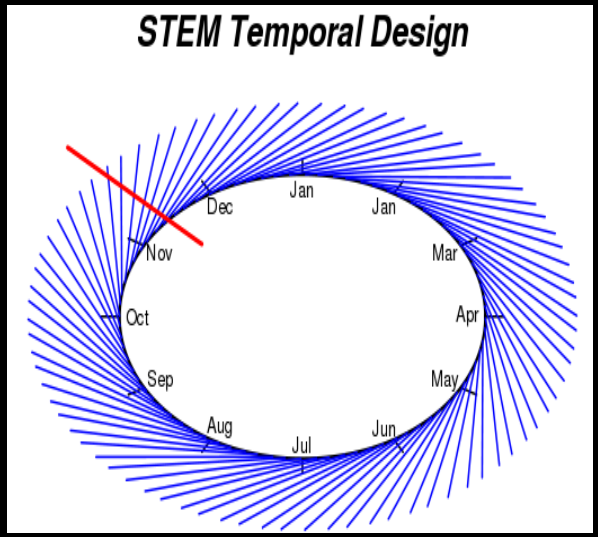
i^{th} ST explicit base model

Number of models supporting (s,t)

$$\frac{1}{n(s,t)} \sum_{i=1}^{m} f_i(X,s,t)I(s,t \in \theta_i)$$

Theta: The Spatio-Temporal Design

IDEA: ST Slice and dice sufficient overlap

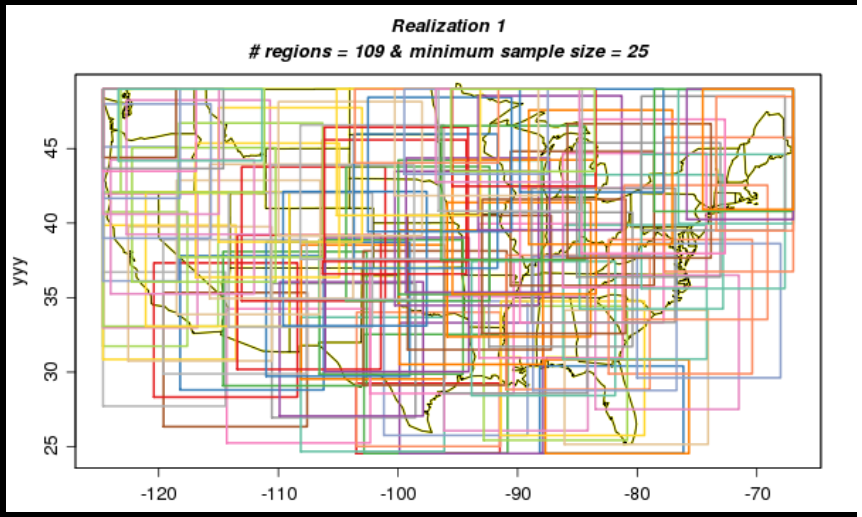


Temporal Design:

- 40 day windows
- 80 evenly spaced windows throughout year

Spatial Design

- For each time window
- Random Sample rectangles (constant size)
- With at least 25 unique locations.



The Spatio-Temporal Ensemble

$$F(X,s,t) = \frac{1}{n(s,t)} \sum_{i=1}^{m} f_i(X,s,t) I(s,t \in \theta_i)$$

Quantitative Intuition		**Ecological Intuition**
Statistical Experimental Design: Block over ST regimes to control variance	→	Local predictor-response learning No "long-range" learning
Bagging: resample block-level variance and average	→	"Local" averaging allows large-scale patterns emerge from local-scale

- *Add essential spatiotemporal structure to existing techniques*
- *Models a wide variety of dynamic processes automatically*

Base Models: Decision Trees

f(Habitat, Effort, LAT, LON, Year, Julian Date) = Predicted Count

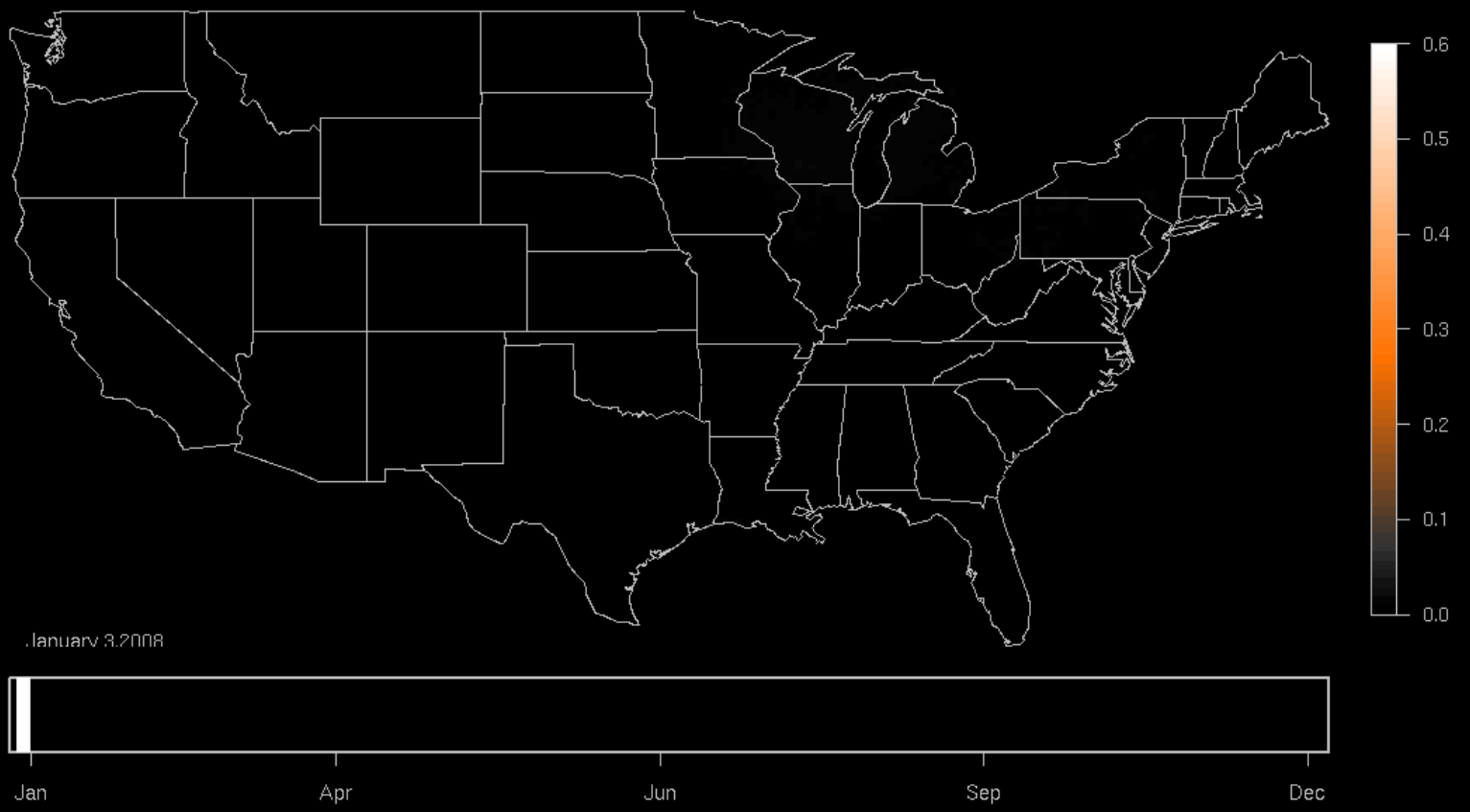
Strengths

- Automatic
- Good Predictive Performance
- Scale well to large data sets

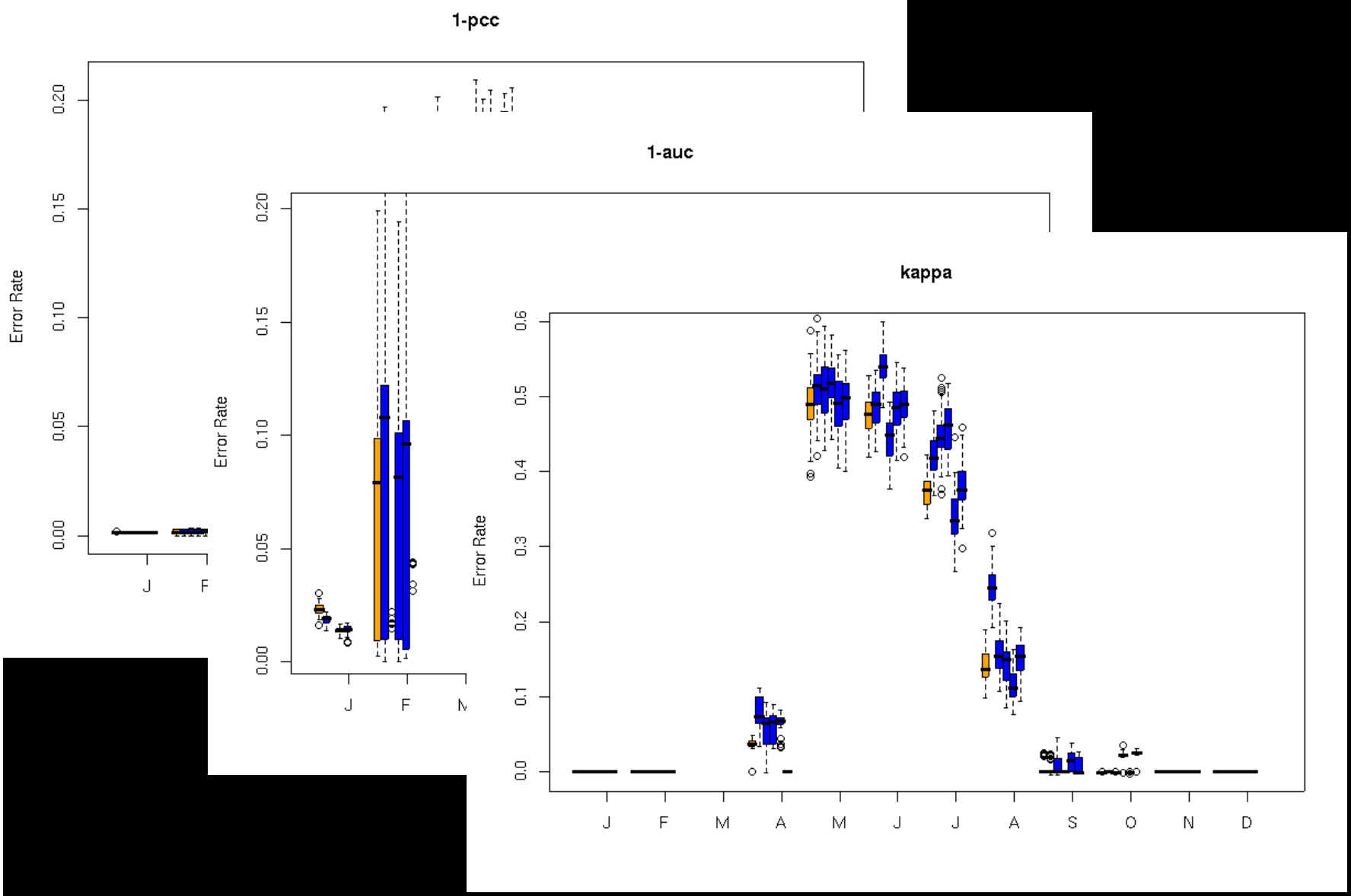
- Monotone transformations
- Nonlinear predictor effects
- Automatic interaction modeling
- Missing covariates

➔Predictor Selection is the main analyst decision

Indigo Bunting | Full Year

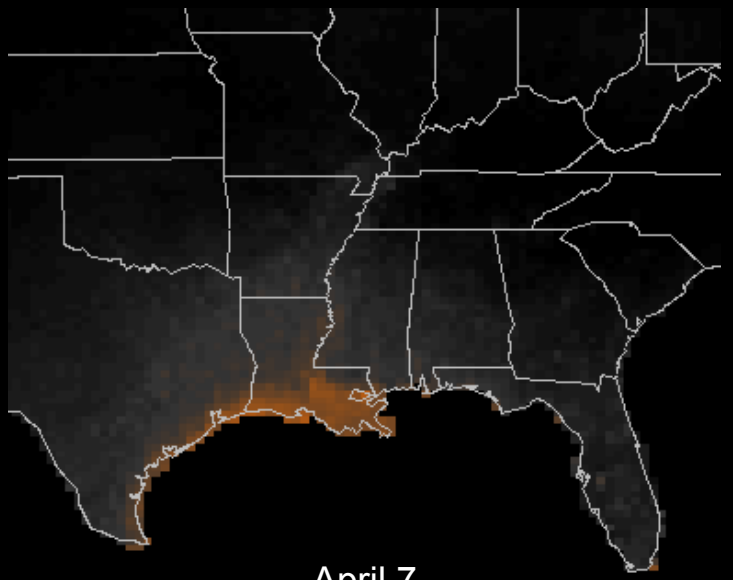


Indigo Bunting | Monthly Predictive Performance

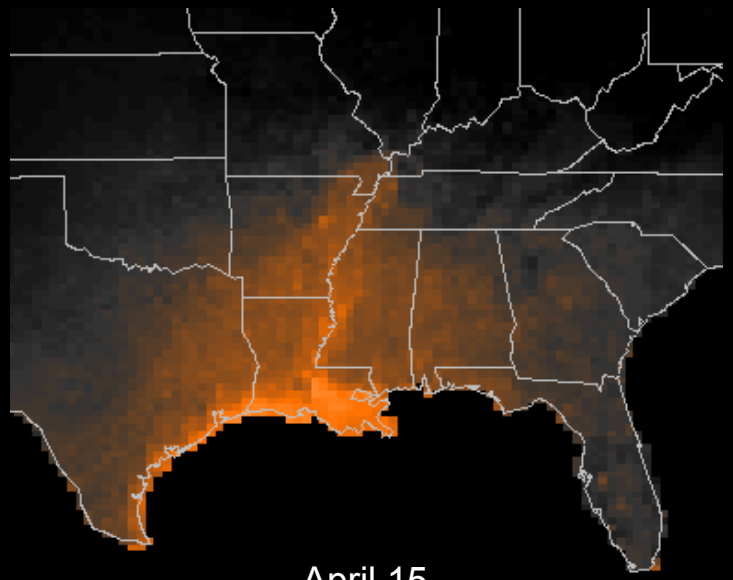


Indigo Bunting

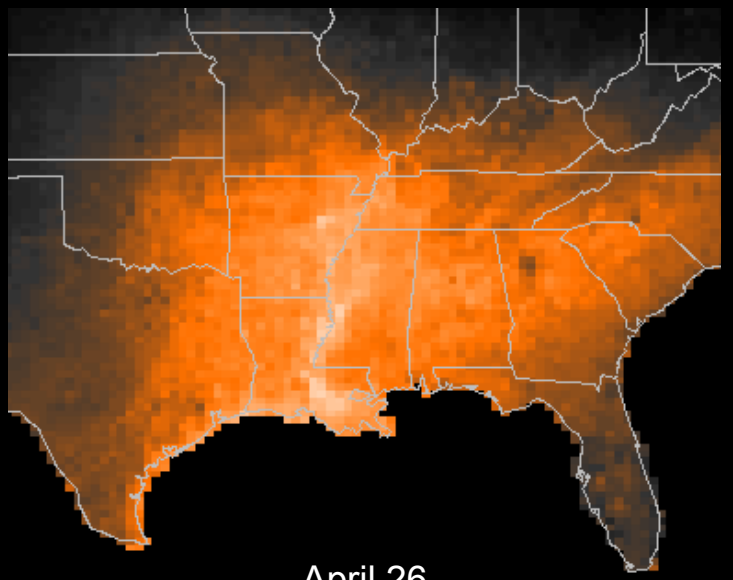
Spring Migration



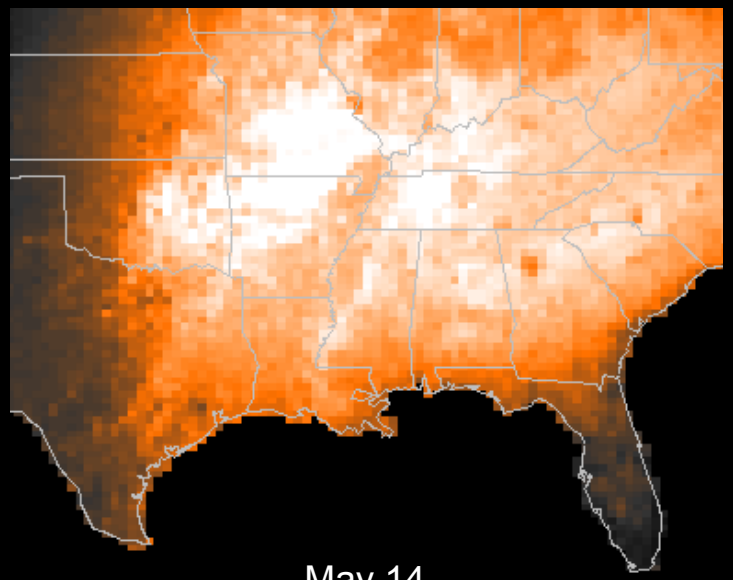
April 7



April 15

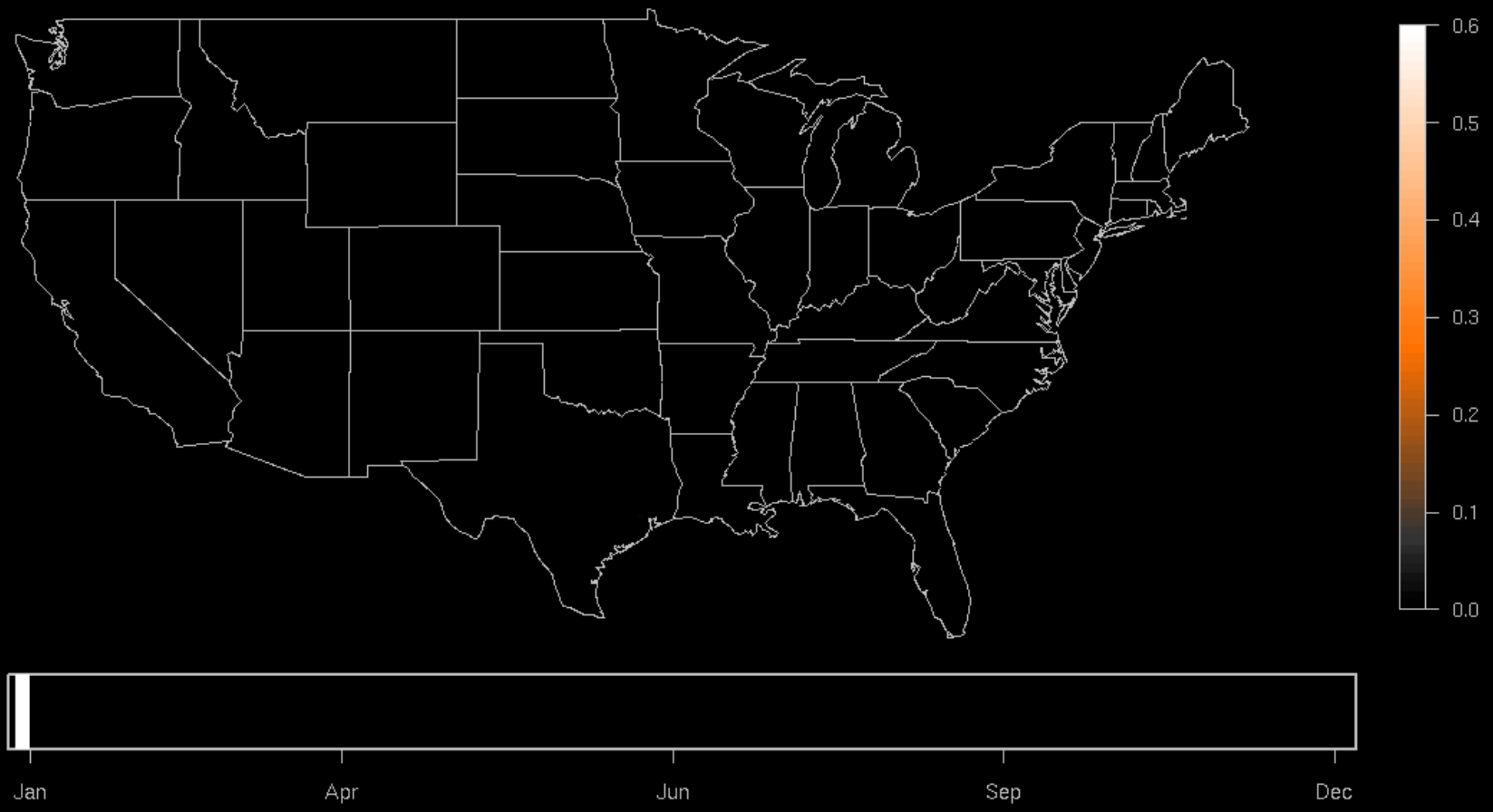


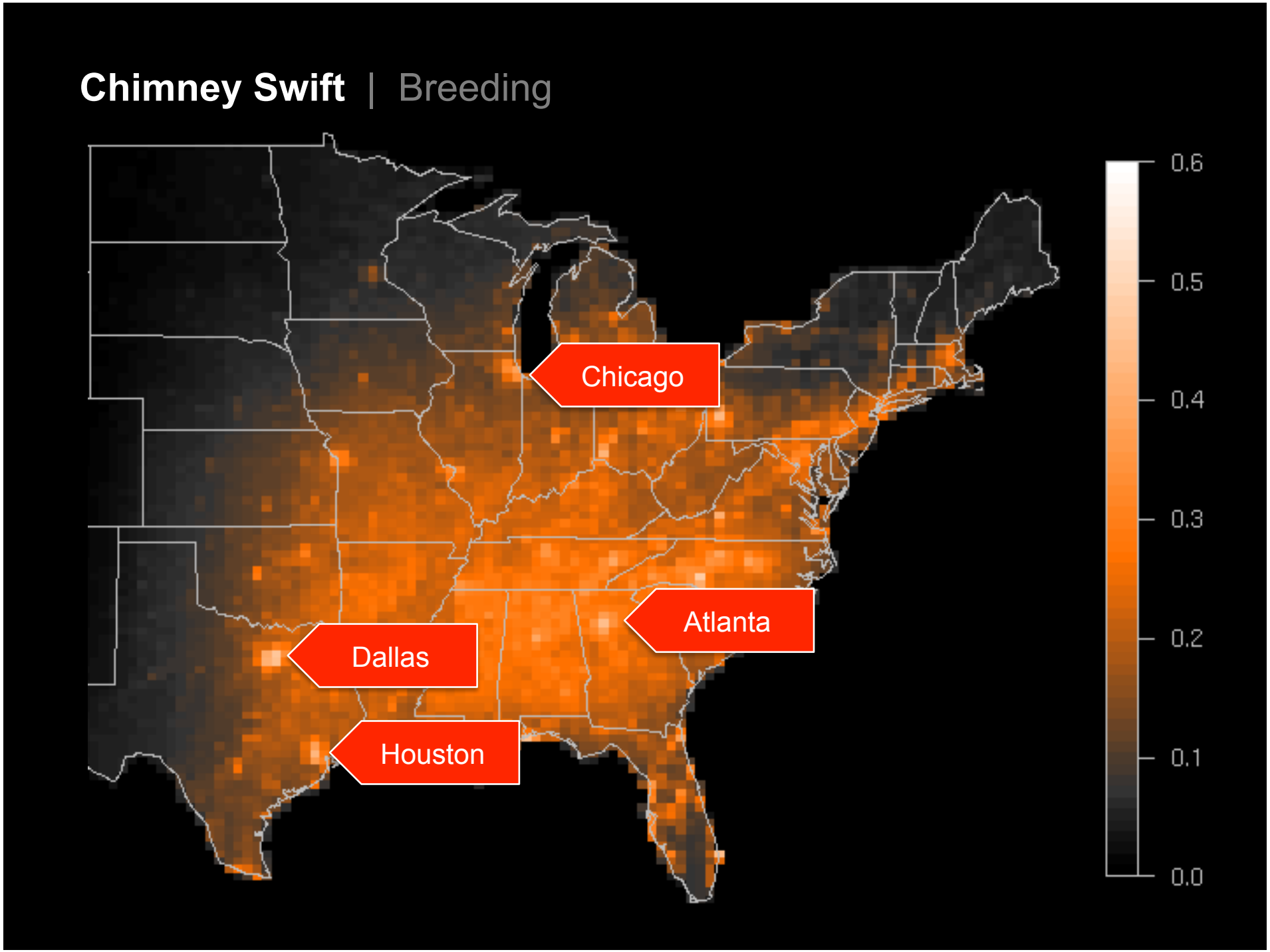
April 26



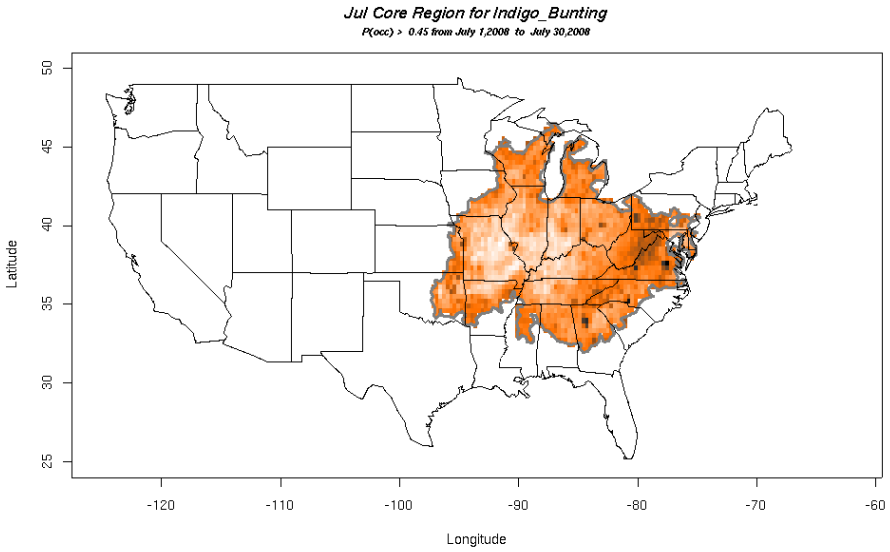
May 14

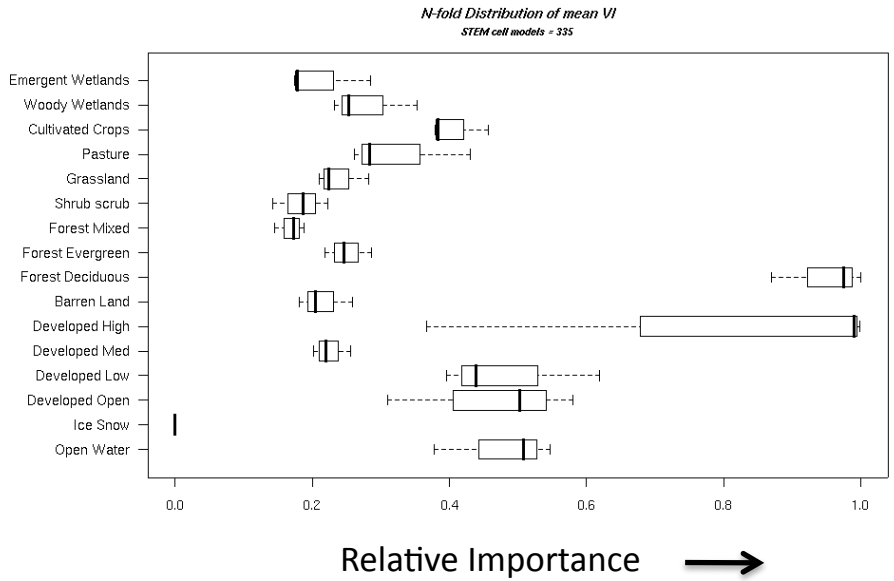
Chimney Swift | Full Year

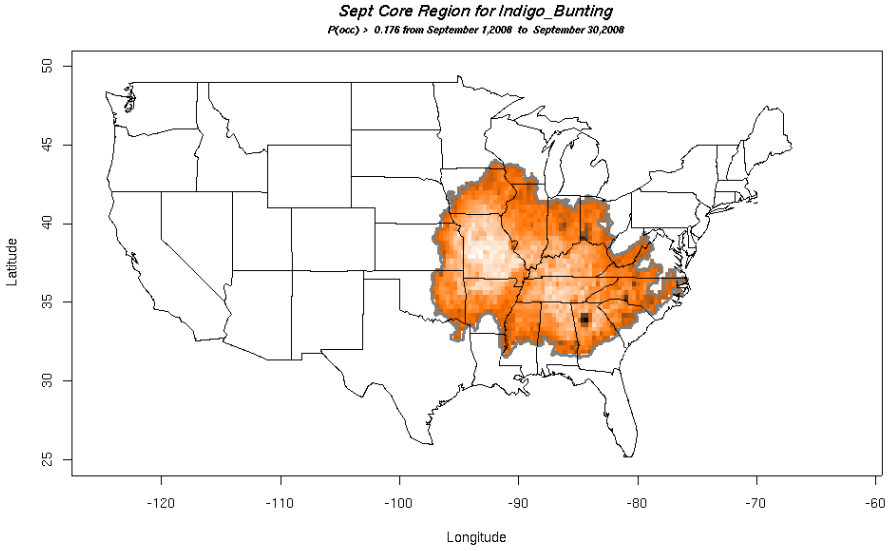


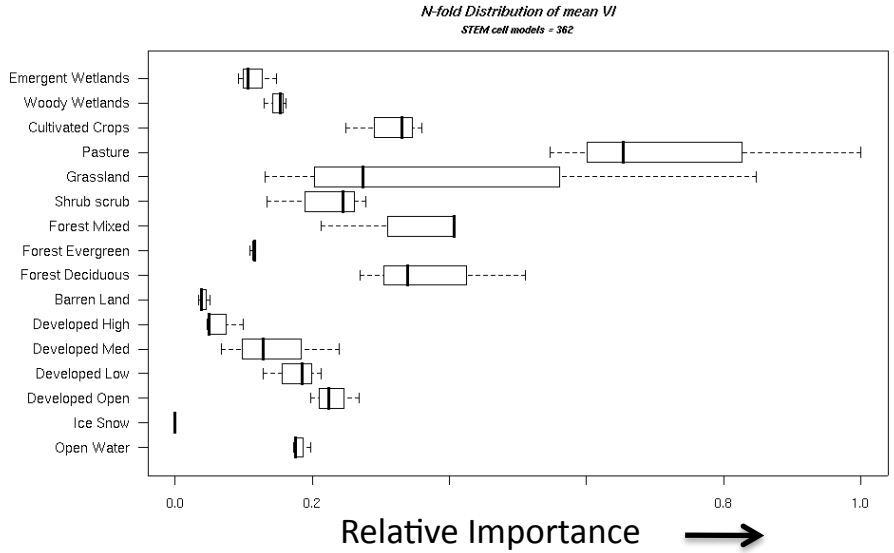

Chimney Swift | Breeding
Chicago
Atlanta
Dallas
Houston
0.6
0.5
0.4
0.3
0.2
0.1
0.0

Indigo Bunting | Relative Variable Importance



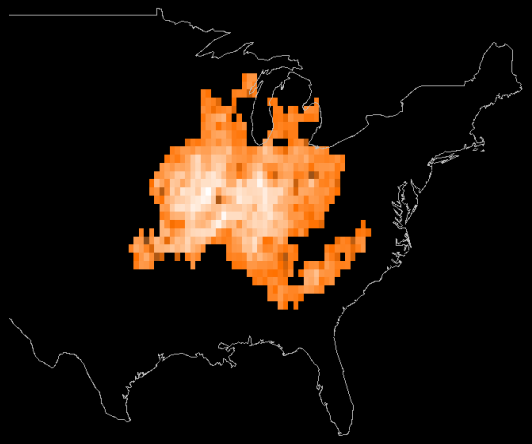


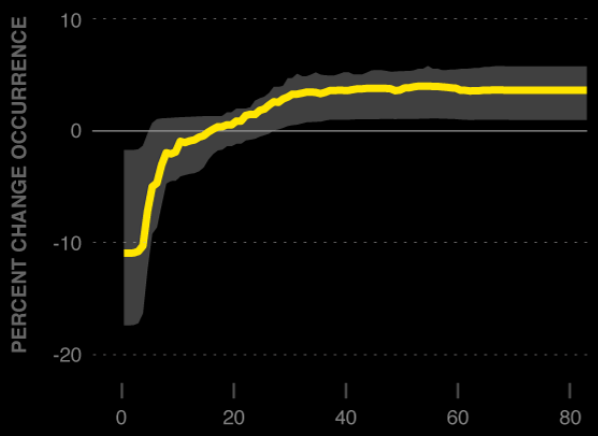


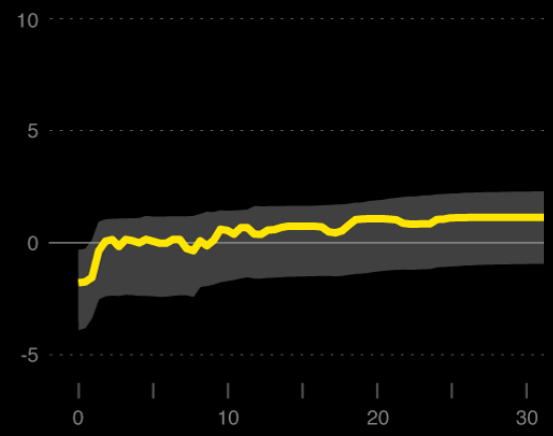


Indigo Bunting | Habitat Preference

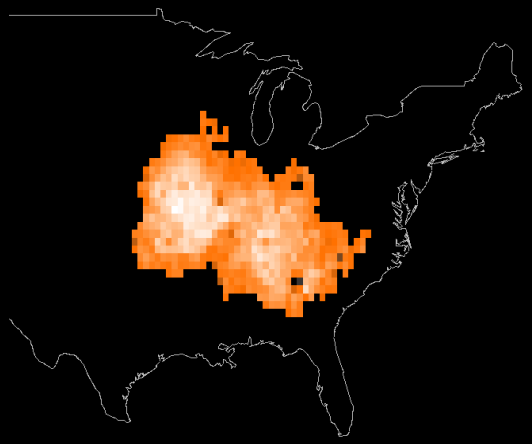
July

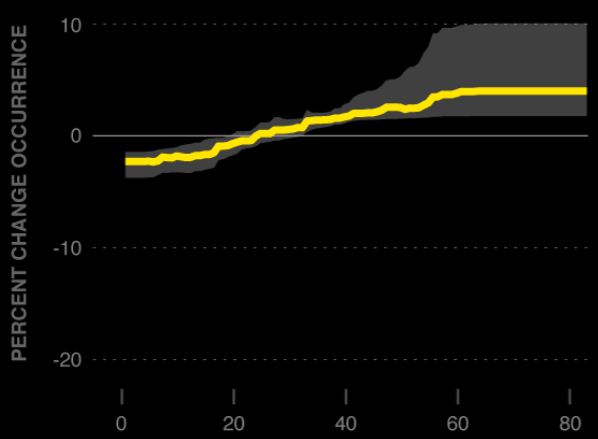


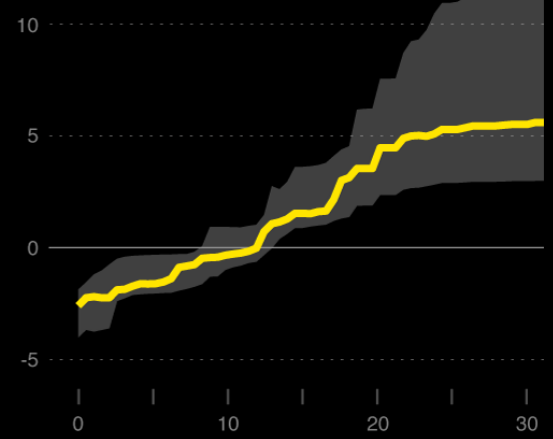




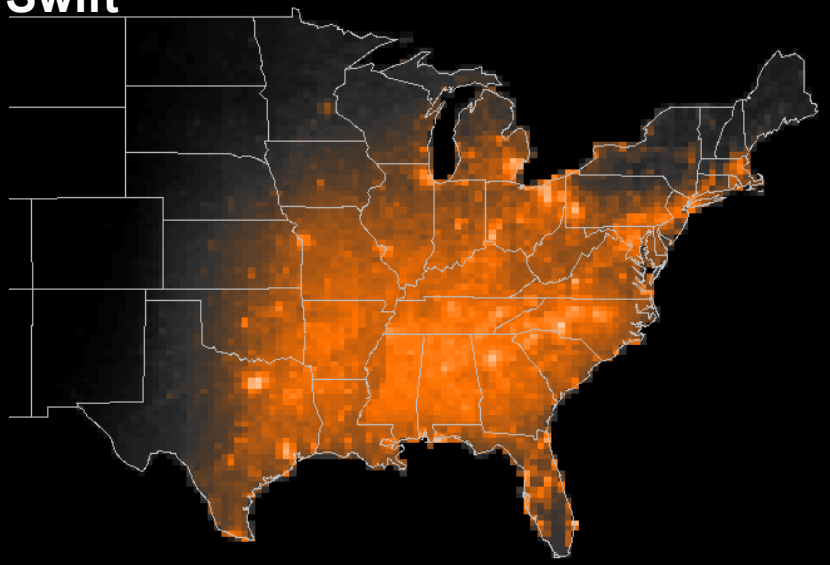
September

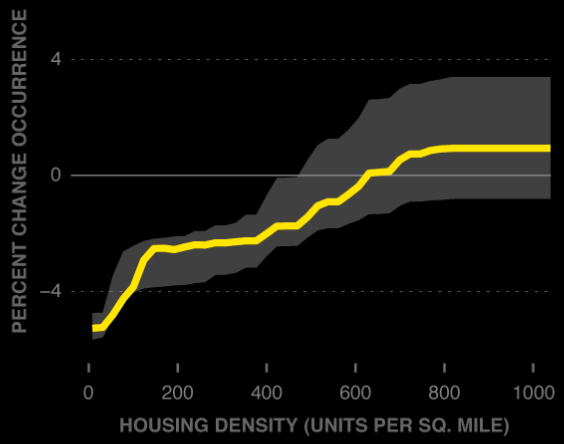




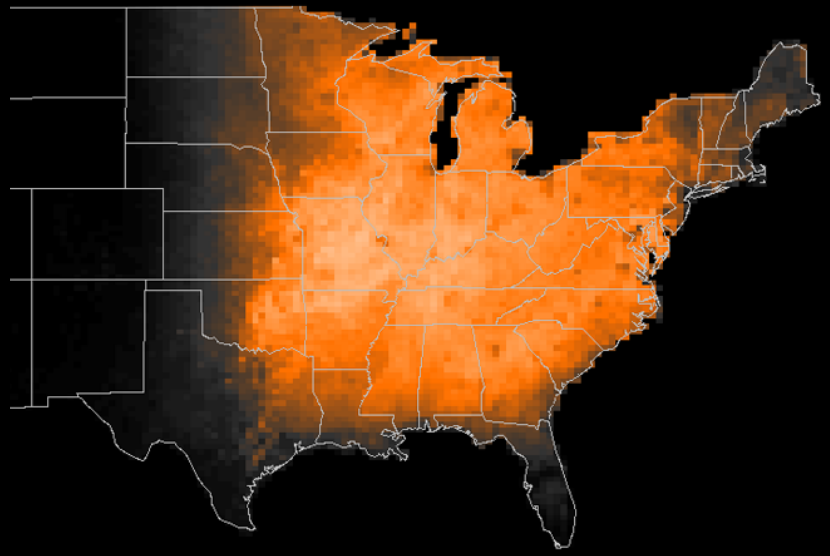


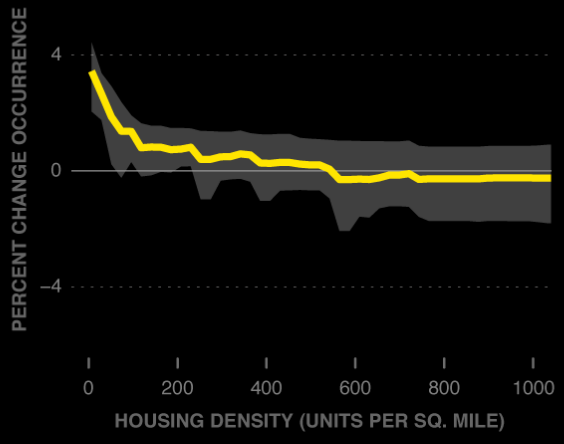
Chimney Swift



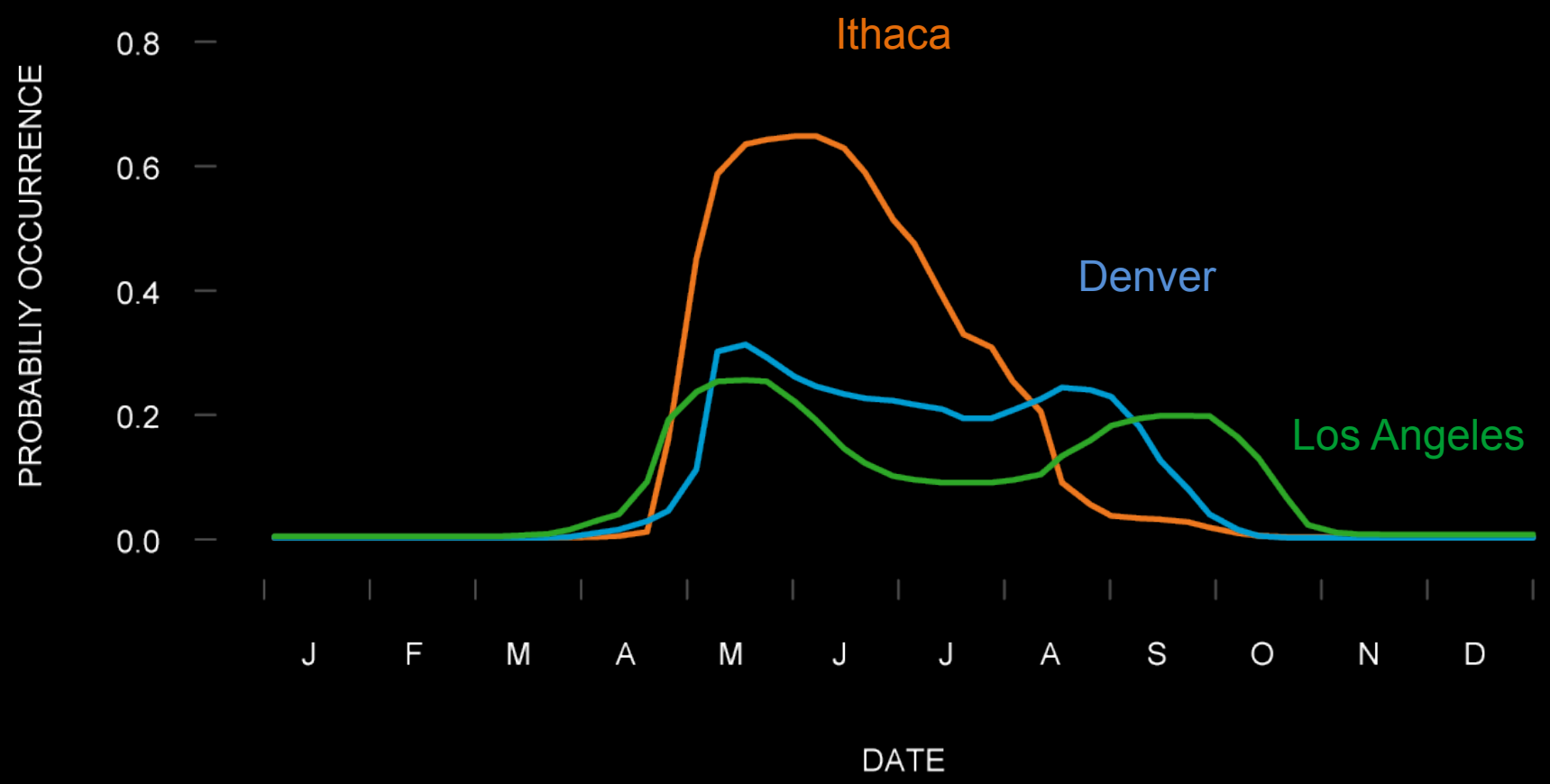


Indigo Bunting

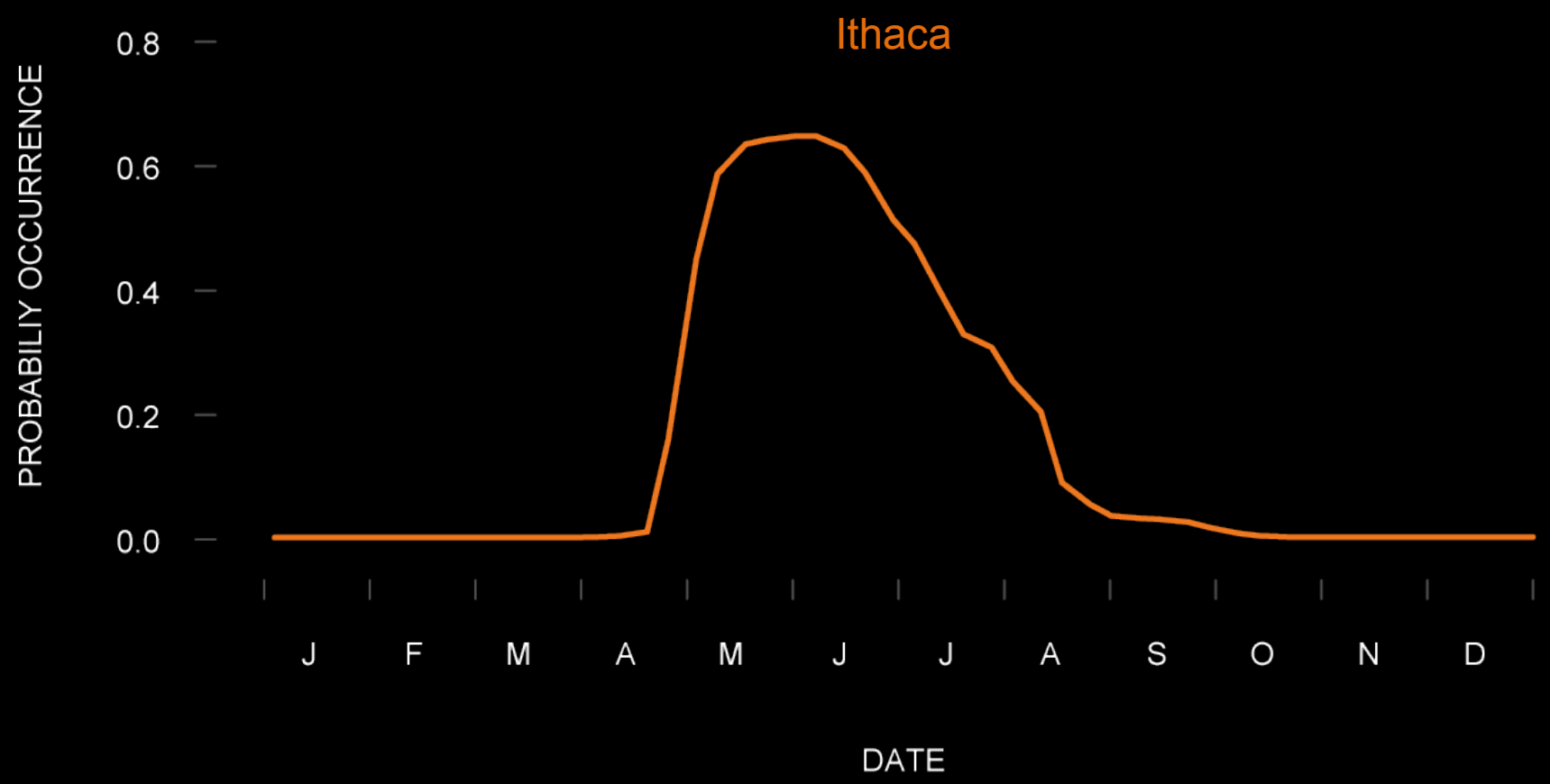




Yellow Warbler | Occurrence Trajectories

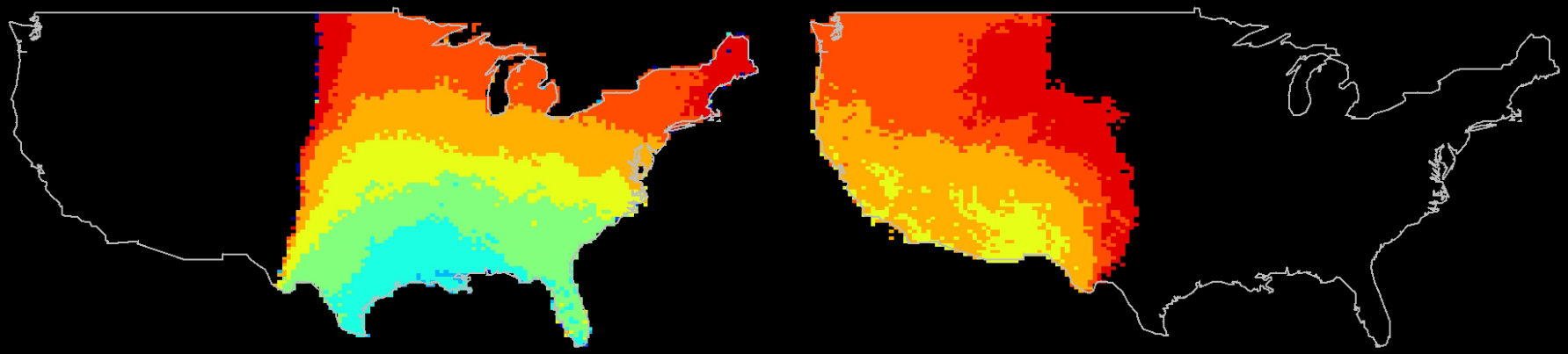


Yellow Warbler | Migration Dates

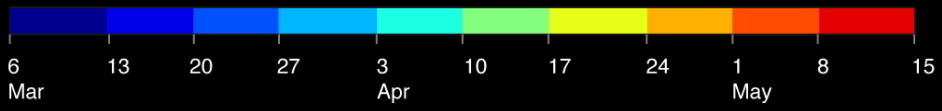


Indigo Bunting

Western Wood-Peewee

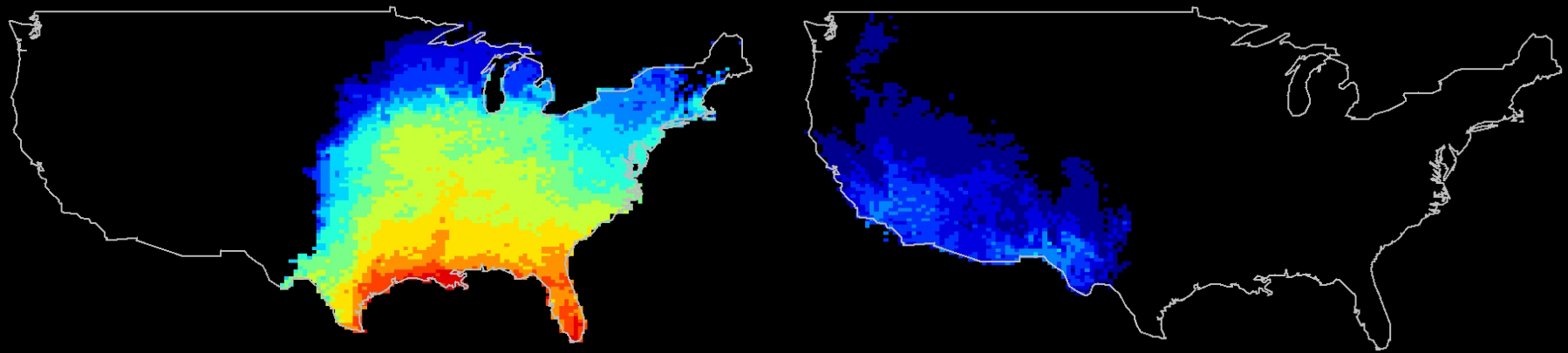


Spring Arrival Dates

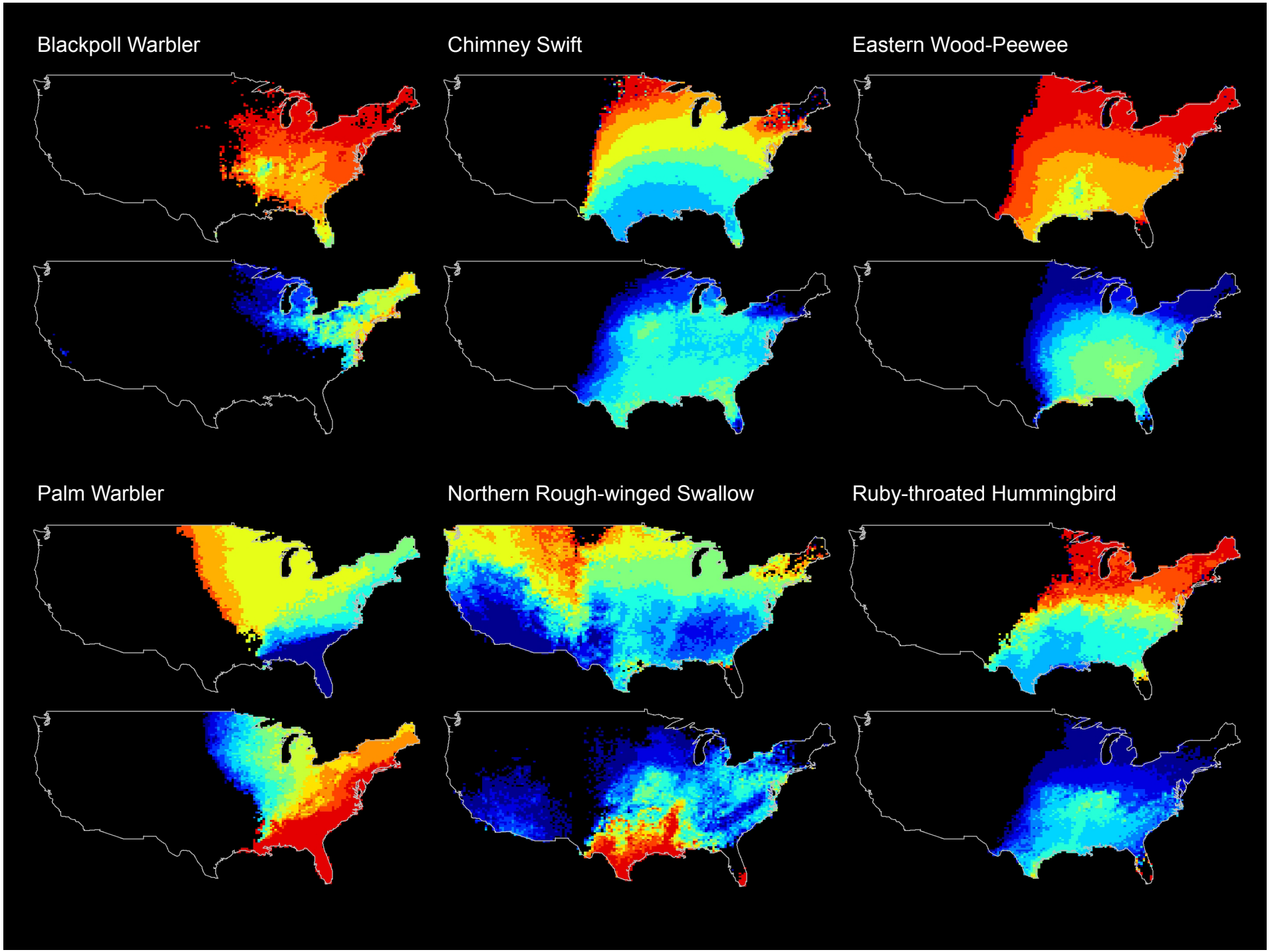


Indigo Bunting

Western Wood-Peewee



Fall Departure Dates


Blackpoll Warbler
Chimney Swift
Eastern Wood-Peewee
Palm Warbler
Northern Rough-winged Swallow
Ruby-throated Hummingbird

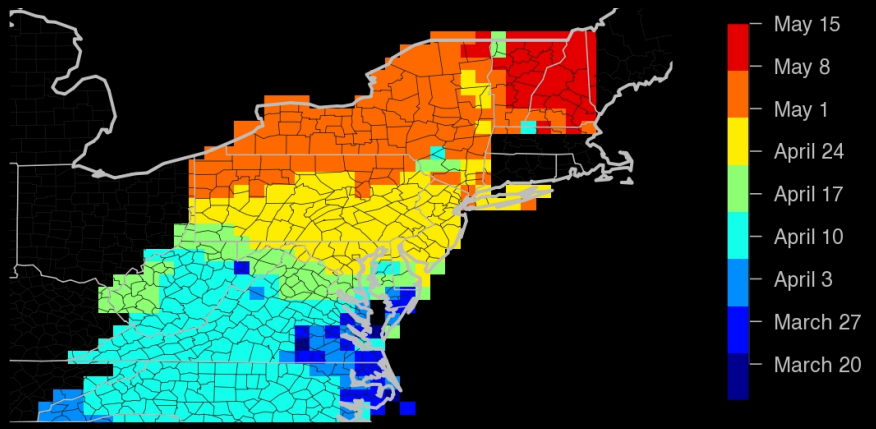
Environmental Cues of Migration

Include spatiotemporally varying covariates to study environmental cues of migration timing: NDVI

Questions

1. Is Red-eyed Vireo migration timing associated with NDVI?
2. If so, how is migration timing, direction, and speed affected by NDVI?

Spring Arrival Dates



Environmental Cues of Migration

Include spatiotemporally varying covariates to study environmental cues of migration timing: NDVI

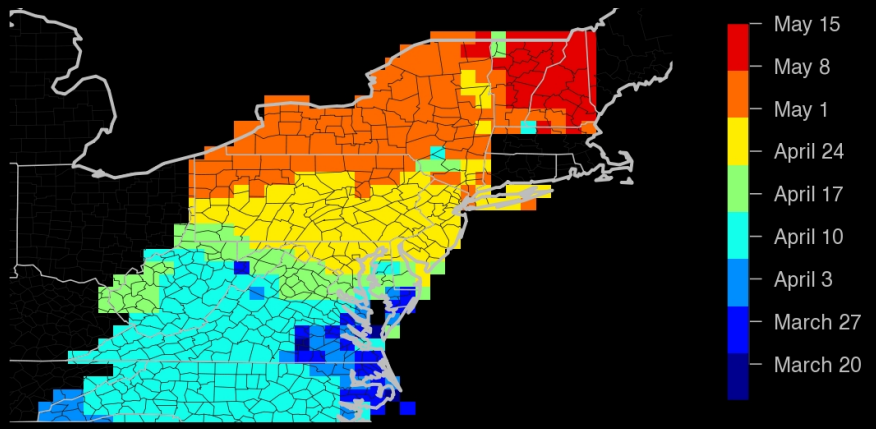
Questions

1. Is Red-eyed Vireo migration timing associated with NDVI?
2. If so, how is migration timing, direction, and speed affected by NDVI?

Experiment

1. Fit STEM with NDVI predictor
2. Advance “greening dates” by 14 days

Spring Arrival Dates



Environmental Cues of Migration

Include spatiotemporally varying covariates to study environmental cues of migration timing: NDVI

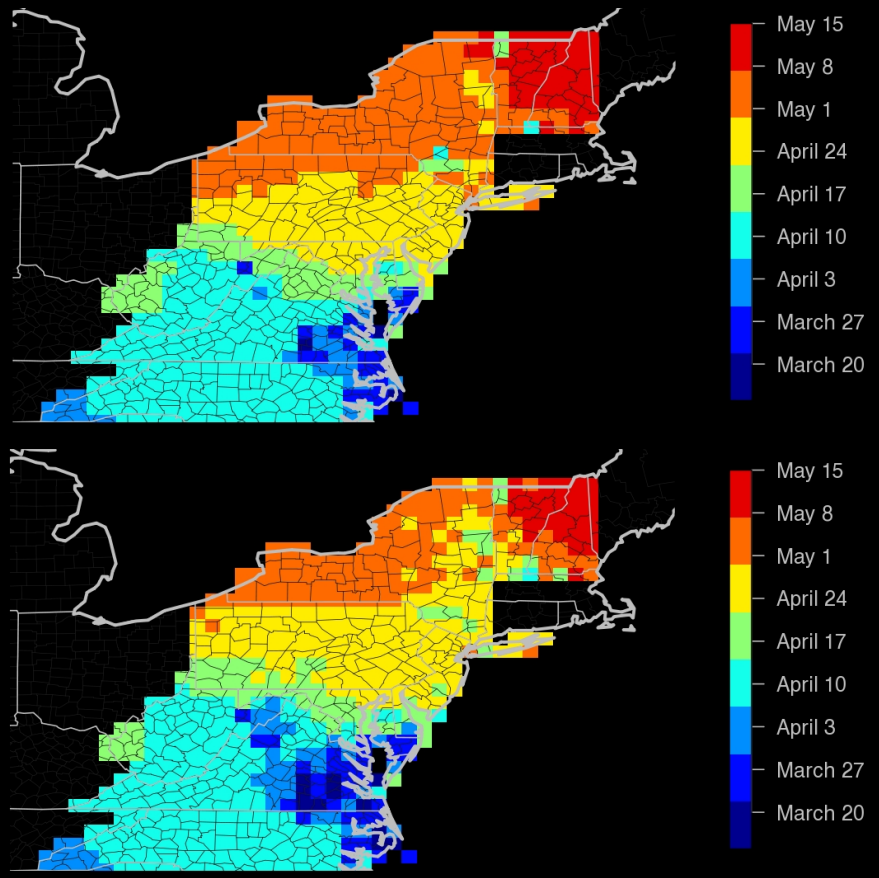
Questions

1. Is Red-eyed Vireo migration timing associated with NDVI?
2. If so, how is migration timing, direction, and speed affected by NDVI?

Experiment

1. Fit STEM with NDVI predictor
2. Advance "greening dates" by 14 days

Spring Arrival Dates



Spring Arrival Dates with Advanced Greening

Environmental Cues of Migration

Include spatiotemporally varying covariates to study environmental cues of migration timing: NDVI

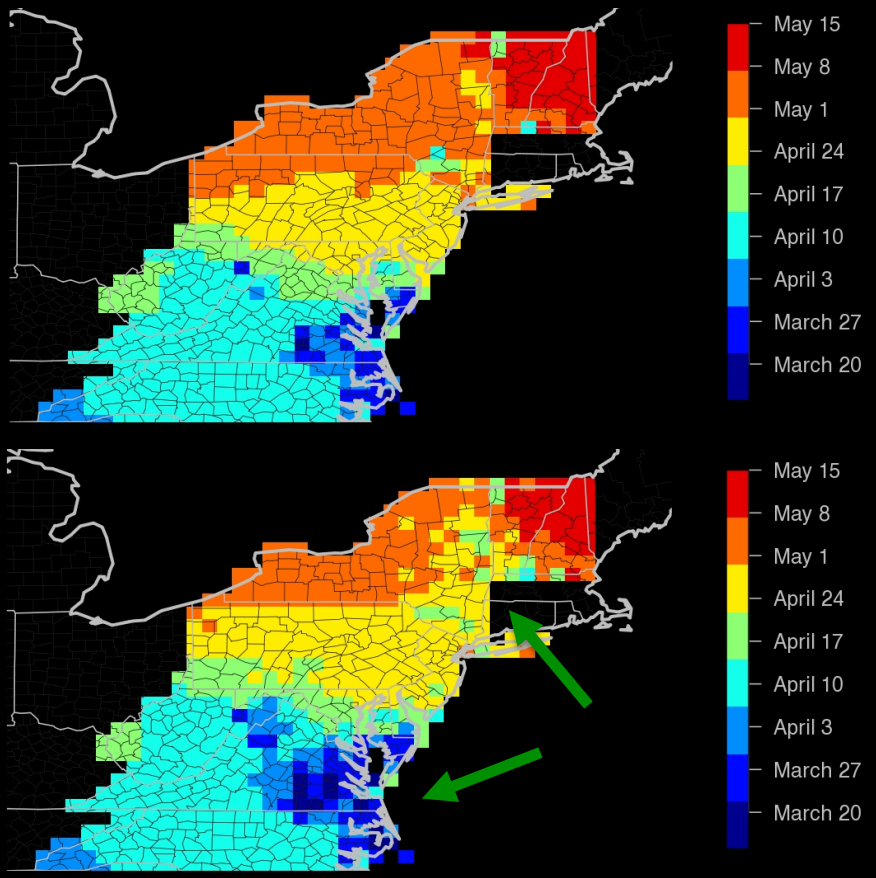
Questions

1. Is Red-eyed Vireo migration timing associated with NDVI?
2. If so, how is migration timing, direction, and speed affected by NDVI?

Experiment

1. Fit STEM with NDVI predictor
2. Advance "greening dates" by 14 days

Spring Arrival Dates



Spring Arrival Dates with Advanced Greening

Overview

- Exploratory Dynamic Spatiotemporal Modeling
- Automatically adapts to many dynamic processes
- Multi (bi)–scale approach

Next Steps

- Replicate over Years to explore inter-annual differences
- Improve Validation Methods – Spatial & Temporal correlation
- Identify & control sources of bias
  - Spatial sampling bias
  - Habitat sampling bias
  - Temporal variation in detection rates
  - Etc.
- Expand NDVI analysis
  - More species – compare different migration strategies
  - VI & Partial Dependence for NDVI
  - Explore other ST covariates