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

TheCornellLab of Ornithology

NSF

LEON LEVY FOUNDATION
The Legacy Of Leon Levy

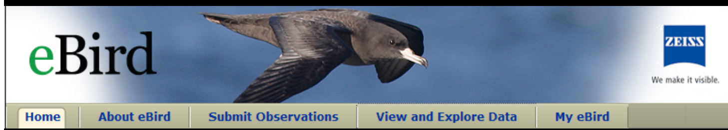
Goal: Modeling Species Distribution Dynamics

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

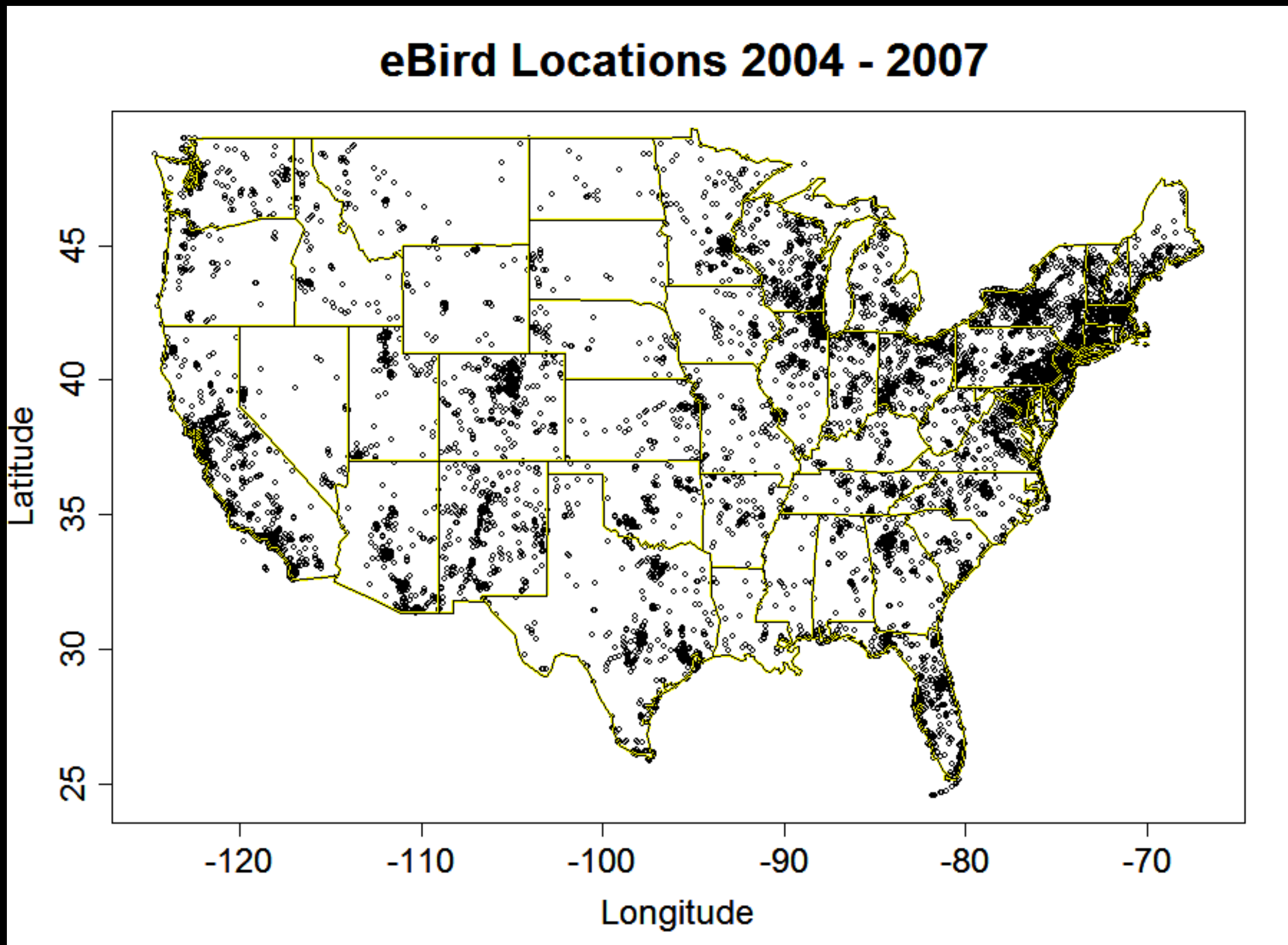
~150,000 observations

~30,000 locations

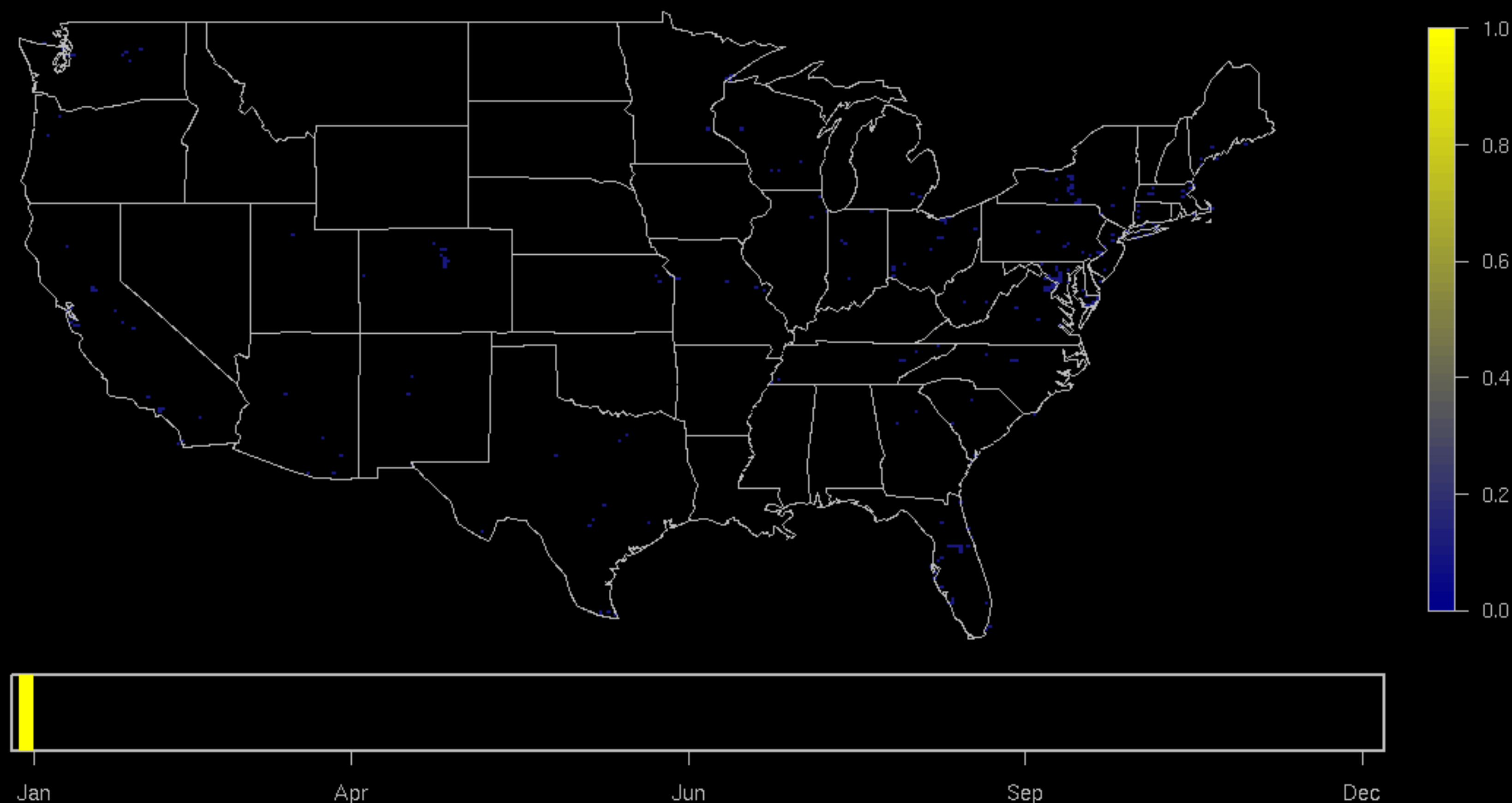
National Landcover Data (NLCD)

– 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
Longitude
25
30
35
40
45
-120
-110
-100
-90
-80
-70

eBird data | 2008 traveling count protocol



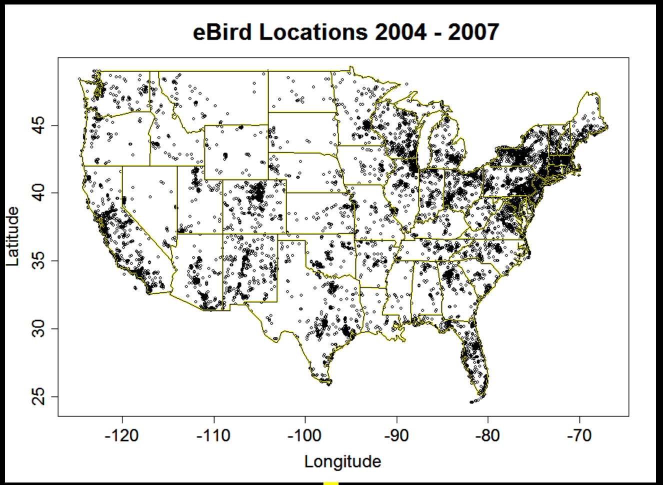
Species Distribution Modeling

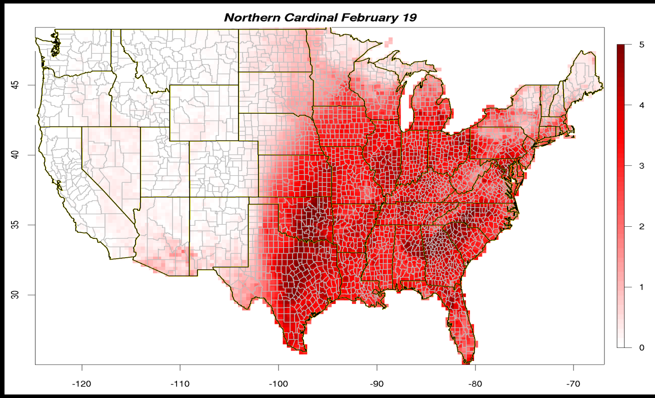
Goal: broad-scale & fine resolution

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

Modeling also buys us

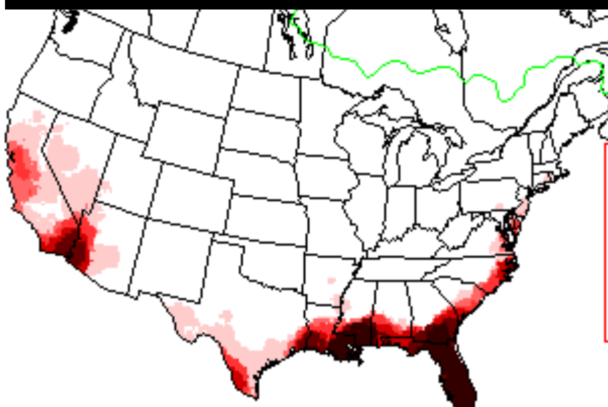
- Control bias
- Quantify uncertainty
- Framework for predictive experimentation



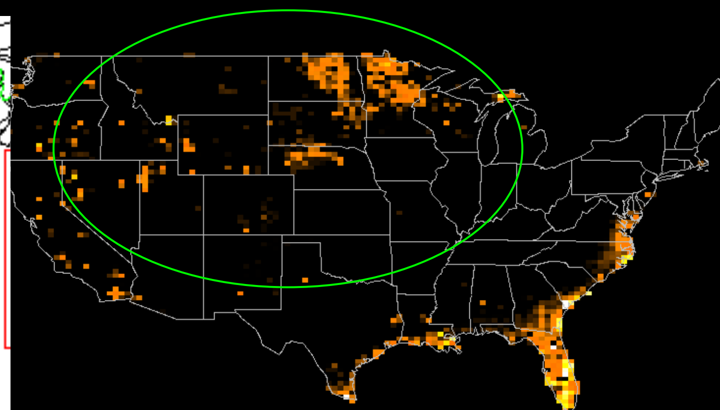


Tree Swallow
Winter Distribution Analysis
Bagged Decision Tree

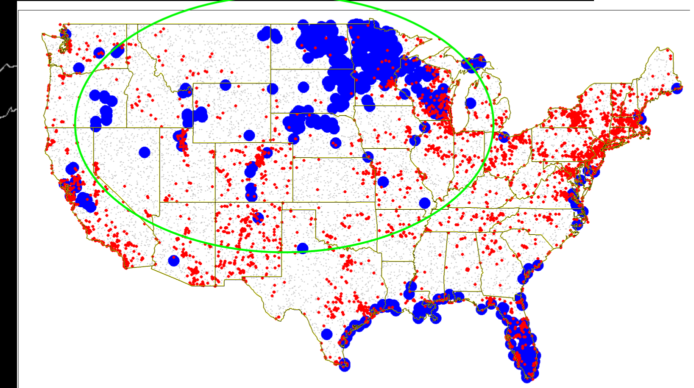
(Tachycineta bicolor)



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)

$$y = f(X)$$

Restricted Support Set (Q)

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

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

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

SpatioTemporal Exploratory Model (STEM)

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

$$y = f(X)$$

Restricted Support Set (q)

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

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

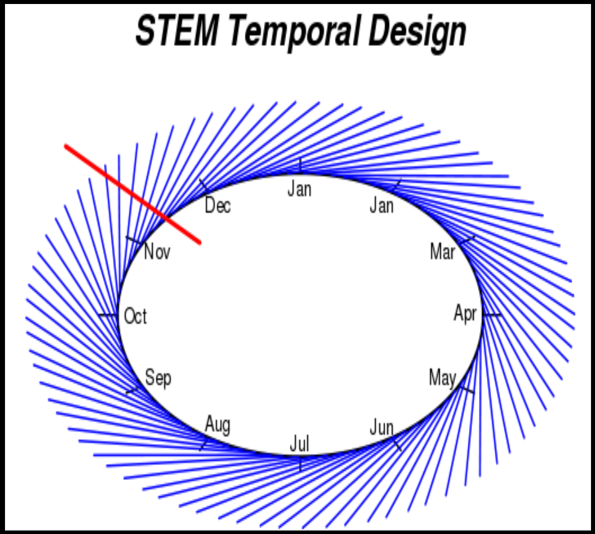
i^{th} ST explicit base model

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

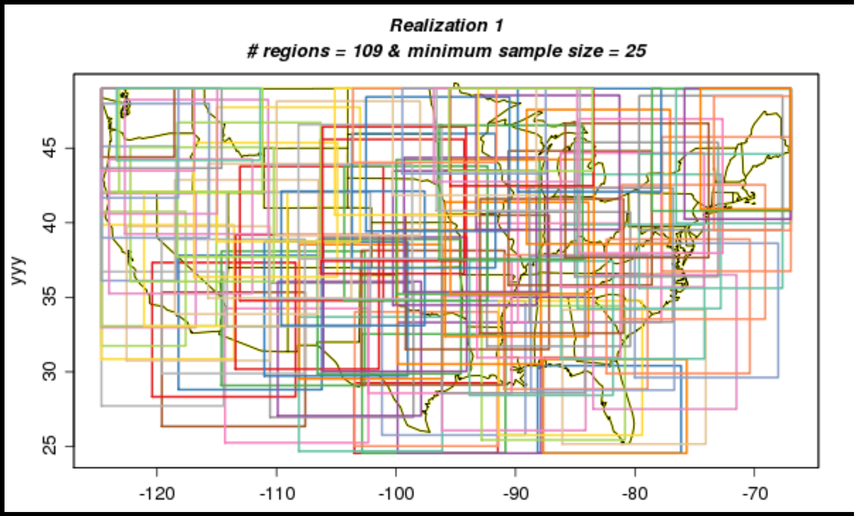


Temporal Design:

- 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

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

<u>Quantitative Intuition</u>

Statistical Experimental Design:
Block over ST regimes to control variance

Bagging: resample block-level variance and average

<u>Ecological Intuition</u>

Local predictor-response learning
No “long-range” learning

“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(Predictors)

Predictors → → Observations

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

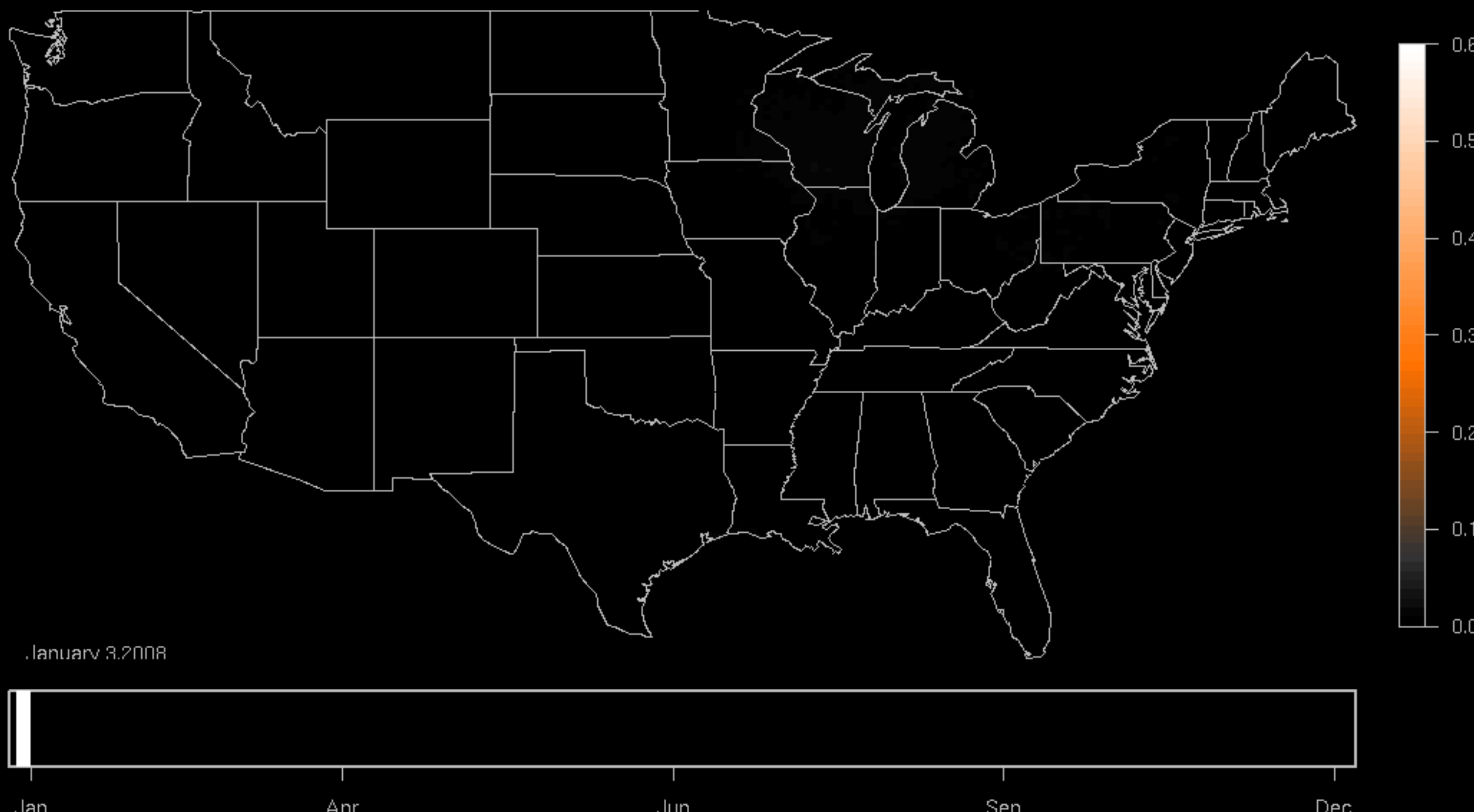
Strengths

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

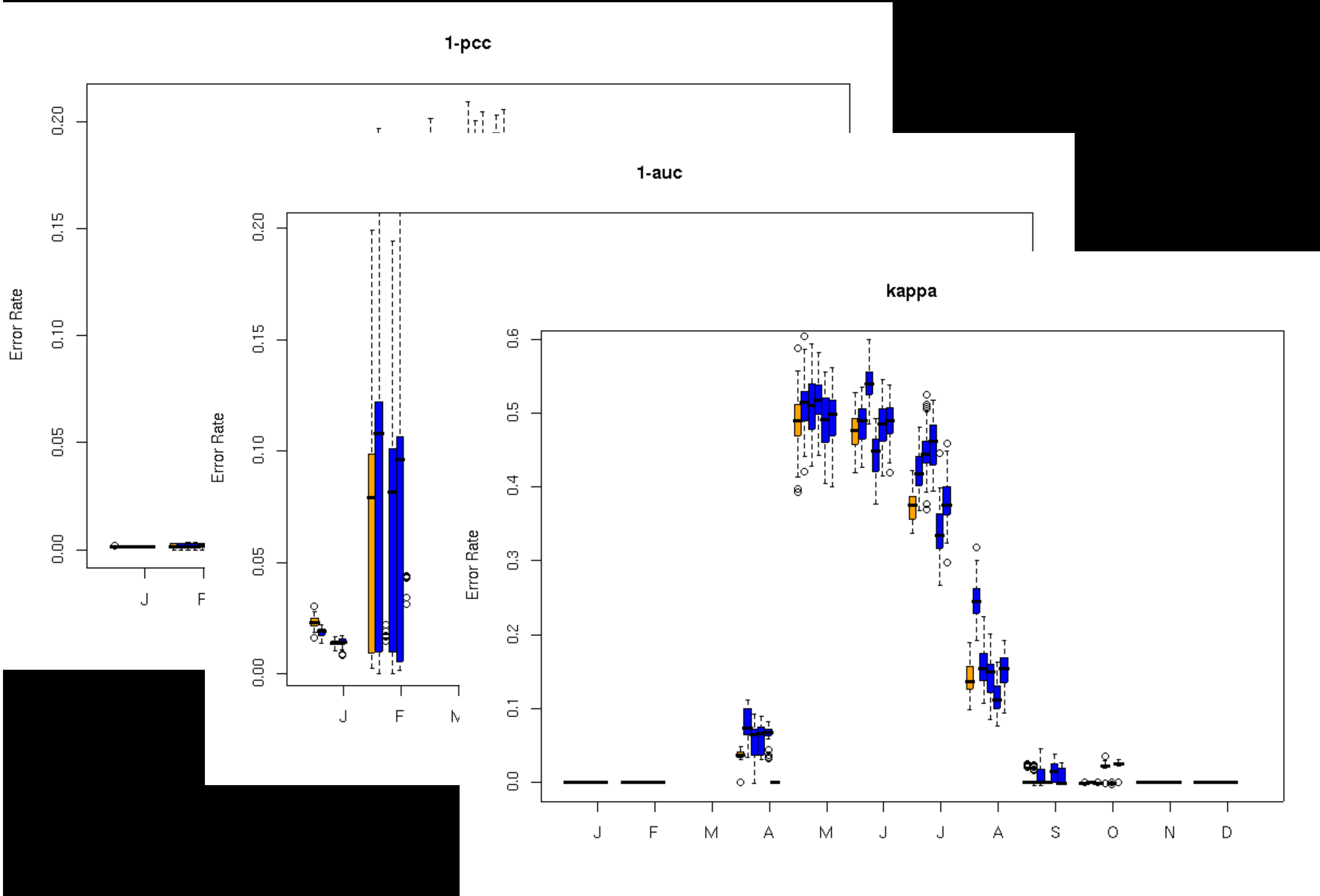
- Monotone transformations
- Nonlinear predictor effects
- Automatic interaction modeling
- Missing covariates

➔ Predictor Selection is the main analyst decision

Indigo Bunting | Full Year

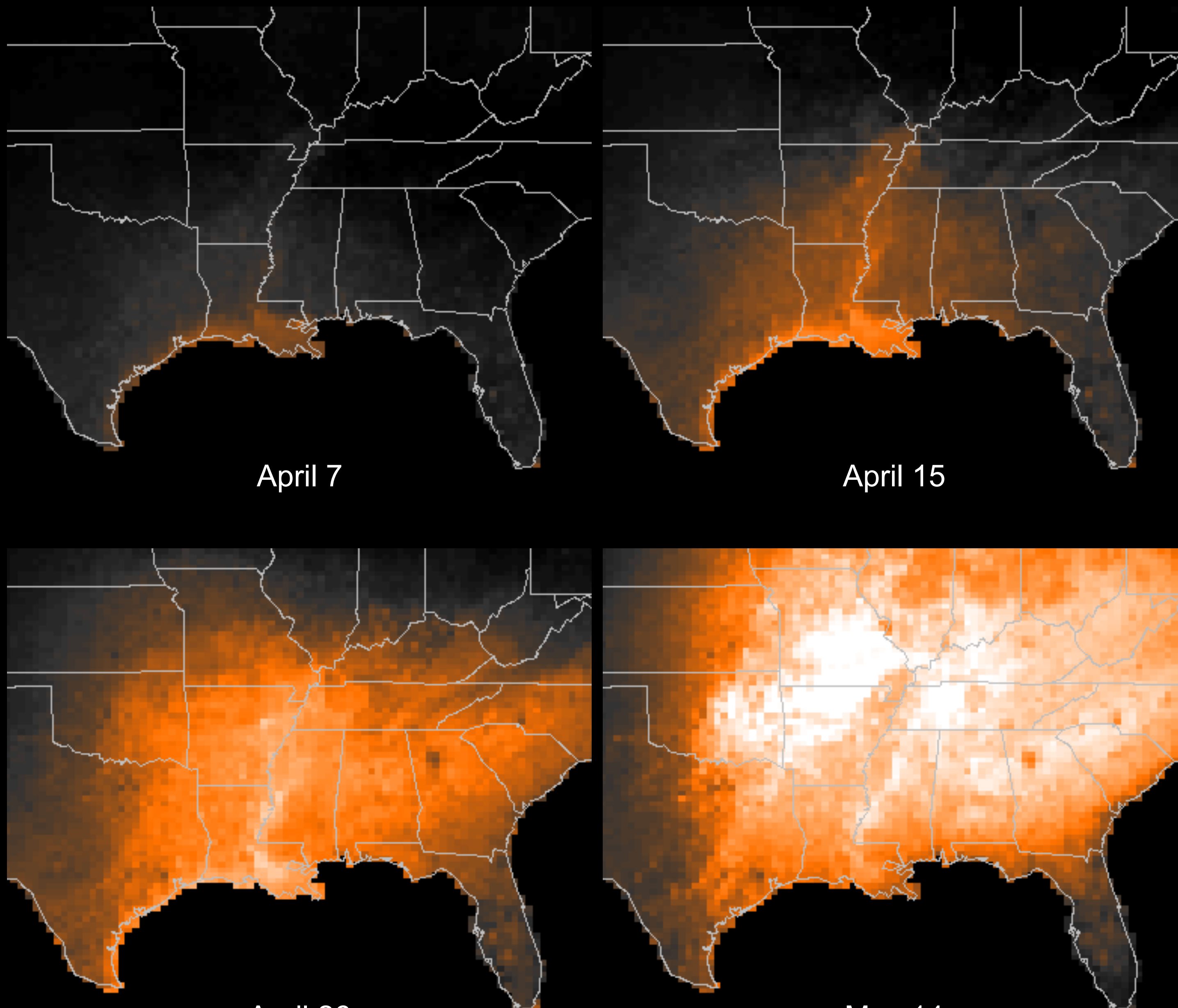


Indigo Bunting | Monthly Predictive Performance



Indigo
Bunting

Spring
Migration



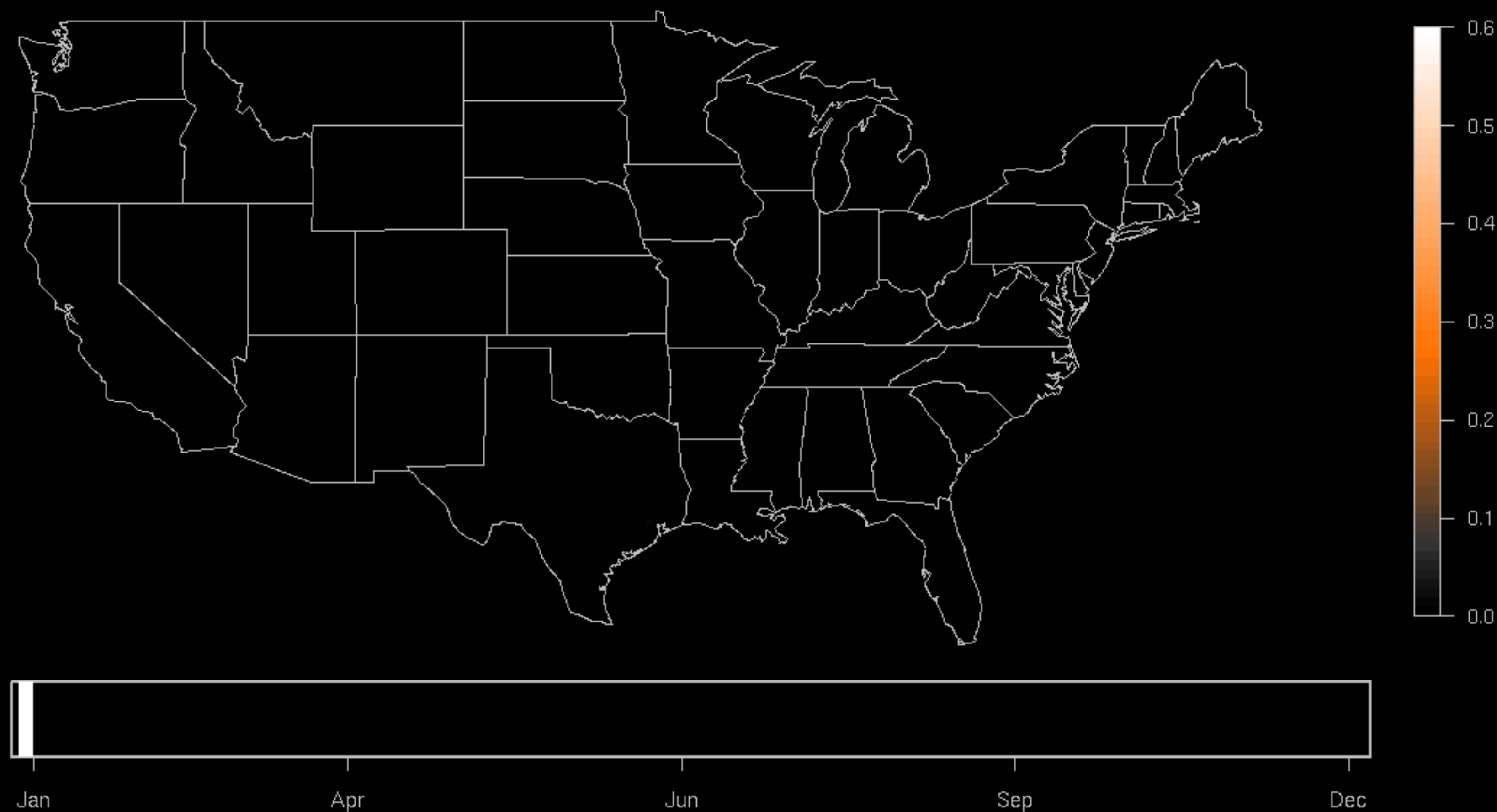
April 7

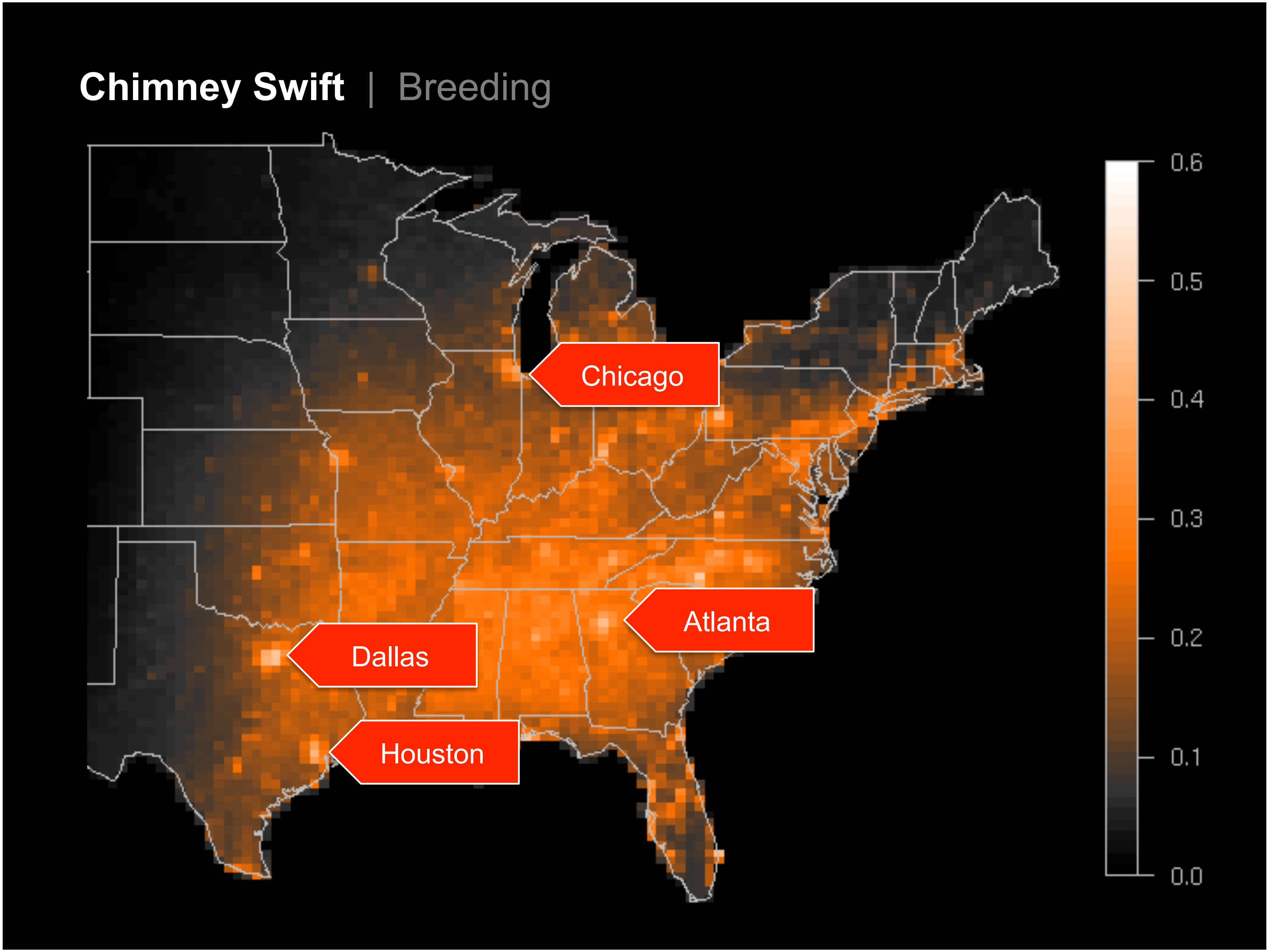
April 15

April 26

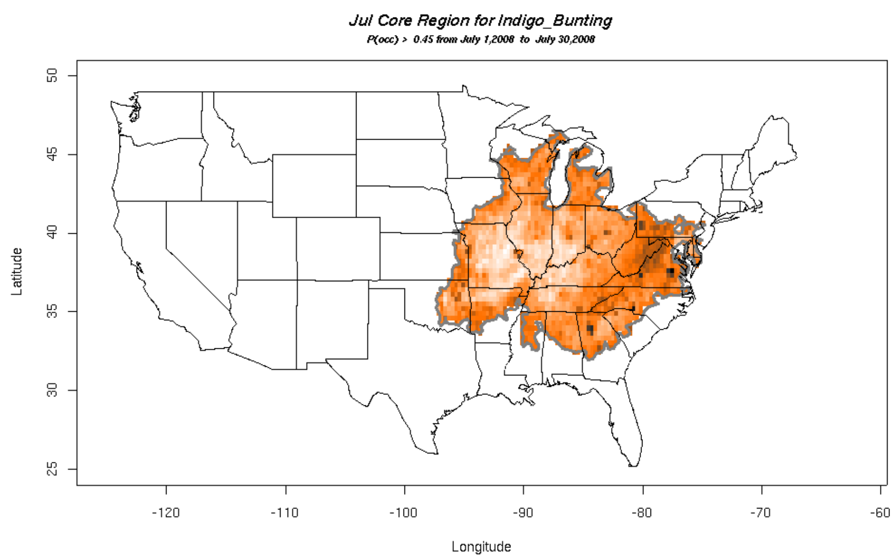
May 14

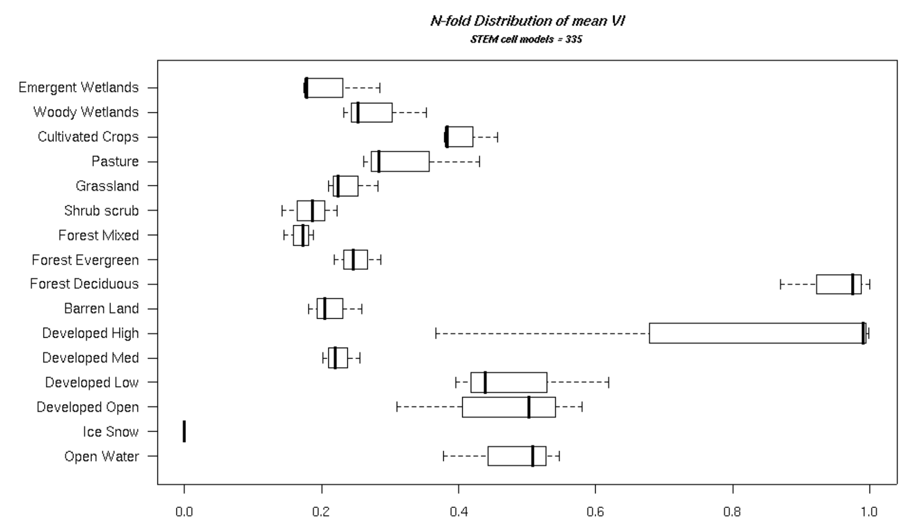
Chimney Swift | Full Year



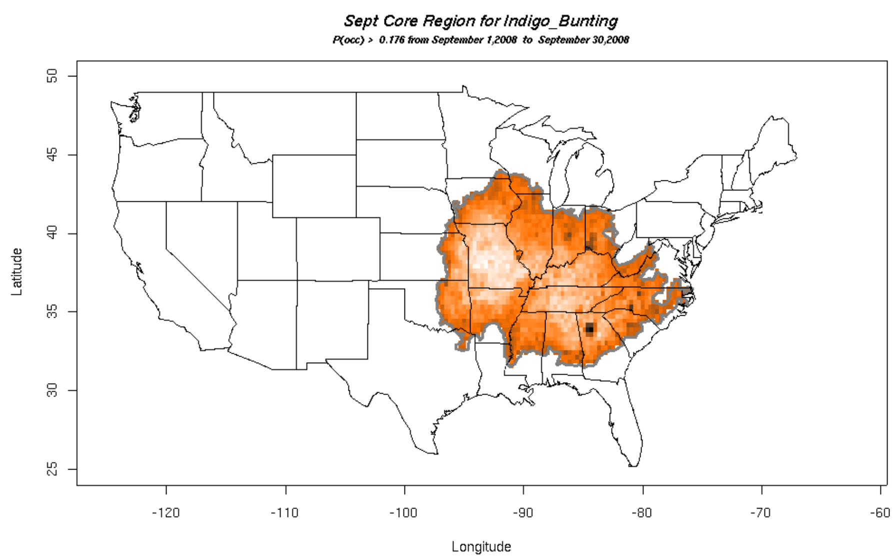

Chimney Swift | Breeding
Chicago
Atlanta
Dallas
Houston
0.6
0.5
0.4
0.3
0.2
0.1
0.0

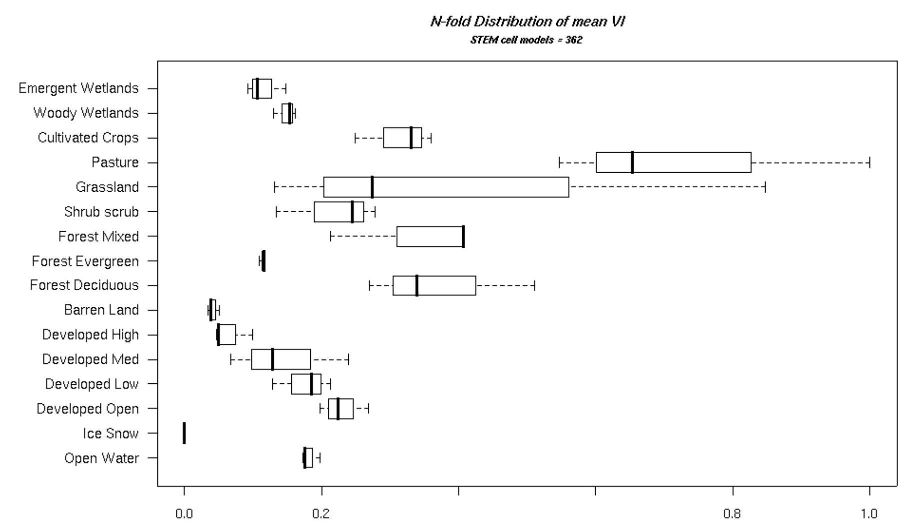
Indigo Bunting | Relative Variable Importance



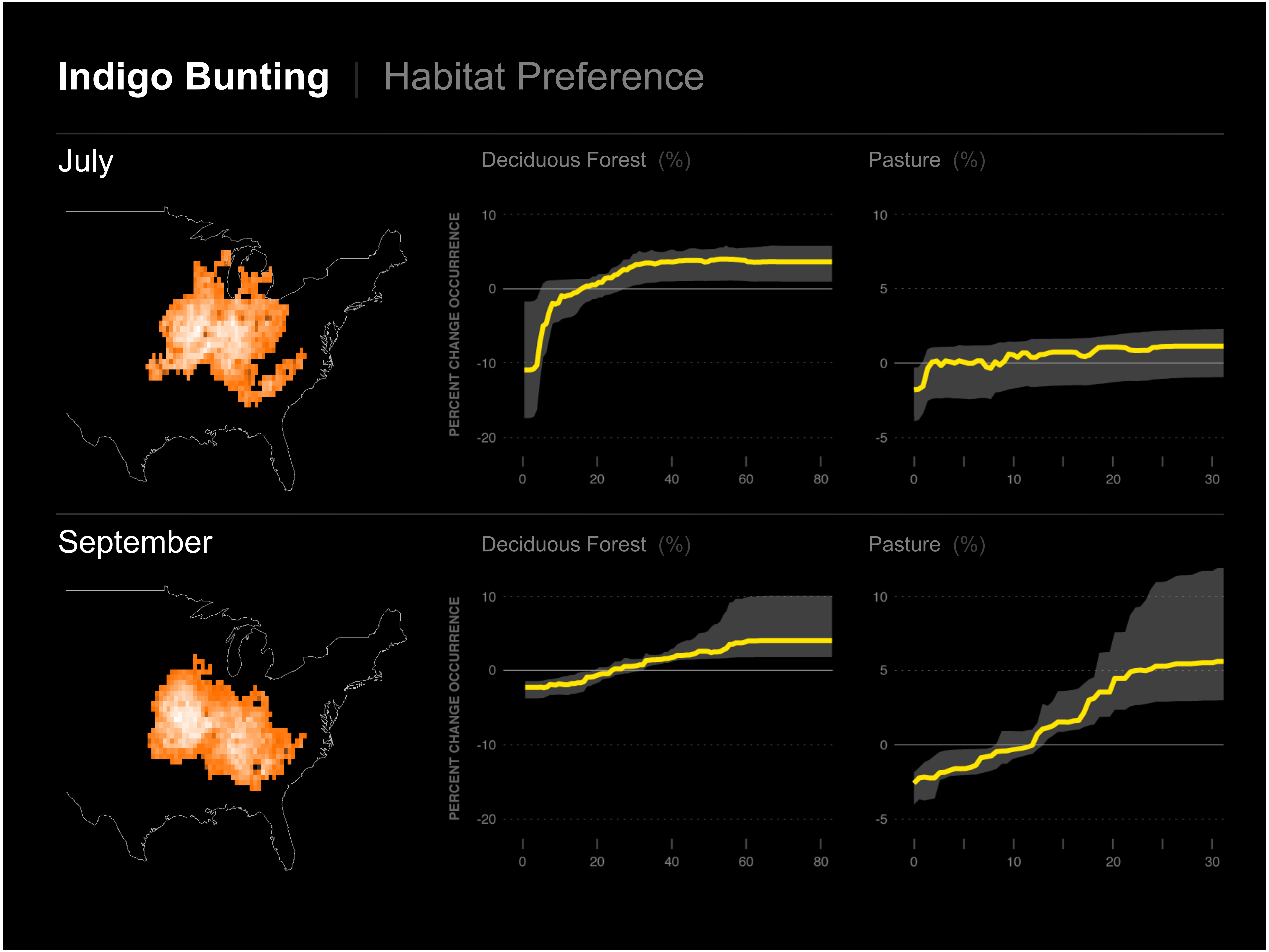


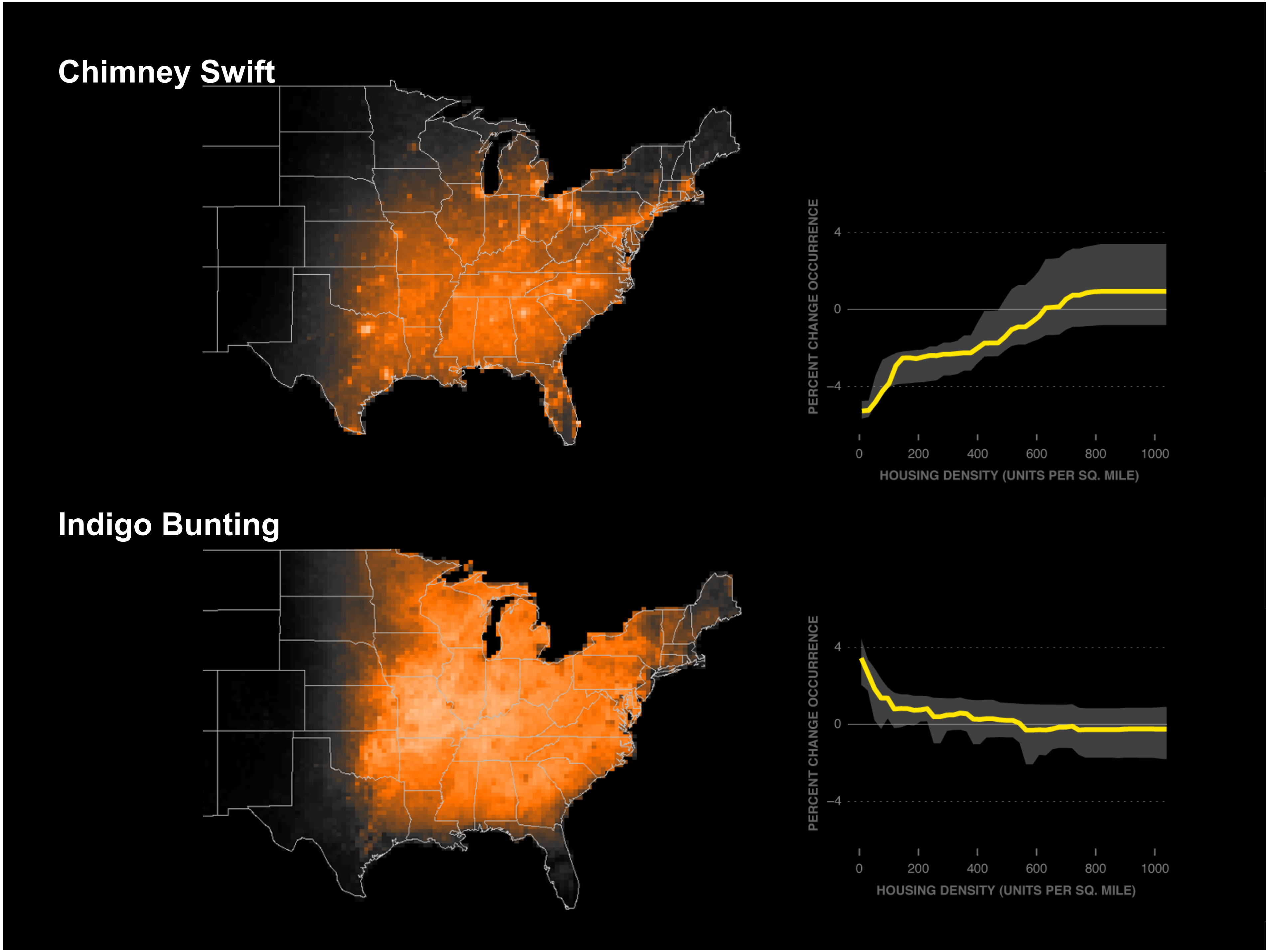
Relative Importance →

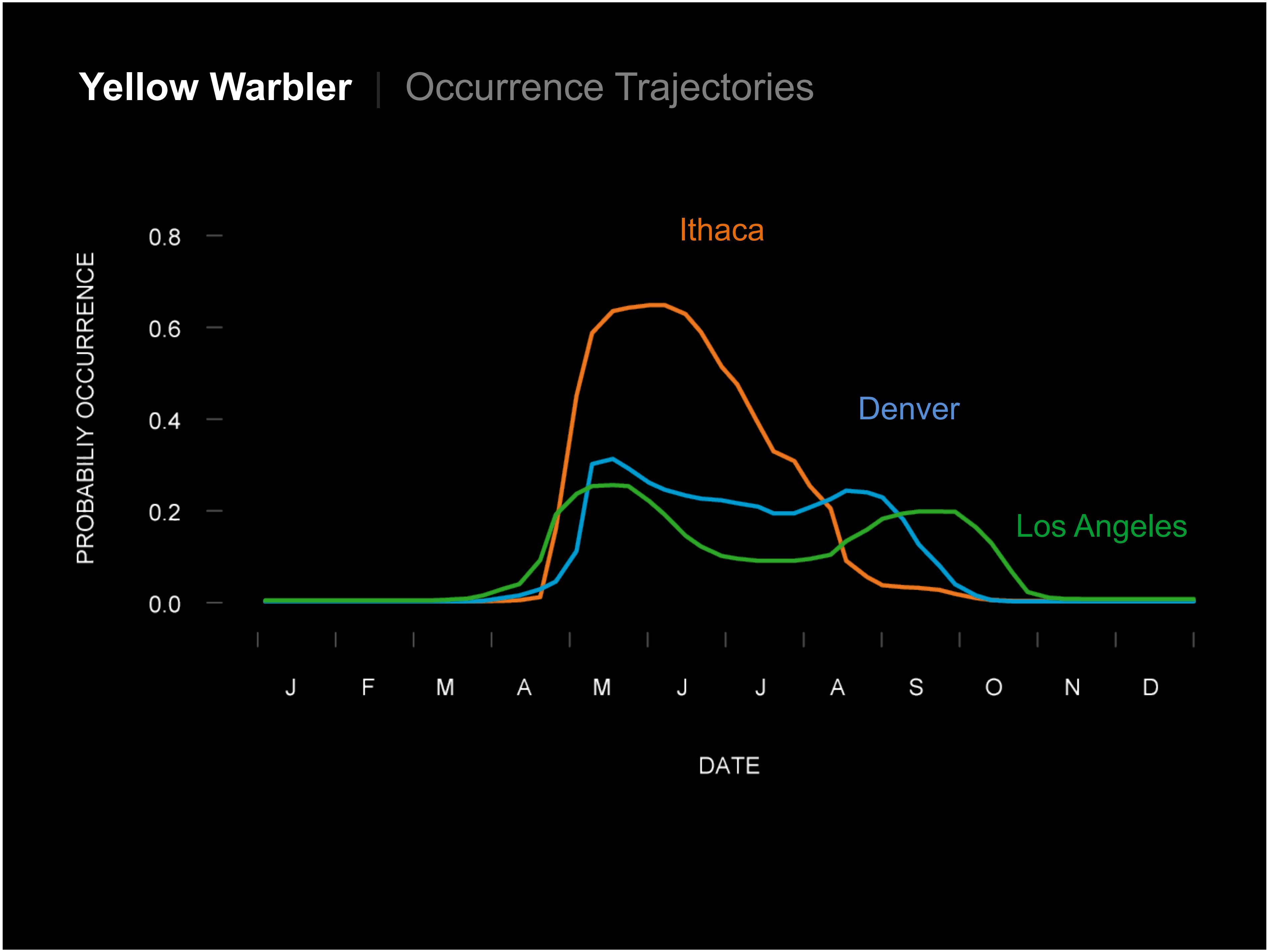




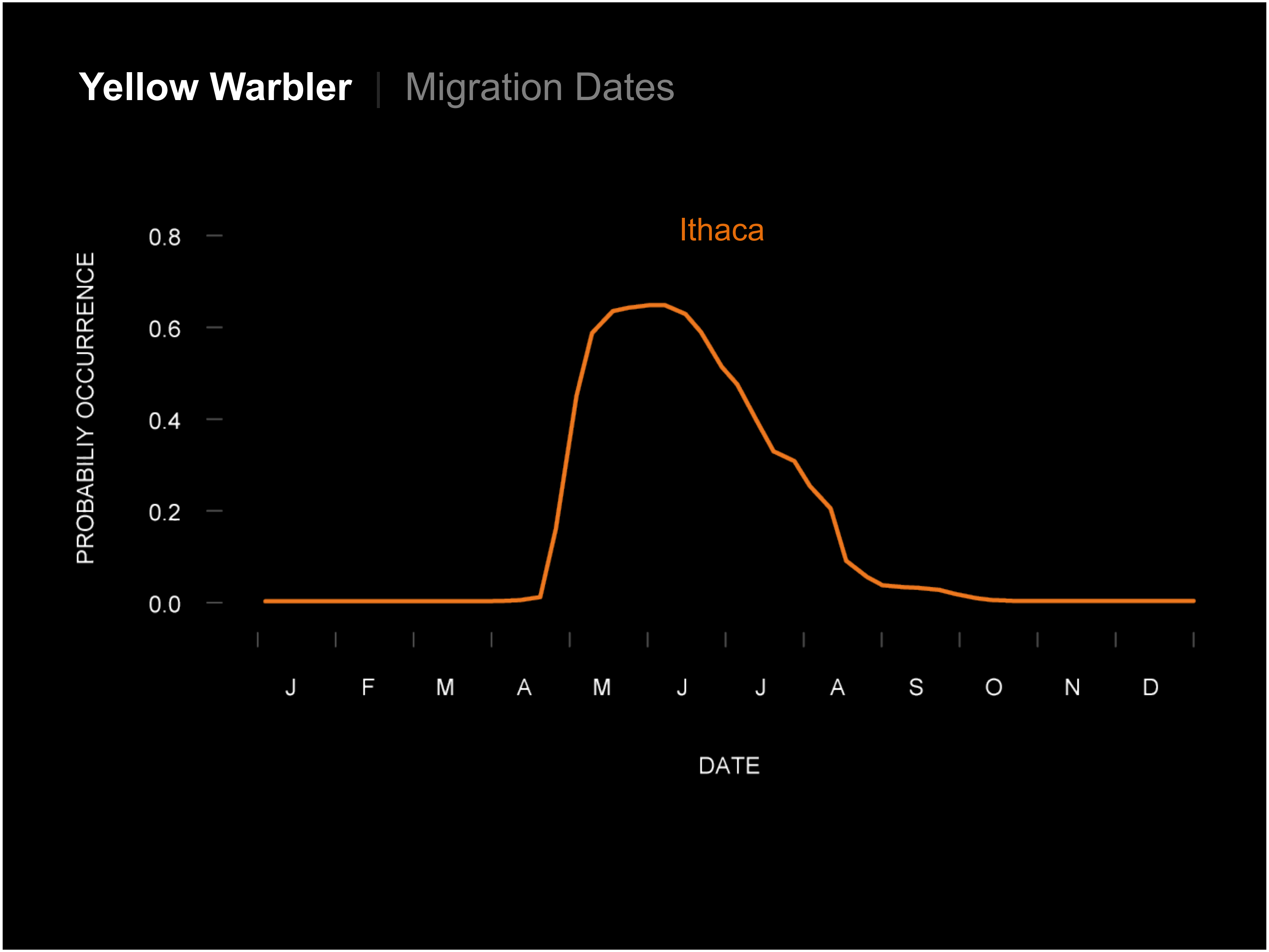
Relative Importance →

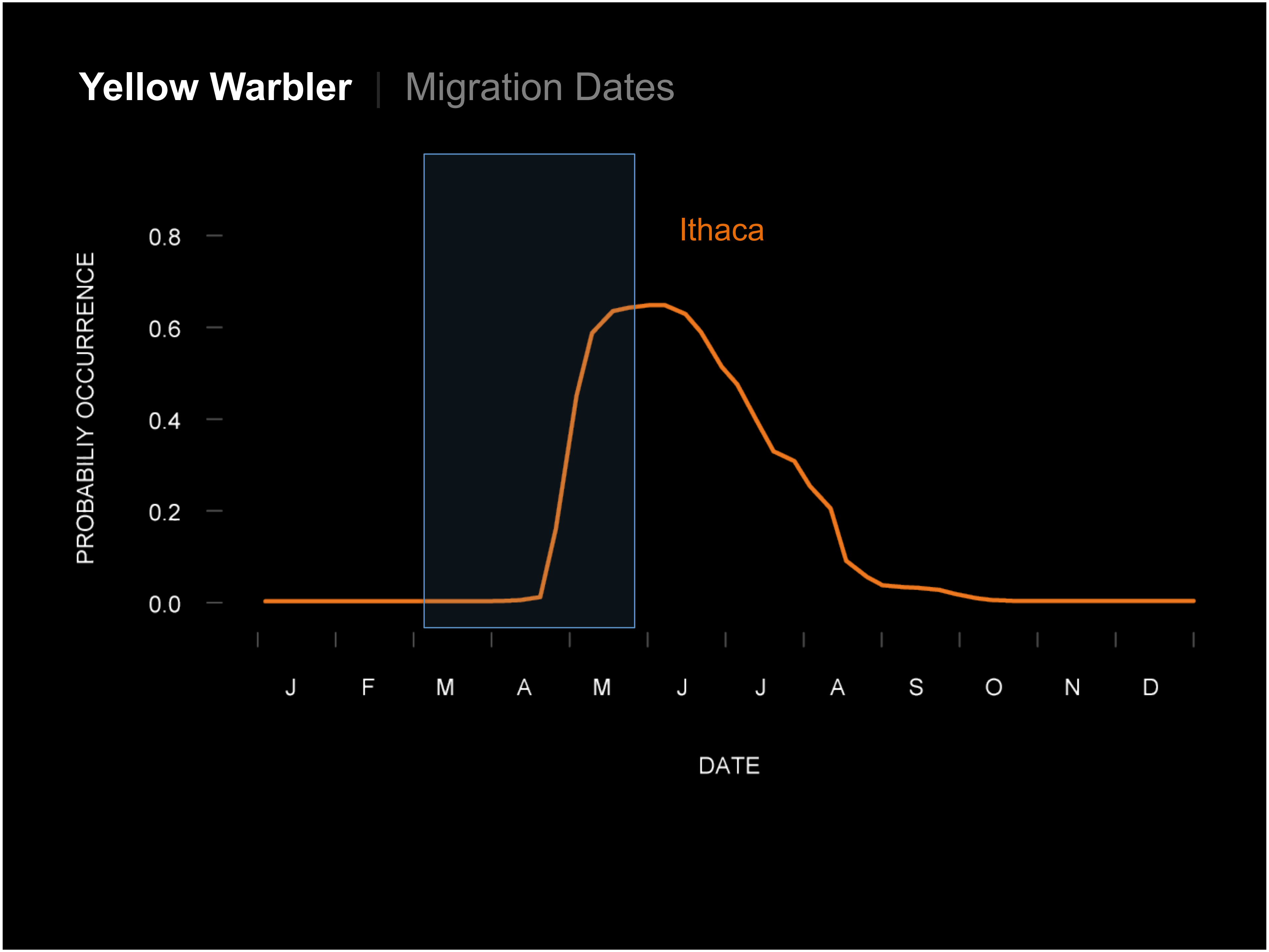

Indigo Bunting | Habitat Preference
July
Deciduous Forest (%)
Pasture (%)
PERCENT CHANGE OCCURRENCE
10
0
-10
-20
0
20
40
60
80
10
5
0
-5
0
10
20
30
September
Deciduous Forest (%)
Pasture (%)
PERCENT CHANGE OCCURRENCE
10
0
-10
-20
0
20
40
60
80
10
5
0
-5
0
10
20
30

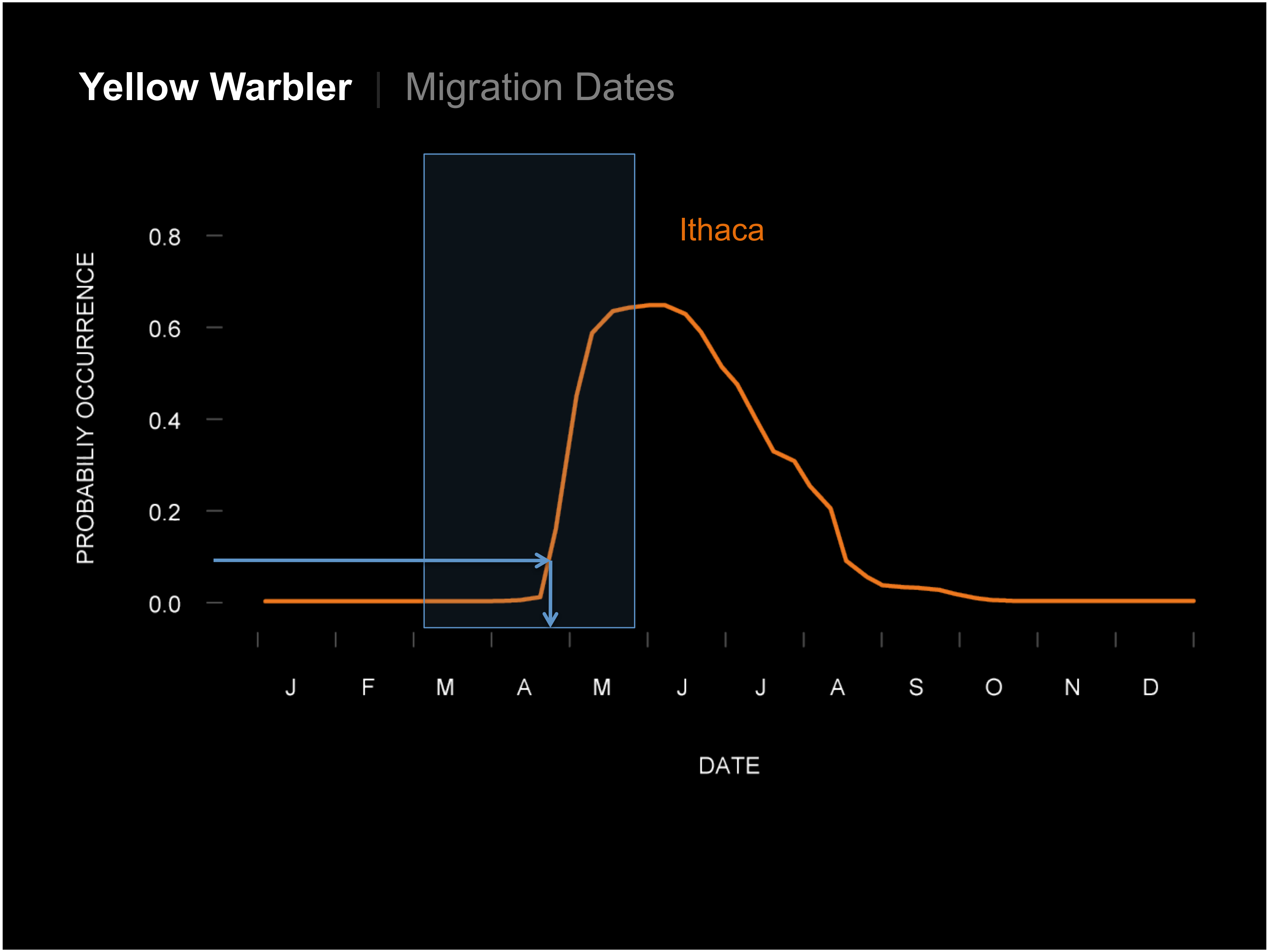

Chimney Swift
PERCENT CHANGE OCCURRENCE
4
0
−4
0
200
400
600
800
1000
HOUSING DENSITY (UNITS PER SQ. MILE)
Indigo Bunting
PERCENT CHANGE OCCURRENCE
4
0
−4
0
200
400
600
800
1000
HOUSING DENSITY (UNITS PER SQ. MILE)

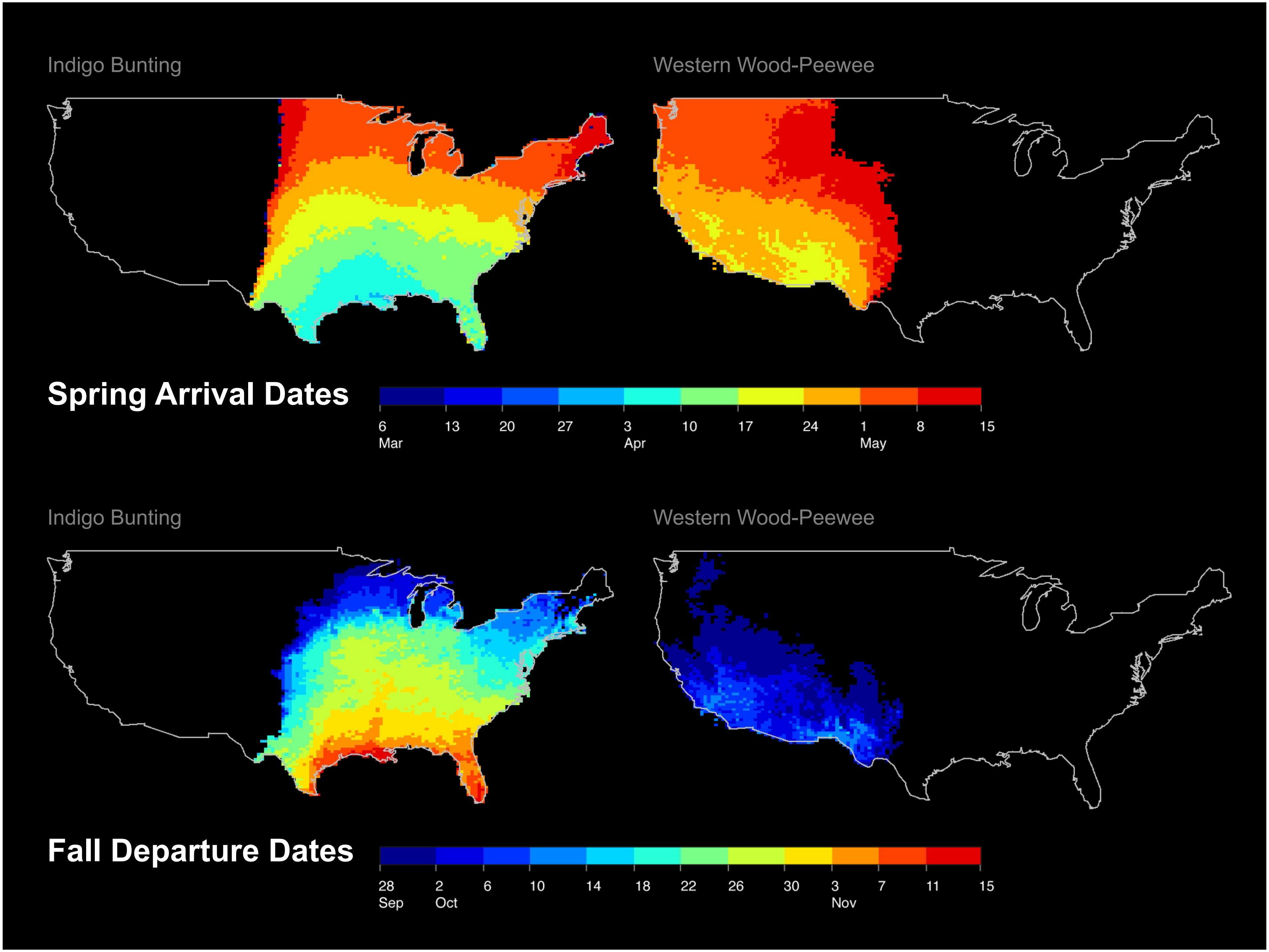


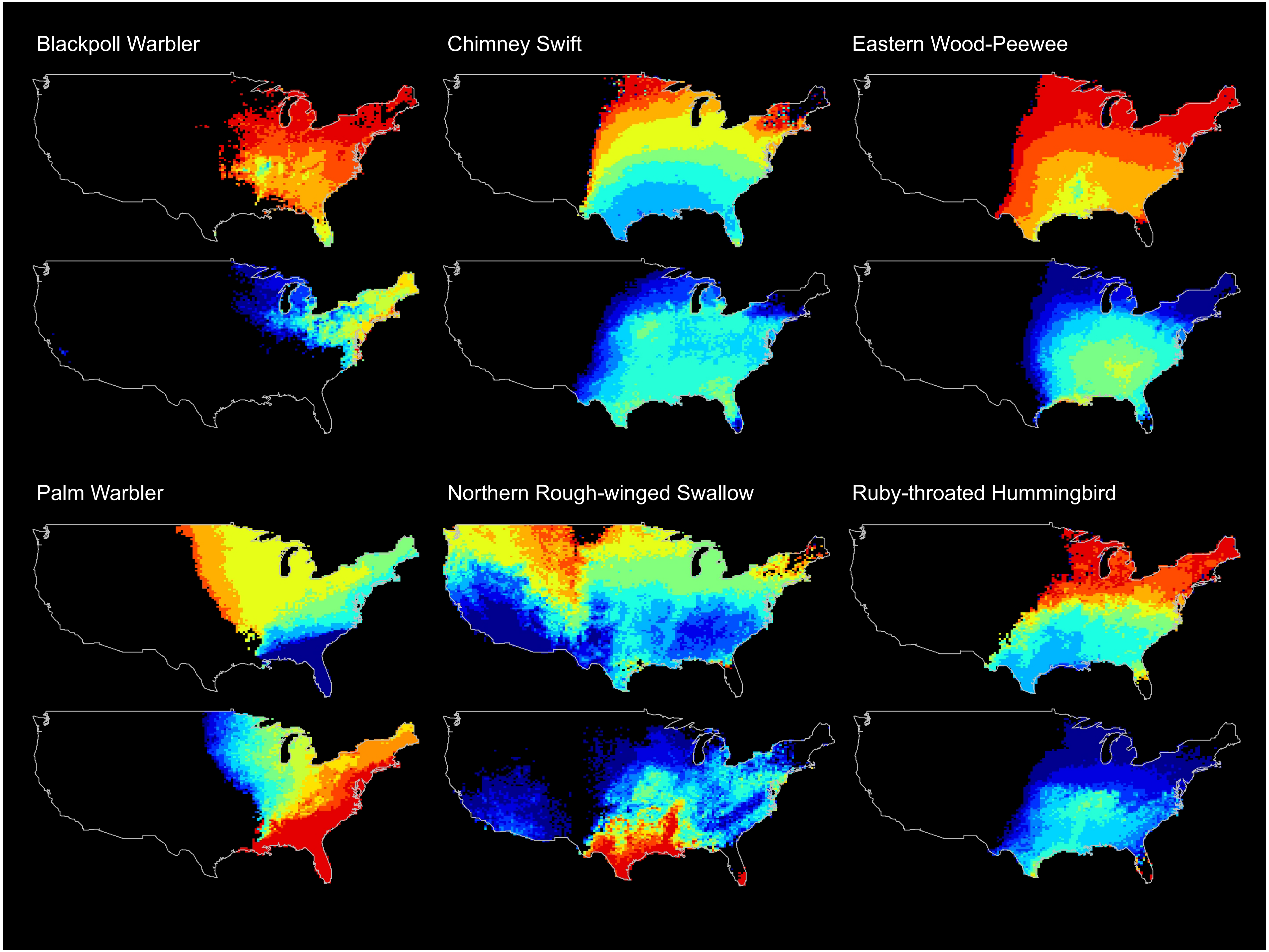
Yellow Warbler | Occurrence Trajectories
Ithaca
Denver
Los Angeles
PROBABILIY OCCURRENCE
0.8
0.6
0.4
0.2
0.0
J
F
M
A
M
J
J
A
S
O
N
D
DATE


Yellow Warbler | Migration Dates
Ithaca
PROBABILIY OCCURRENCE
0.8
0.6
0.4
0.2
0.0
J F M A M J J A S O N D
DATE


Yellow Warbler | Migration Dates
Ithaca
PROBABILIY OCCURRENCE
0.8
0.6
0.4
0.2
0.0
J F M A M J J A S O N D
DATE


Yellow Warbler
Migration Dates
Ithaca
PROBABILIY OCCURRENCE
0.8
0.6
0.4
0.2
0.0
J
F
M
A
M
J
J
A
S
O
N
D
DATE


Indigo Bunting
Western Wood-Peewee
Spring Arrival Dates
6 Mar
13
20
27
3 Apr
10
17
24
1 May
8
15
Indigo Bunting
Western Wood-Peewee
Fall Departure Dates
28 Sep
2 Oct
6
10
14
18
22
26
30
3 Nov
7
11
15


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

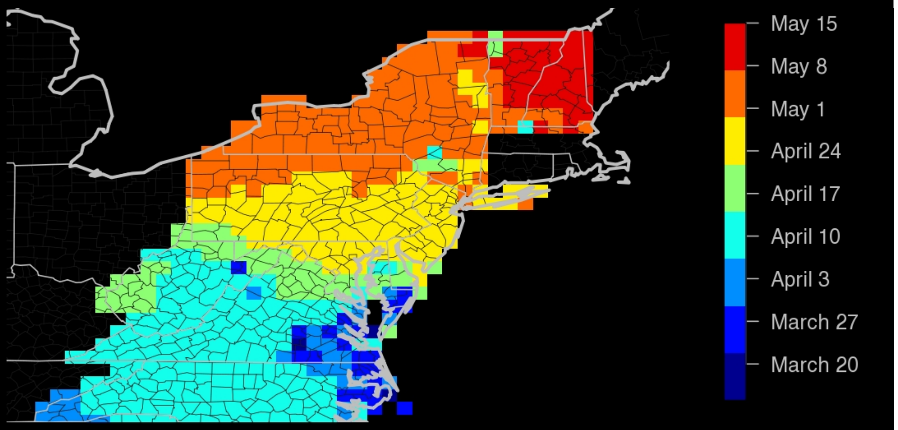
Environmental Cues of Migration

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

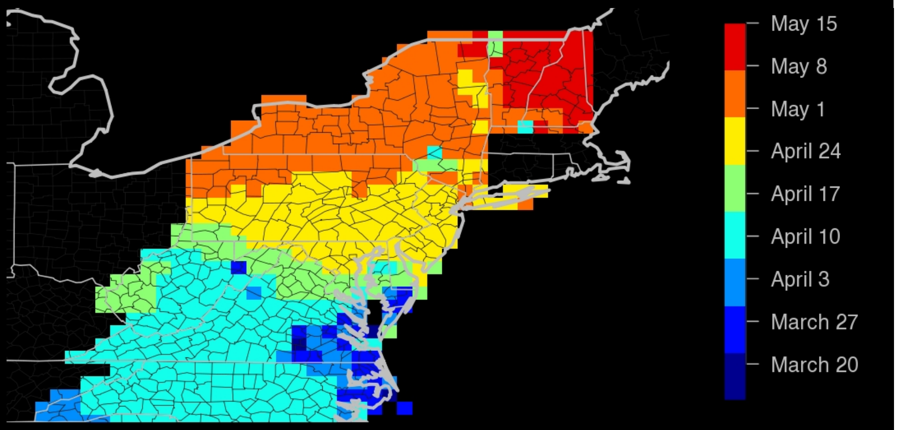
Questions

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

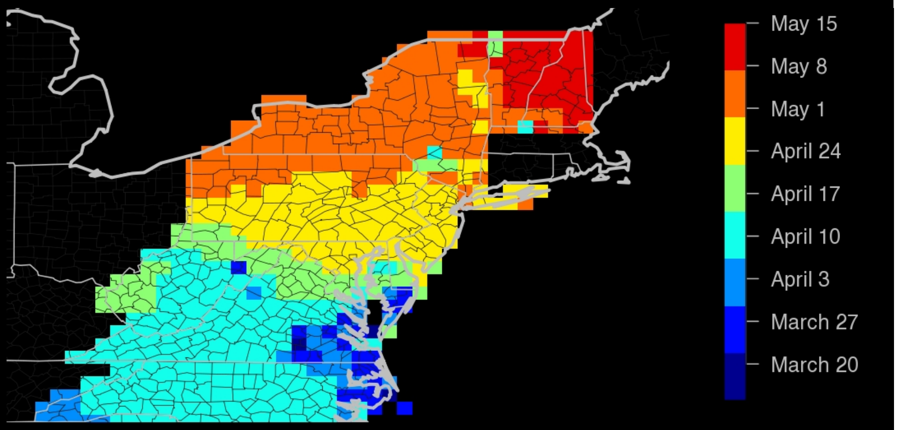
Questions

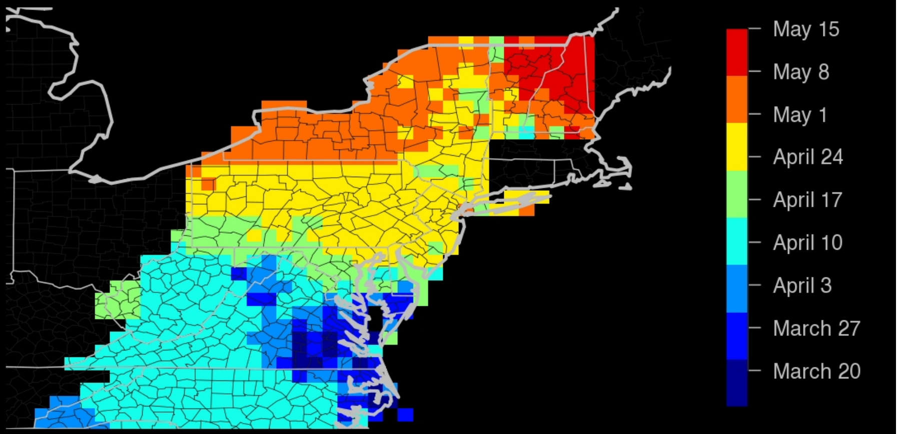
1. Is Red-eyed Vireo migration timing associated with NDVI?
2. If so, how is migration timing, direction, and speed affected by NDVI?

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

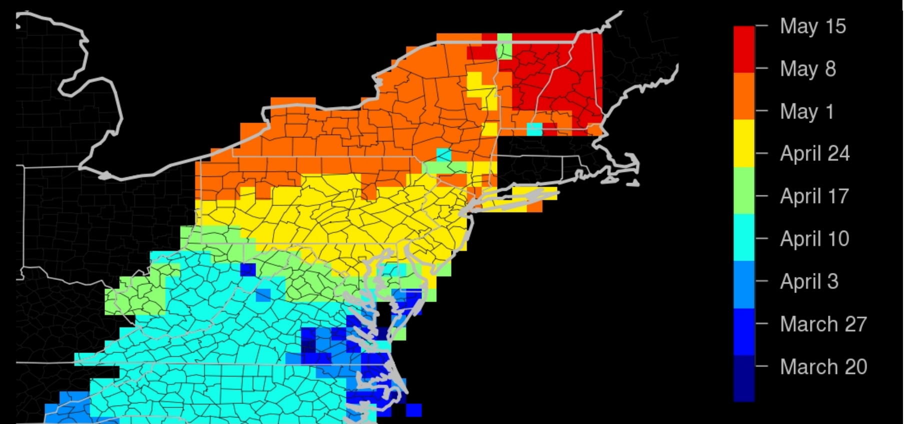
Questions

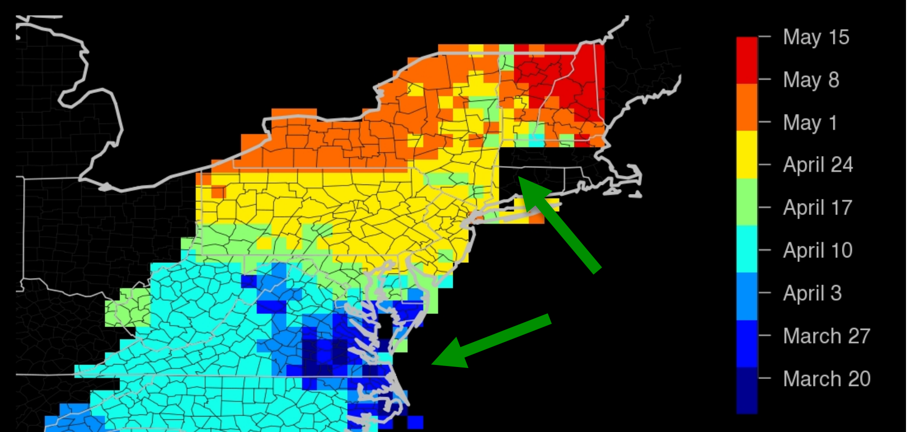
1. Is Red-eyed Vireo migration timing associated with NDVI?
2. If so, how is migration timing, direction, and speed affected by NDVI?

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