DARK ECOLOGY

Unraveling Mysteries of Bird Migration using Weather Radar and Machine Learning

Subhransu Maji

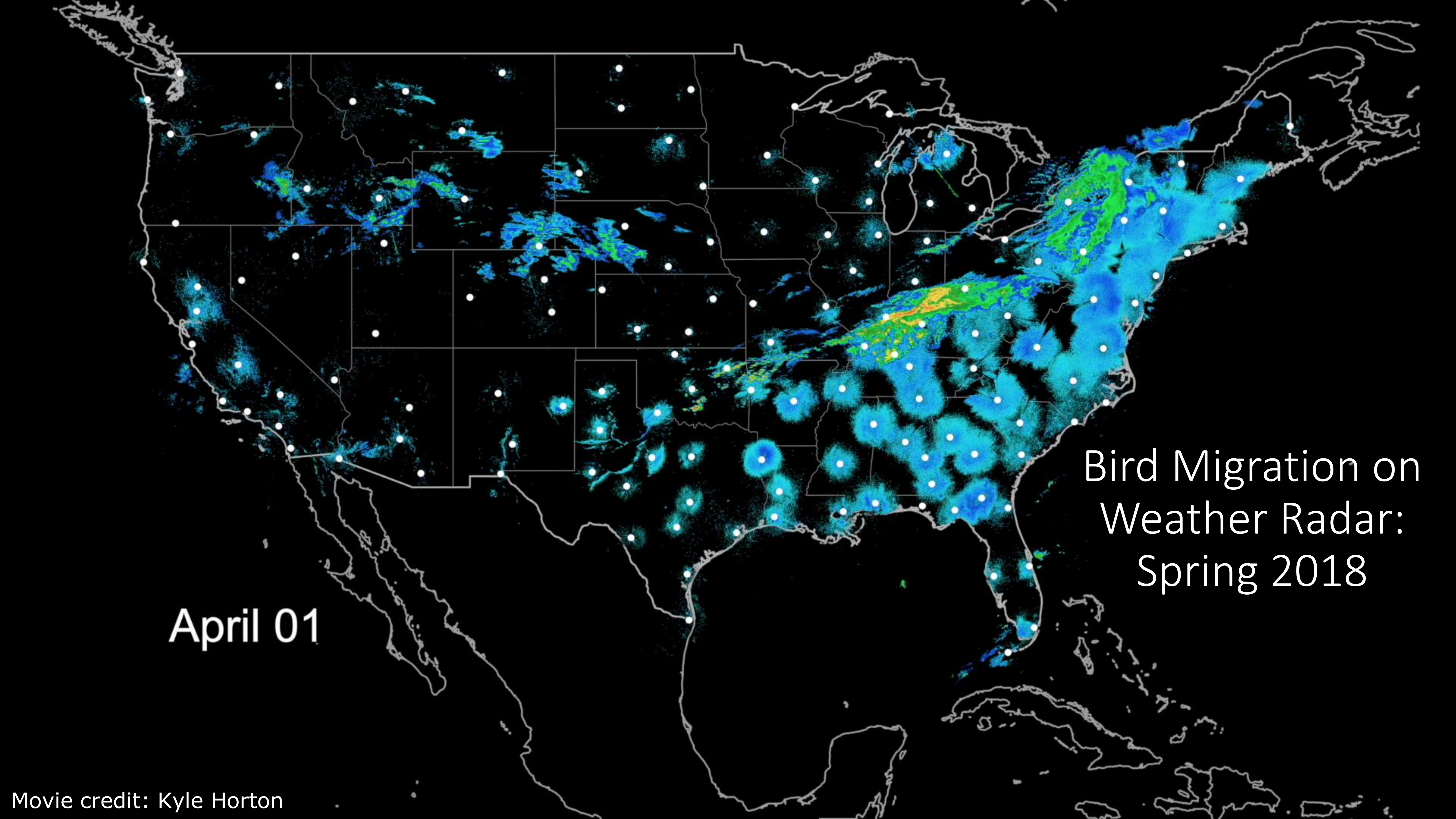
University of Massachusetts, Amherst

UMassAmherst | College of Information & Computer Sciences

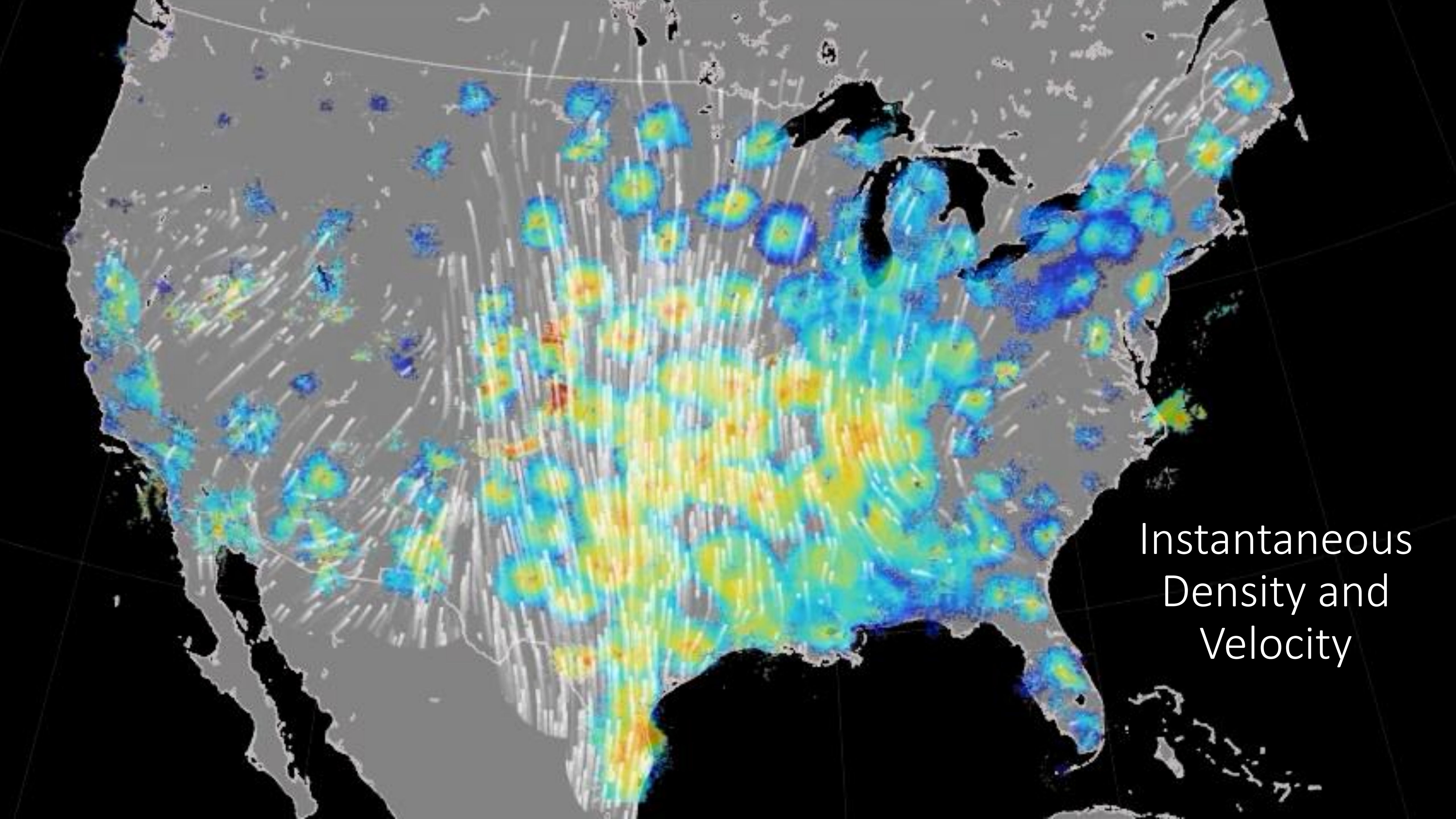
TheCornellLab

NSF

Bird Migration on Weather Radar: Spring 2018



Movie credit: Kyle Horton



Instantaneous Density and Velocity

Dark Ecology Project

- **Goal:** AI to unlock biological information in 25-year US weather radar archive
  - Basic science
  - Conserve biodiversity, ecosystems
  - Climate change

- **Challenges:** big data, not previously automated

UMass Amherst | College of Information & Computer Sciences

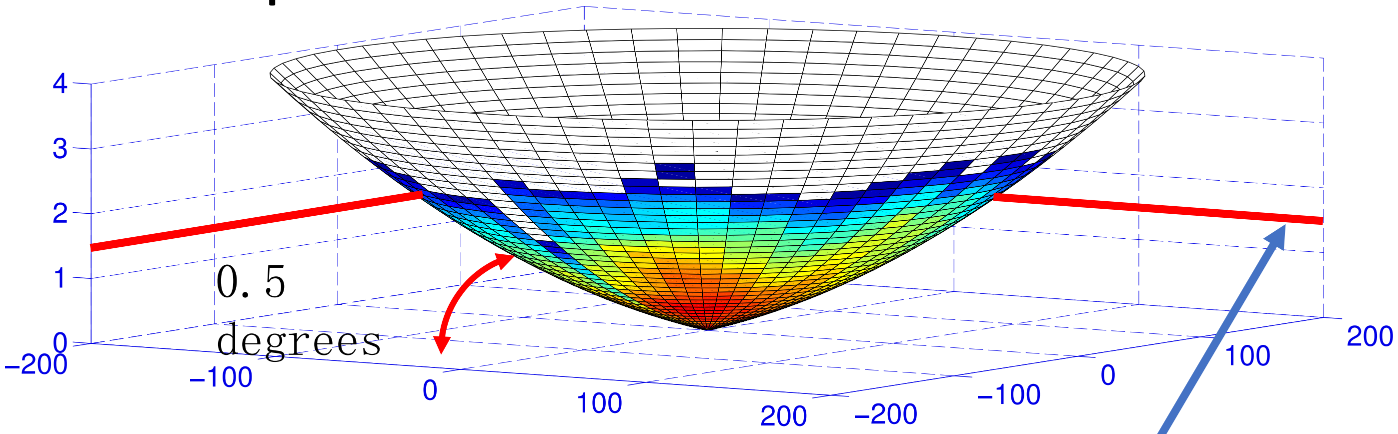
The Cornell Lab of Ornithology

NSF

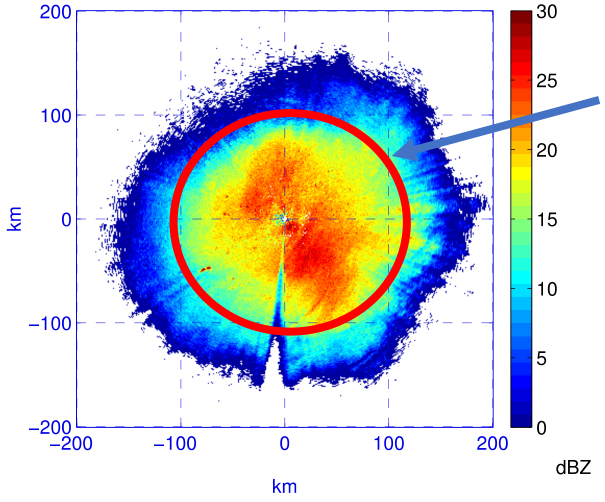
Radar Background: Reflectivity

- Distance–height relationship

Scan pattern



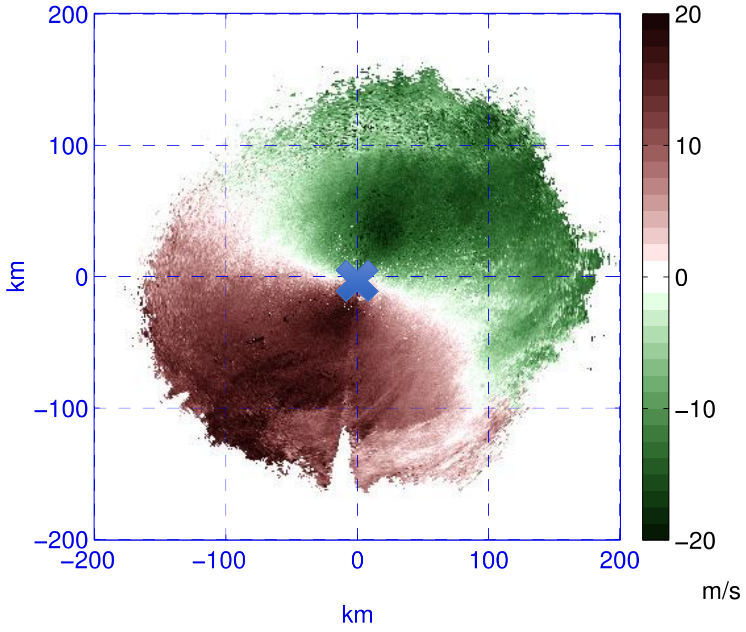
2D image
(top view)

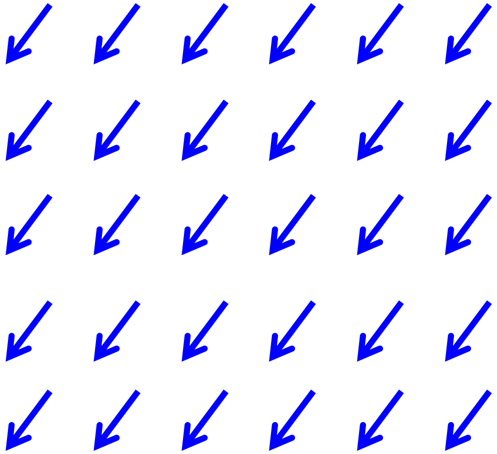


100km range
≈ 1.5km
height

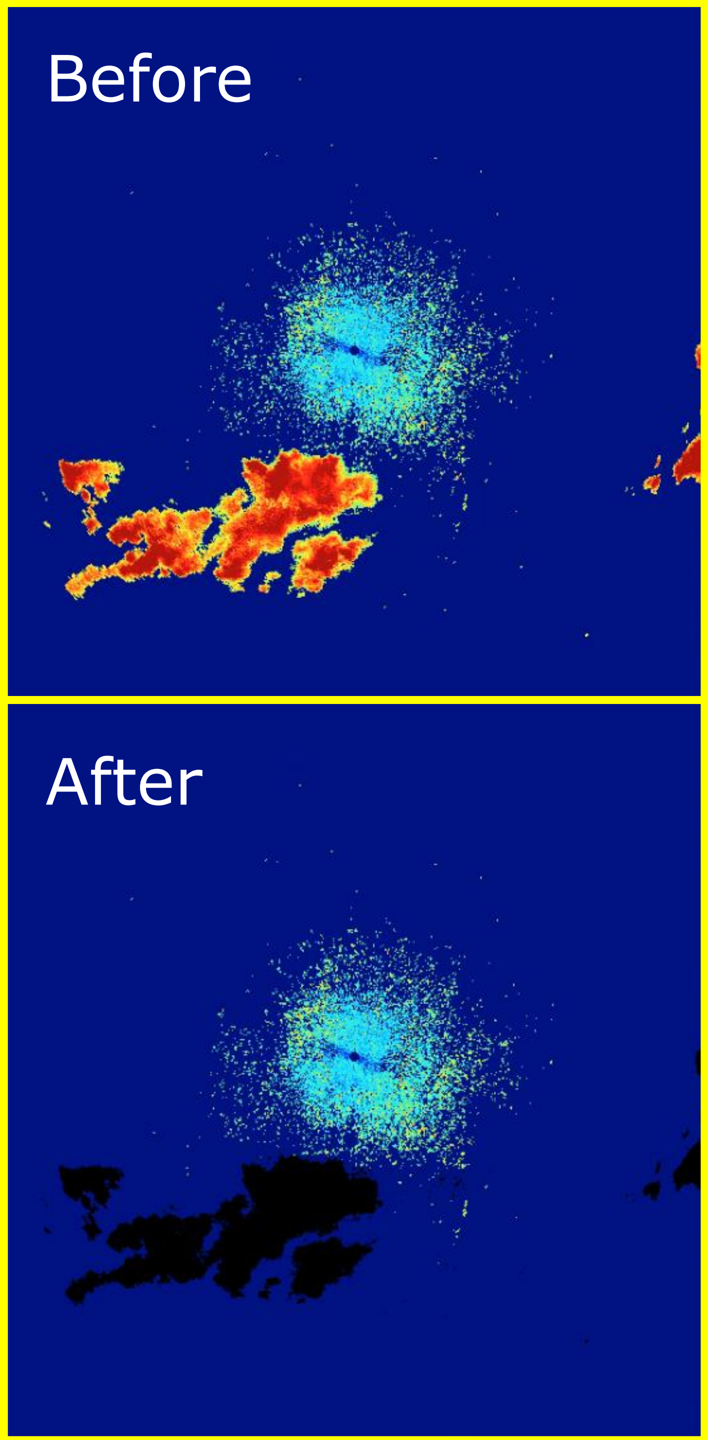
Radar Background: Radial Velocity

- Radial velocity: speed at which targets approach or depart the radar station (Doppler shift)



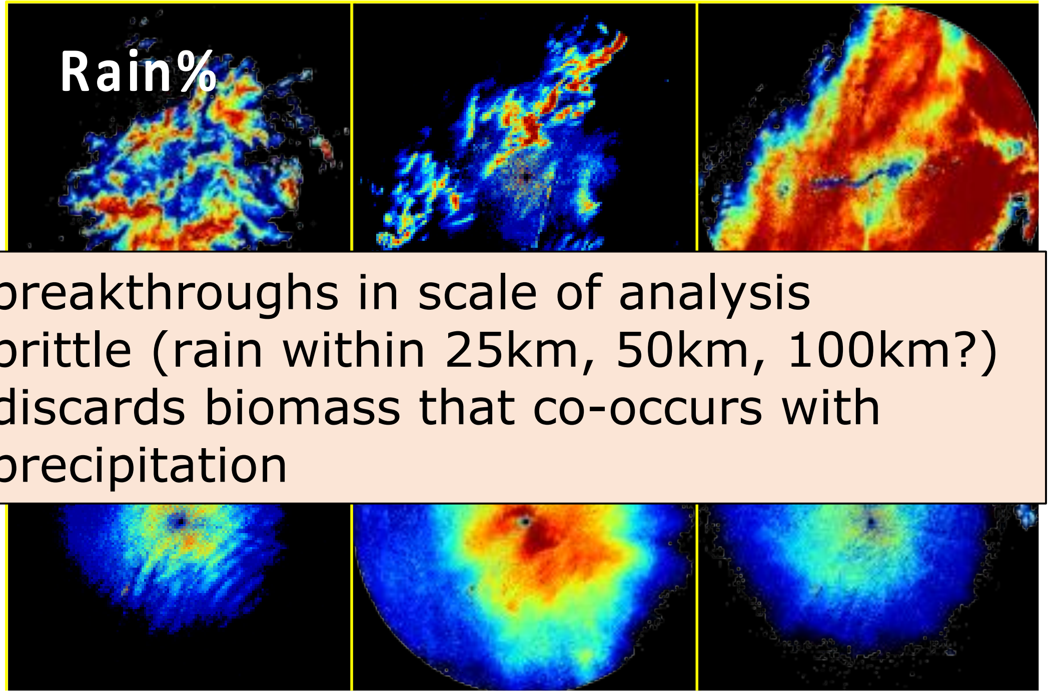


Removing weather contamination



Screening Rain

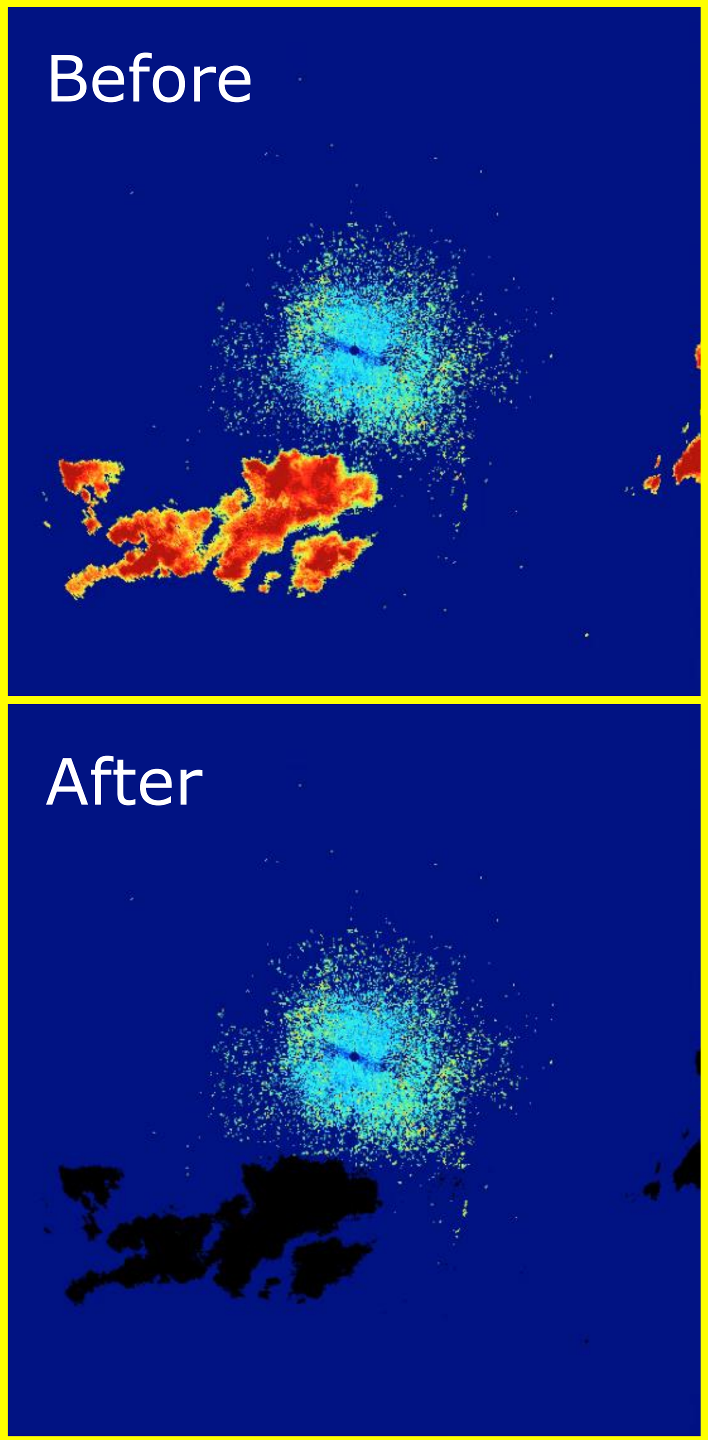
- Manual: most US studies until recently
- "Direct" algorithms for some radar systems
  - Dual-pol: US post 2013
  - European C-band [Dokter et al. 2011]
- ML for *whole-scan classification*
  - Images [Roy-Chowdhury et al. 2016]
  - Vertical profiles [Van-Doren and Horton, 2018]



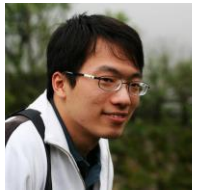
- breakthroughs in scale of analysis
- brittle (rain within 25km, 50km, 100km?)
- discards biomass that co-occurs with precipitation

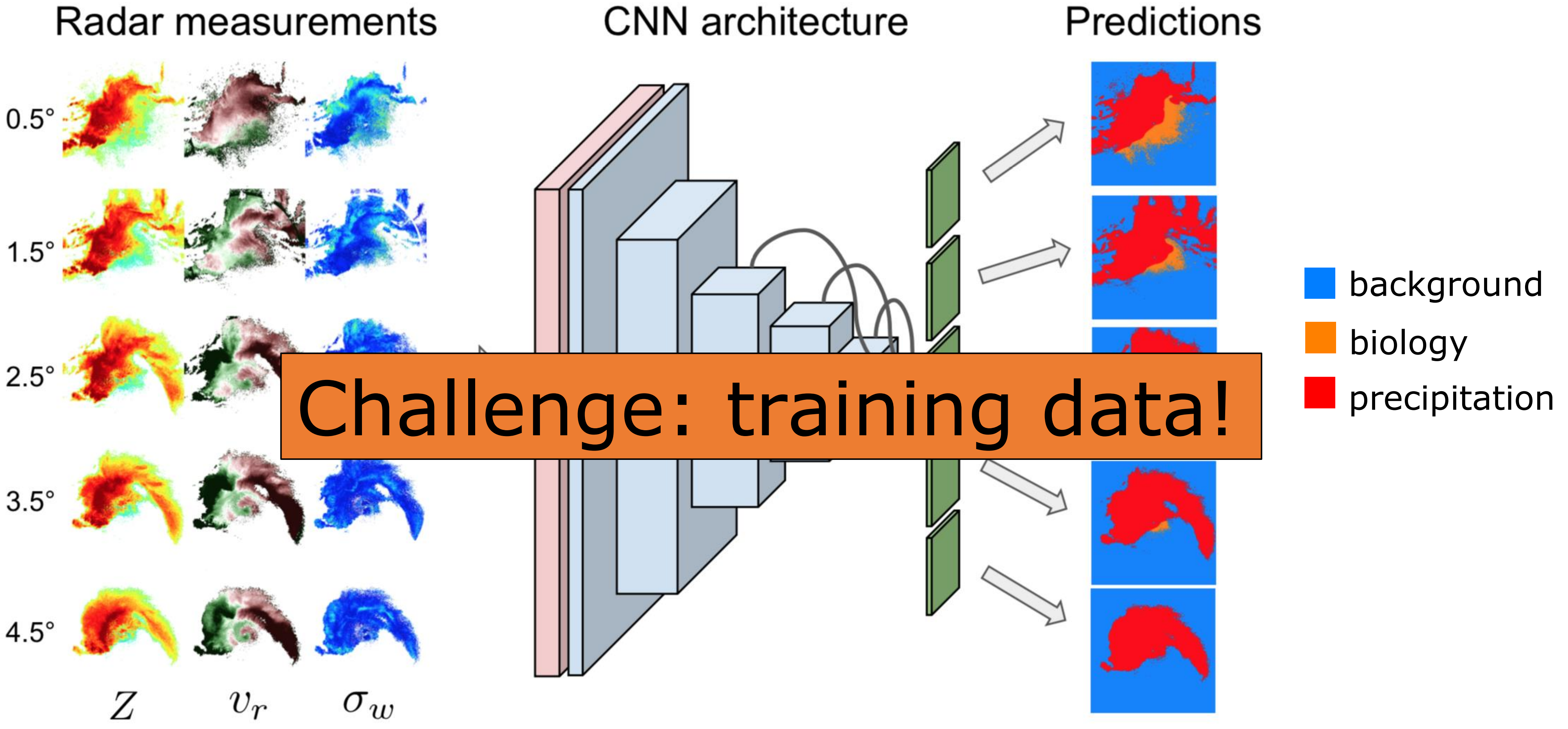
Rain segmentation

- Flexible to many downstream analyses
- Spatially-explicit
- Retains biomass that co-occurs with precipitation (~19%)



MistNet — CNN for Weather Radar





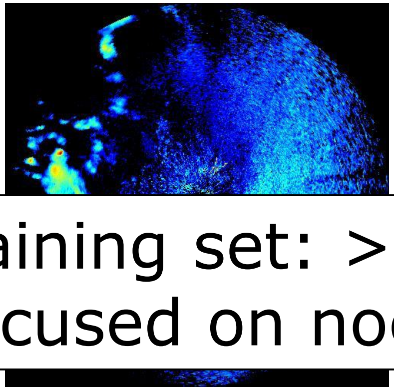
Challenge: training data!

MistNet: Measuring historical bird migration in the US using archived weather radar data and convolutional neural networks, Tsung-Yu Lin et al., Methods in Ecology and Evolution, August 2019

Weak Supervision

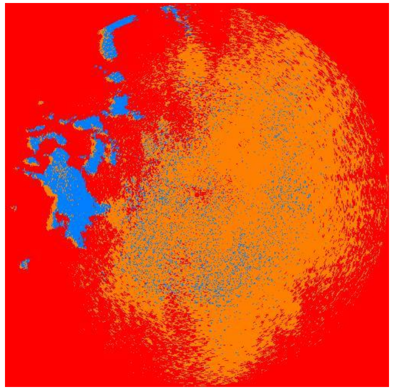
2013 Dual polarization upgrade: abundant noisy labels

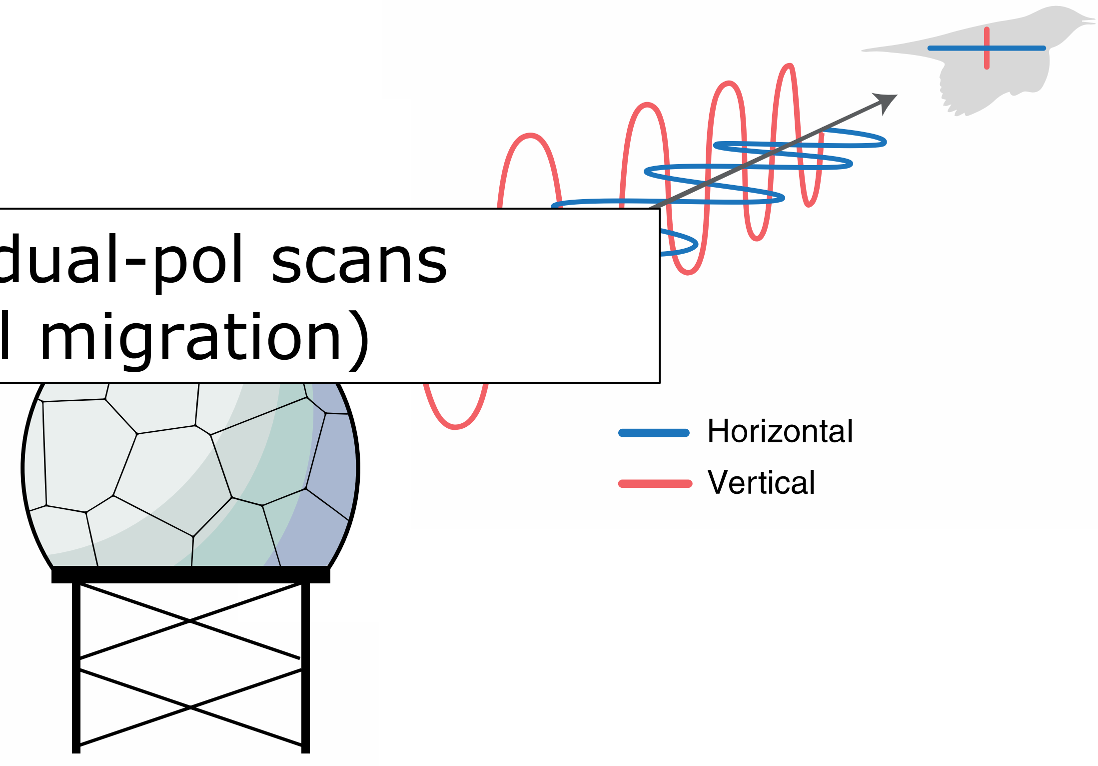
Legacy data



Training set: >200K dual-pol scans (focused on nocturnal migration)

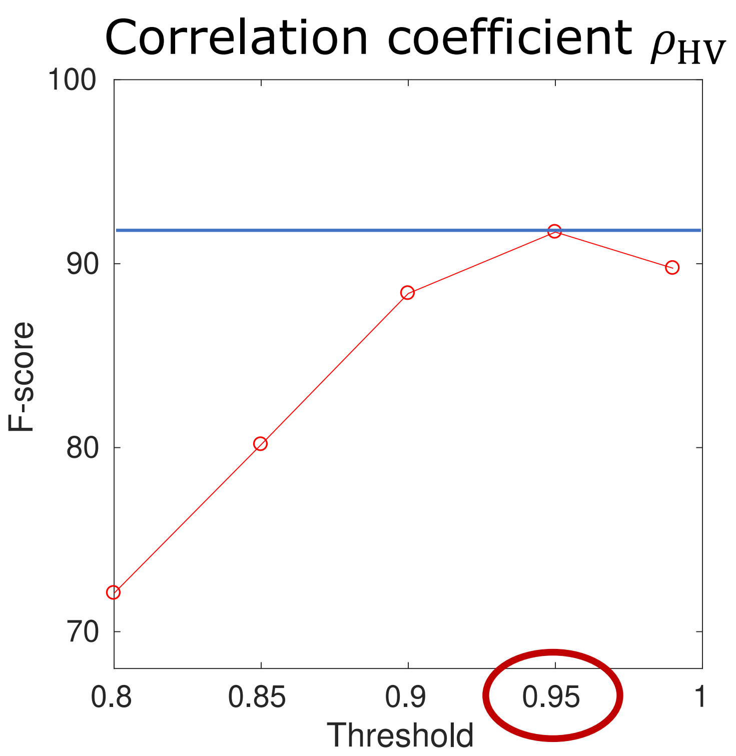
Dual-pol: noisy labels

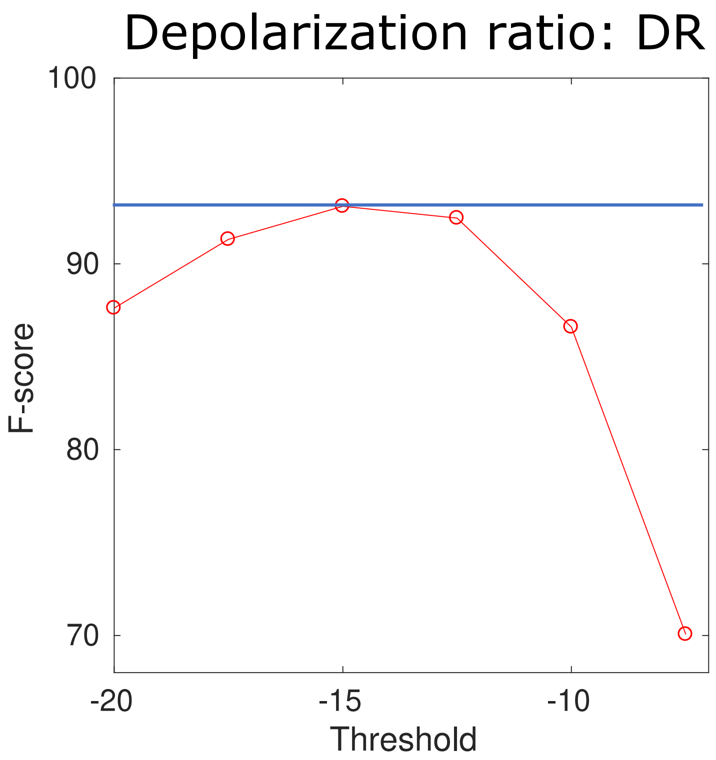




[Figure: Stepanian et al. Ecosphere 2016]

Quality of Dual-Pol Thresholding Rules



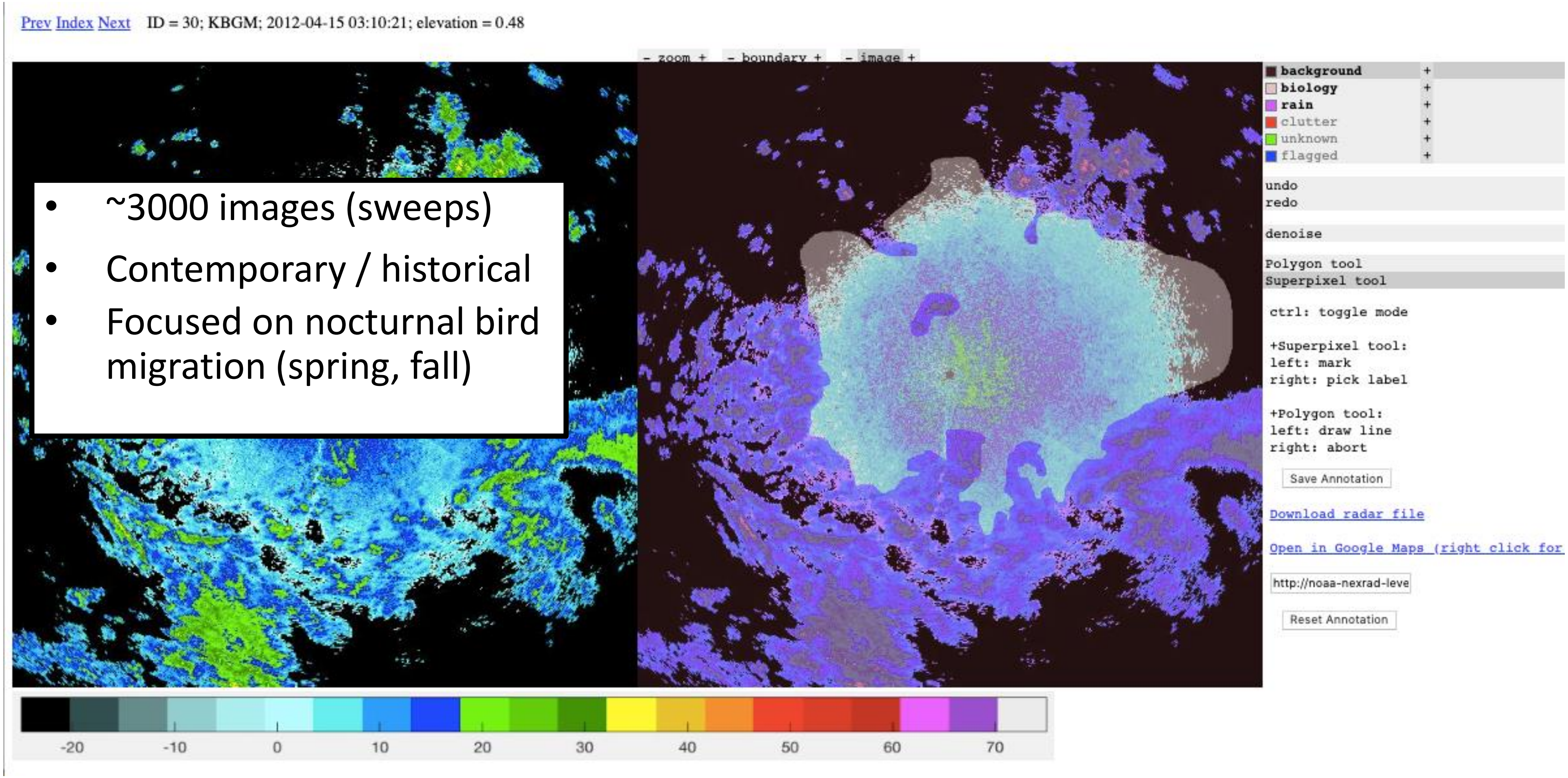


$$\mathrm{DR} = 10\log_{10}\left(\frac{Z_{\mathrm{DR}} + 1 - 2Z_{\mathrm{DR}}^{1/2}\rho_{HV}}{Z_{\mathrm{DR}} + 1 + 2Z_{\mathrm{DR}}^{1/2}\rho_{\mathrm{HV}}}\right)$$

[Kilambi et al., 2018]

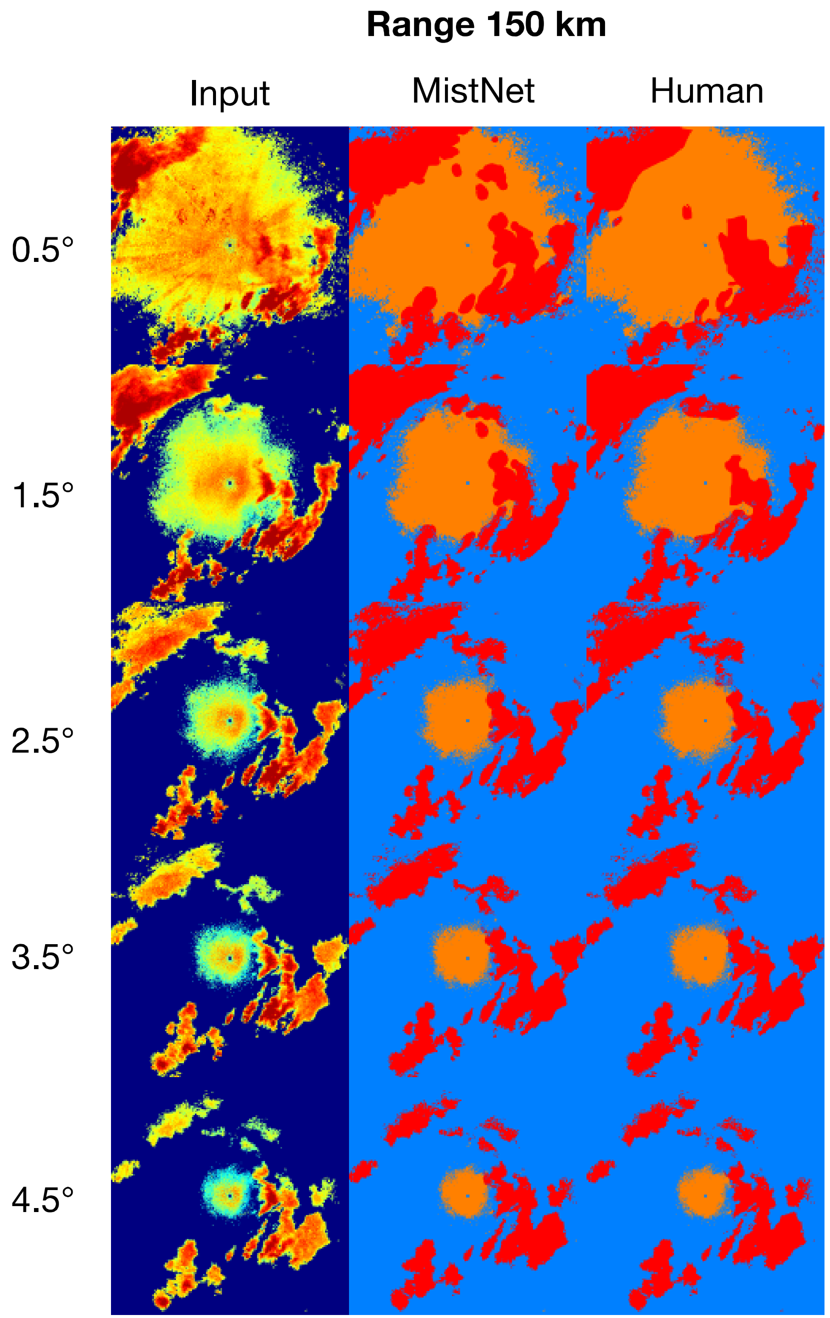
Evaluation: Human Labels

- ~3000 images (sweeps)
- Contemporary / historical
- Focused on nocturnal bird migration (spring, fall)



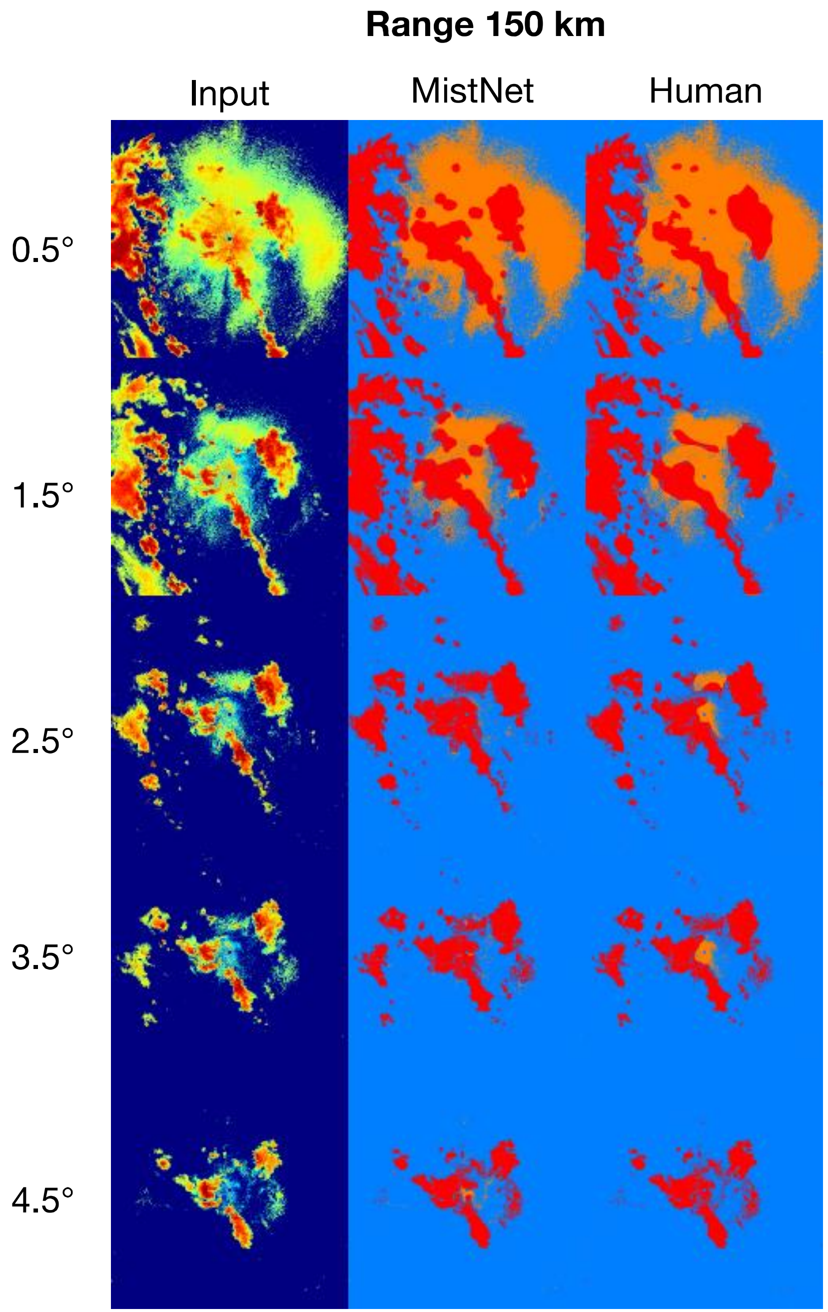
Qualitative Results

Mobile, AL, Sep 1 2007, 3:10 UTC



Qualitative Results

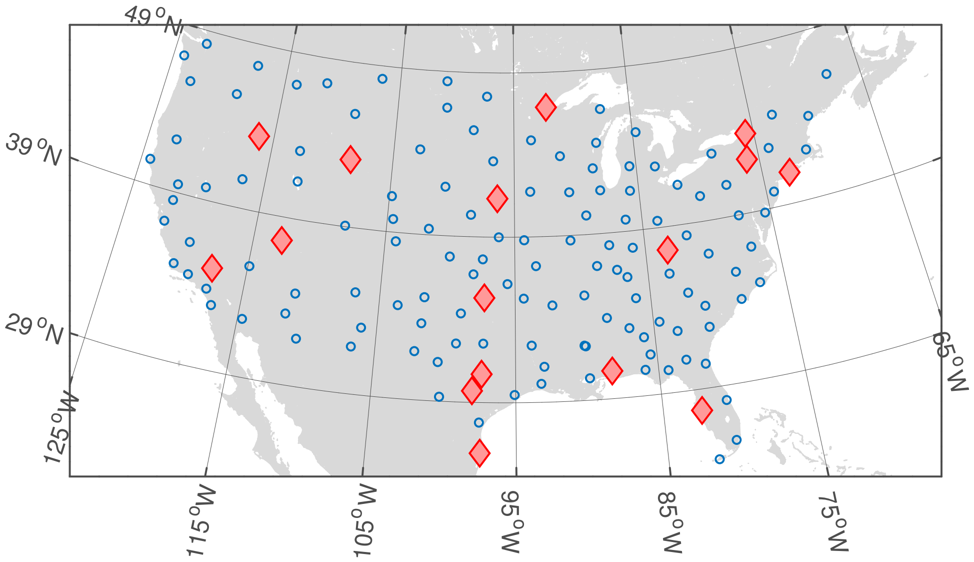
Binghamton, NY, Oct 1, 2014, 2:18 UTC



Quantitative Results

Method	Precision	Recall	F-score
$\rho_{HV} > 0.95$	90.1	93.4	91.7
$DR < -15$	89.0	96.6	93.1
MistNet	99.1	96.7	97.9

Contemporary (2017)



Historical
(2 stations 1995–2017)

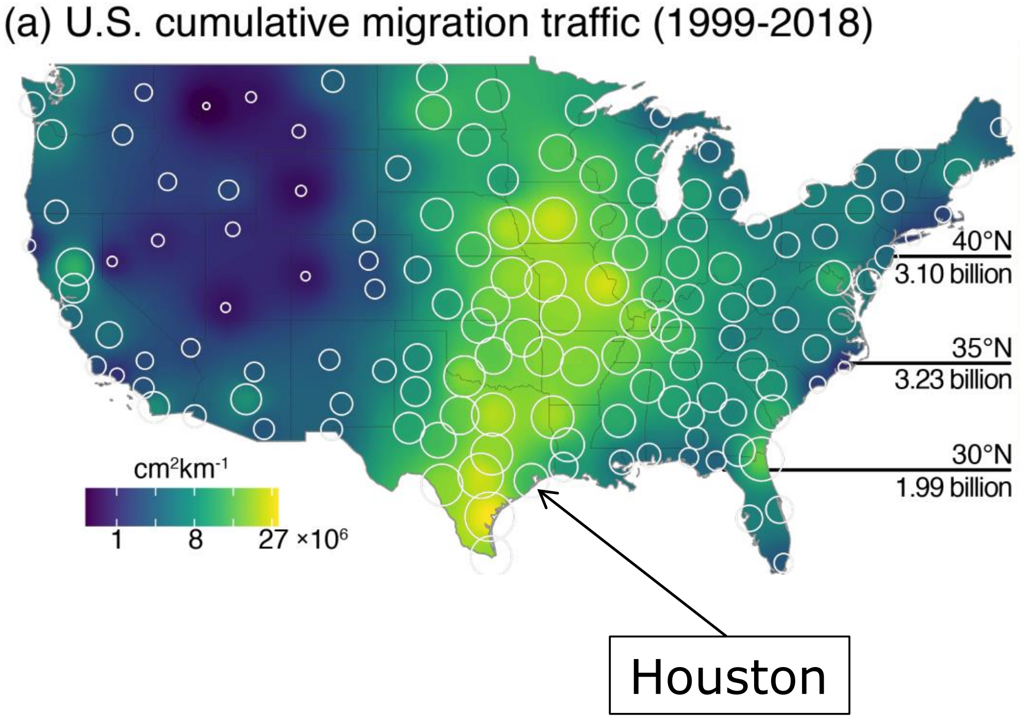
Precision	Recall
98.7	95.9

Preserve 95.9% of biomass

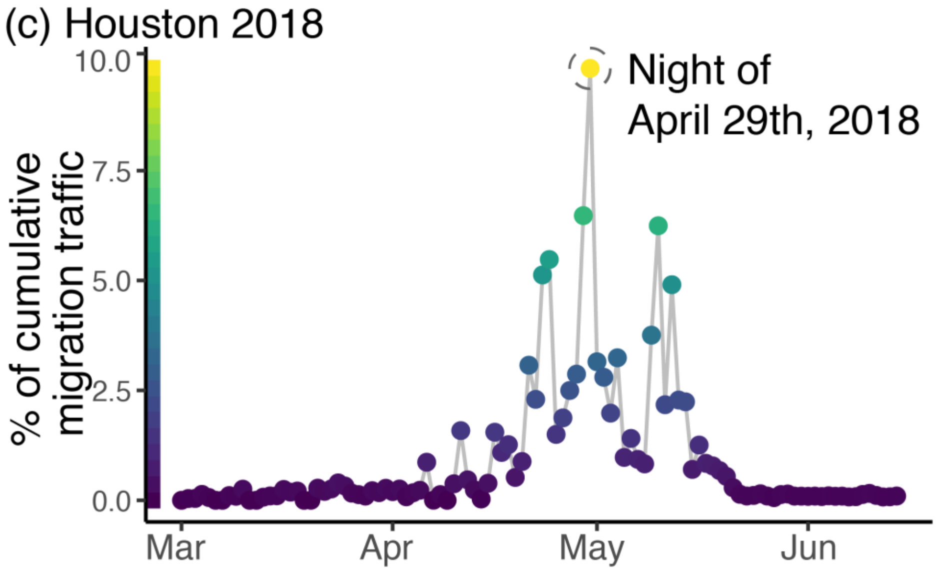
98.7% of what we keep is biology

Case Studies

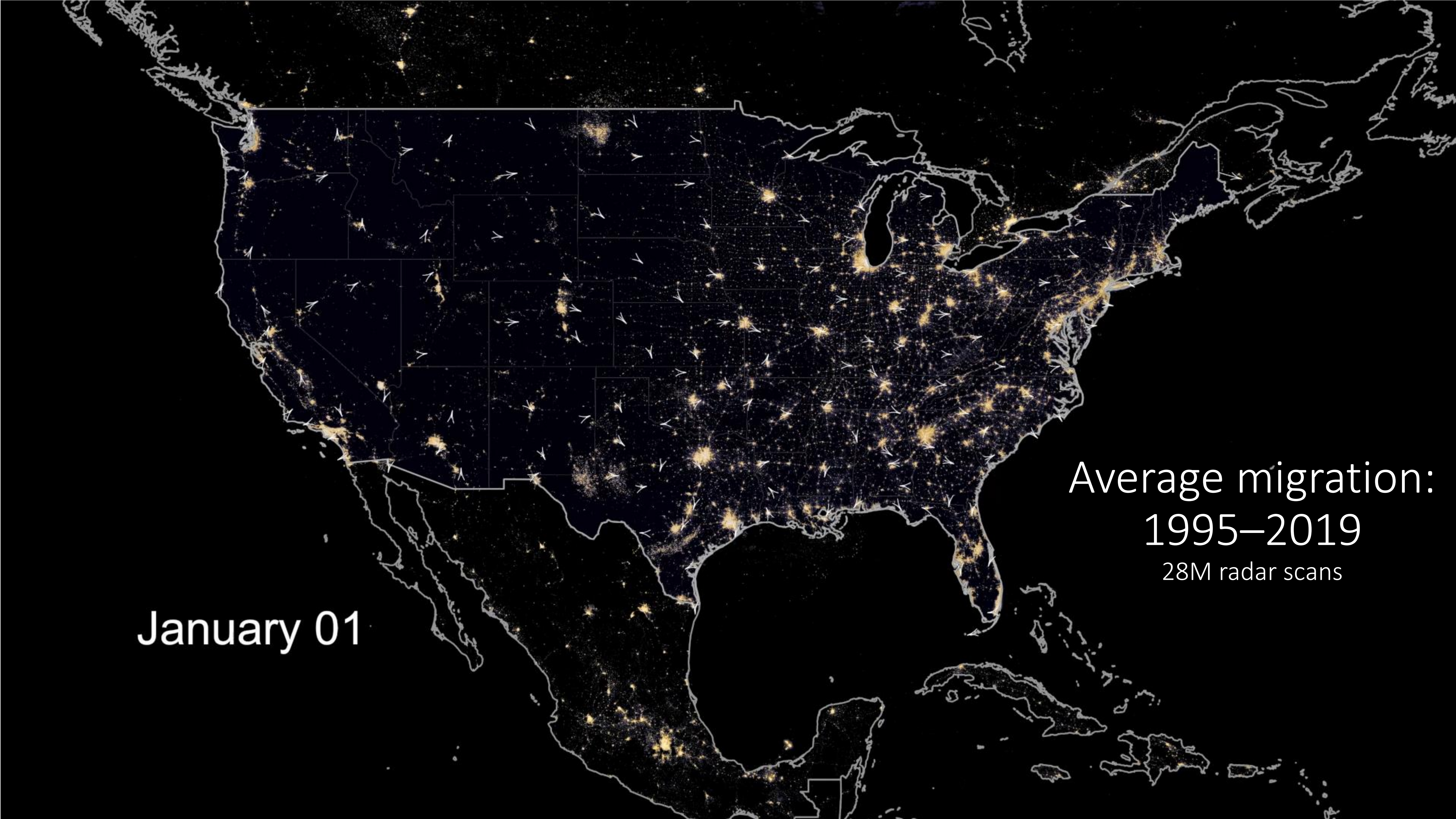
13 million scans, 1.2 years compute time



Spatial mapping



Nightly variation


January 01
Average migration:
1995–2019
28M radar scans

MistNet: Outlook

- Code available: https://github.com/darkecology/wsrlib

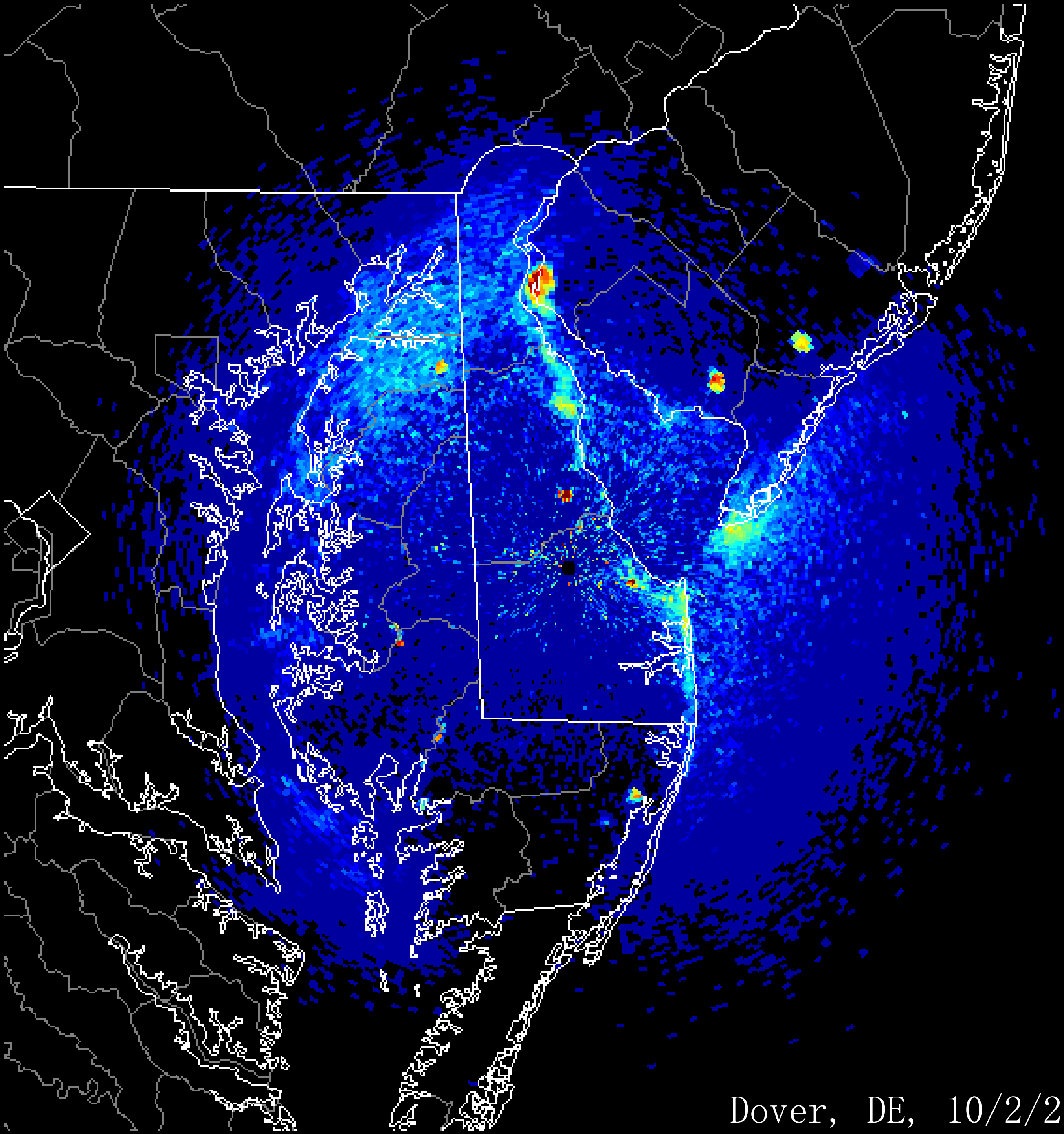
- Coming soon: R, Python, versions
- European version (w/ Bart Hoekstra @ Univ. Amsterdam)
- Dataset: vertical profiles of biomass for entire US radar archive (>200 million scans; >90% complete)
- Spatially-explicit analyses

Detecting and Tracking Swallow Roosts

Migration behavior of a single species

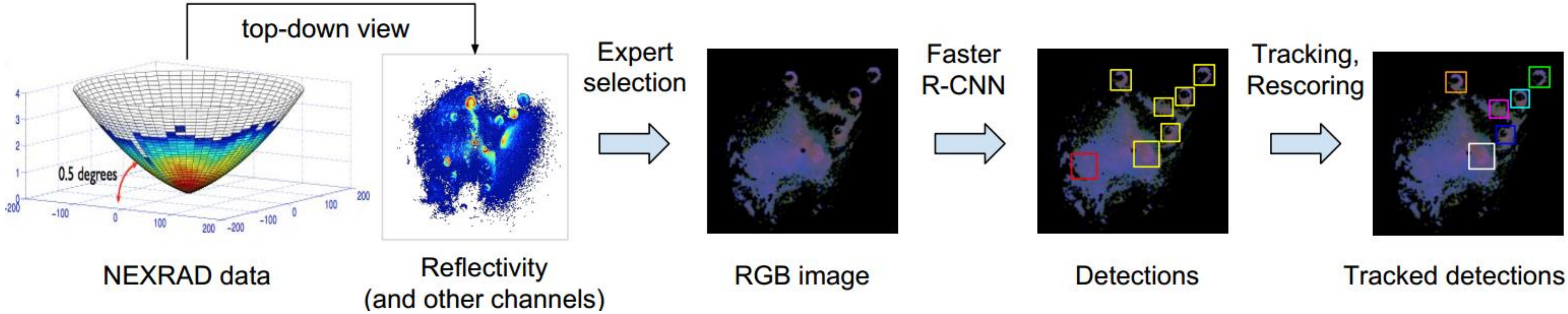
[Cheng et al., 20XX]

Communal Roosts



Dover, DE, 10/2/2010@6:52AM

Approach: Detect and Track

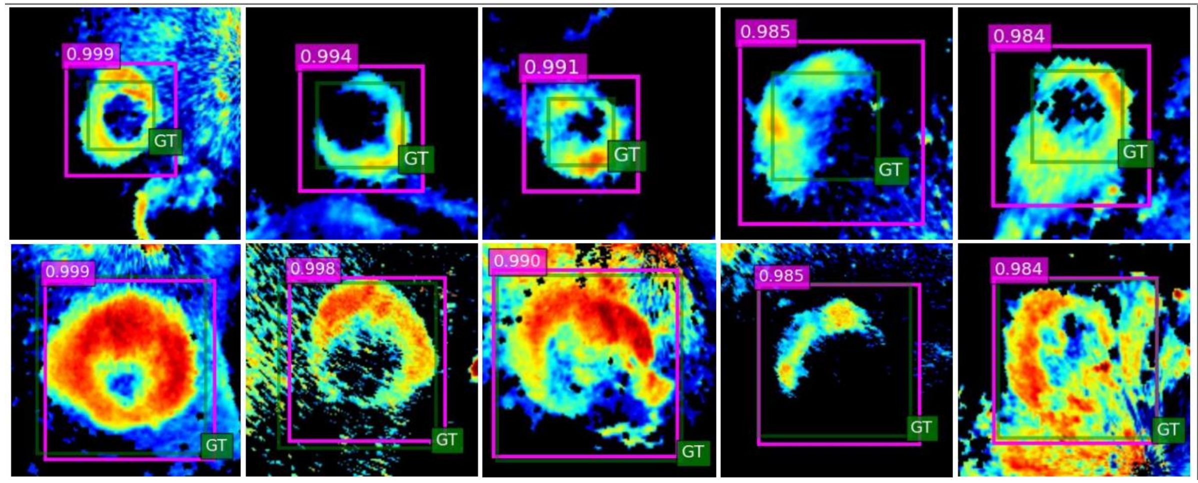


- **Faster R-CNN for radar image**
  - ImageNet pre-training is useful
  - Deeper networks are better

Model	w/ ImageNet pretraining	w/o ImageNet pretraining
VGG-M	**41.0**	34.8
Shallow VGG-M	37.7	33.1

Challenge: Variable annotation styles

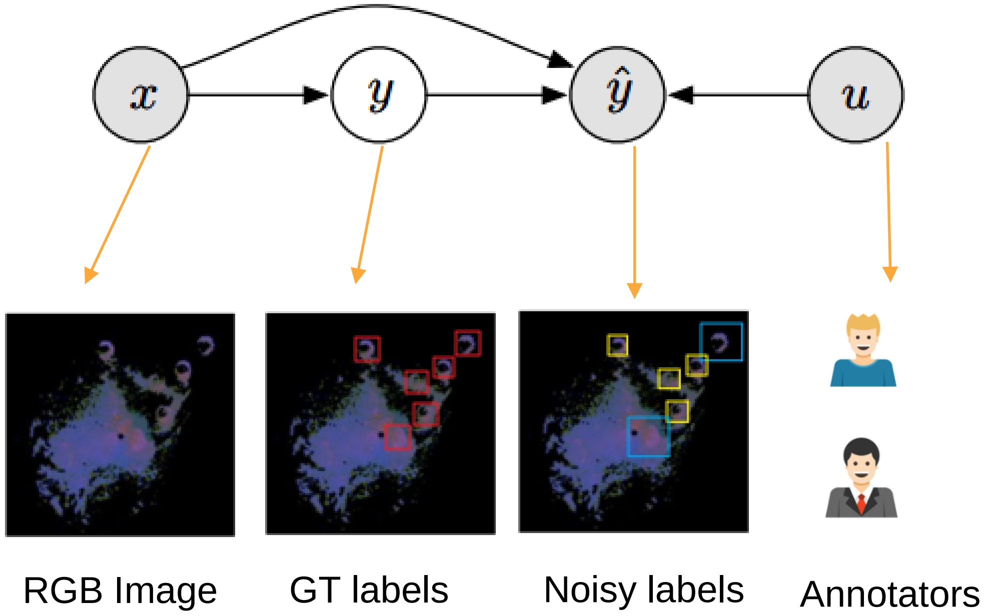
Station	Size of trainset	Faster RCNN (mAP)
KDOX	902	9.1



Noisy roost annotations lead to inaccurate evaluation

- **Challenge**: **Roost labels are abundant (more than 60K) but very noisy:** Considerable labeling variability, *much of which is specific to individual users*.→ Inaccurate evaluation and potentially hurt training.

An EM approach for learning with noisy annotations

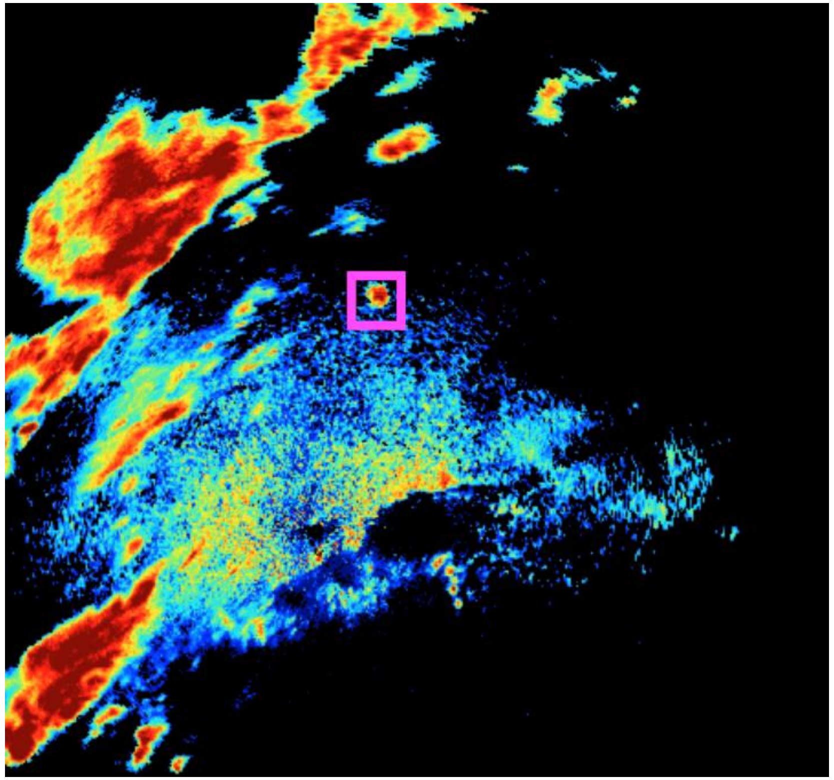


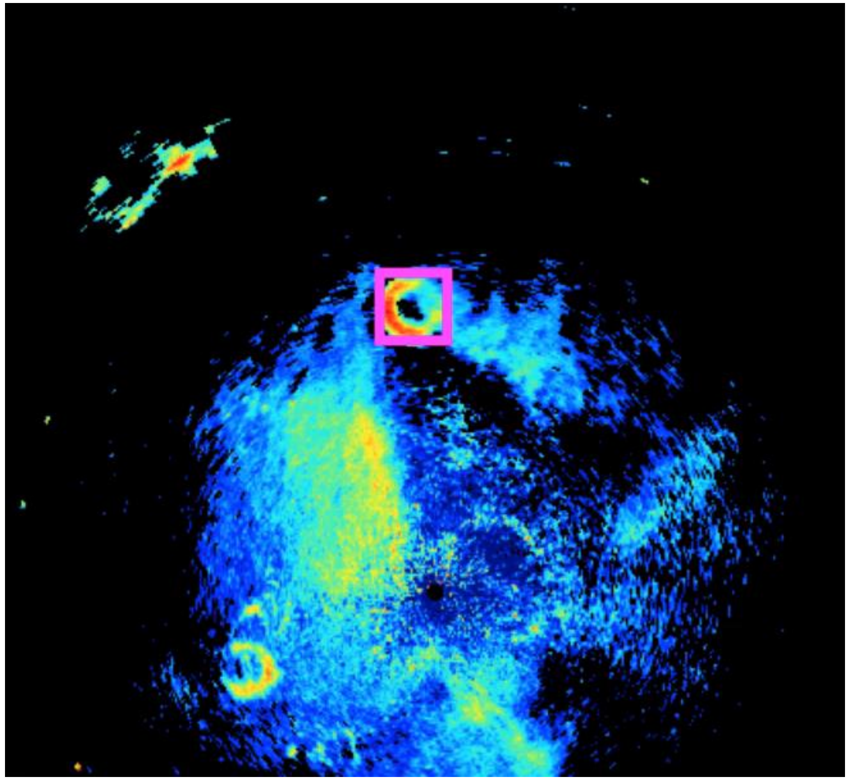
$p_\theta(y|x)$ is the *detection model*.

$p_\beta(\hat{y} \mid x, y, u)$ is the *forward user model*

$p_{\theta,\beta}(y \mid x, \hat{y}, u)$ is the *reverse user model*. $\longrightarrow$ variational reverse user model $q_\phi(y \mid x, \hat{y}, u)$

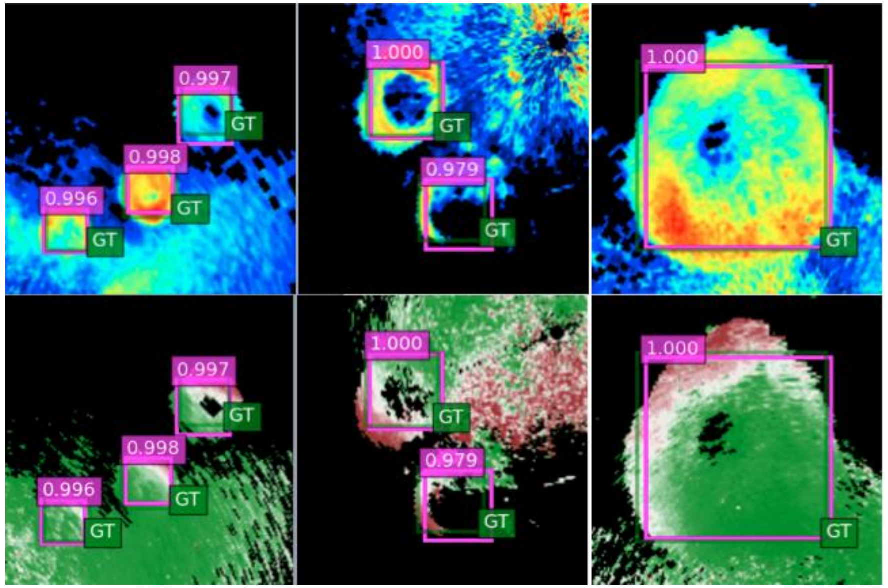
Example Detections and Tracks



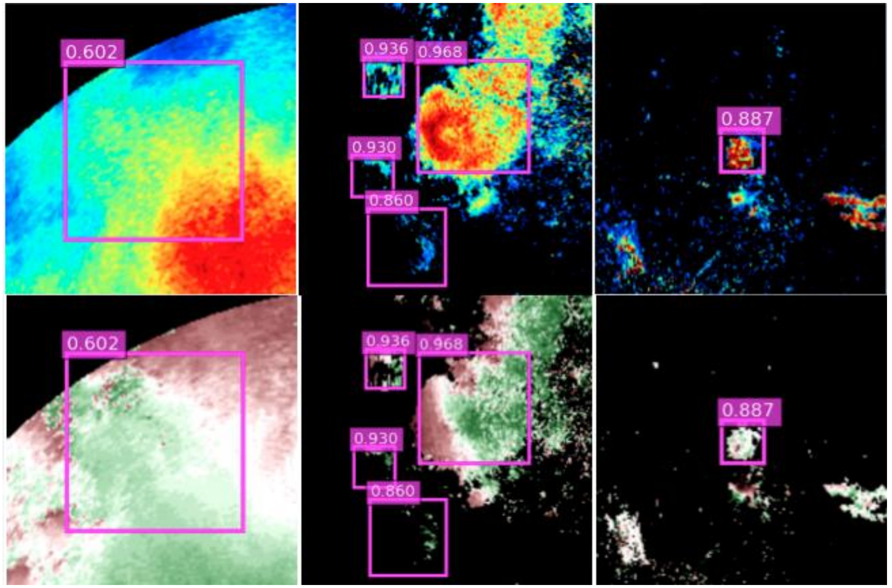


Error Analysis

Post processing	Swallow roost	Precipitation	Wind mills	Other roost	Misc. clutter	Unknown
Before	454	109	47	38	22	8
After	449	5	0	38	21	8

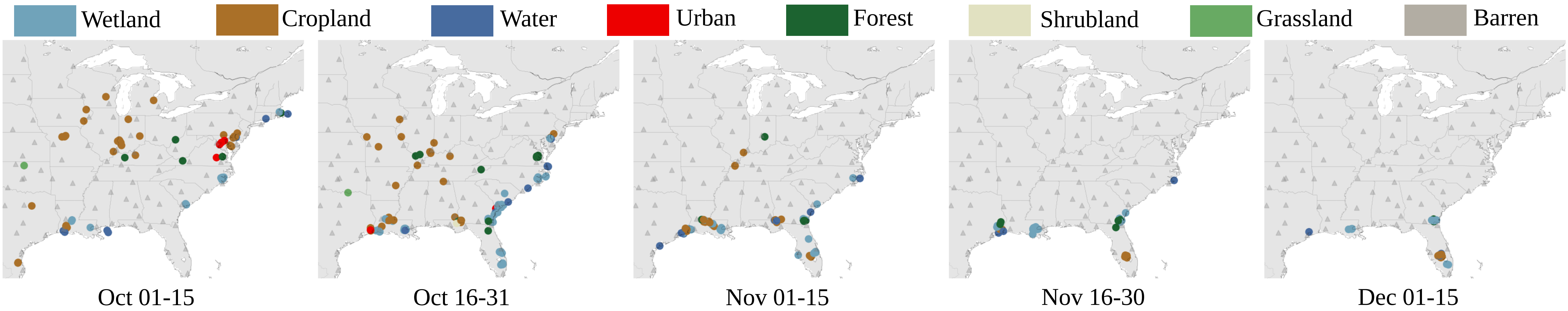


correctly detected roosts



rain, roosts of other species, windmills

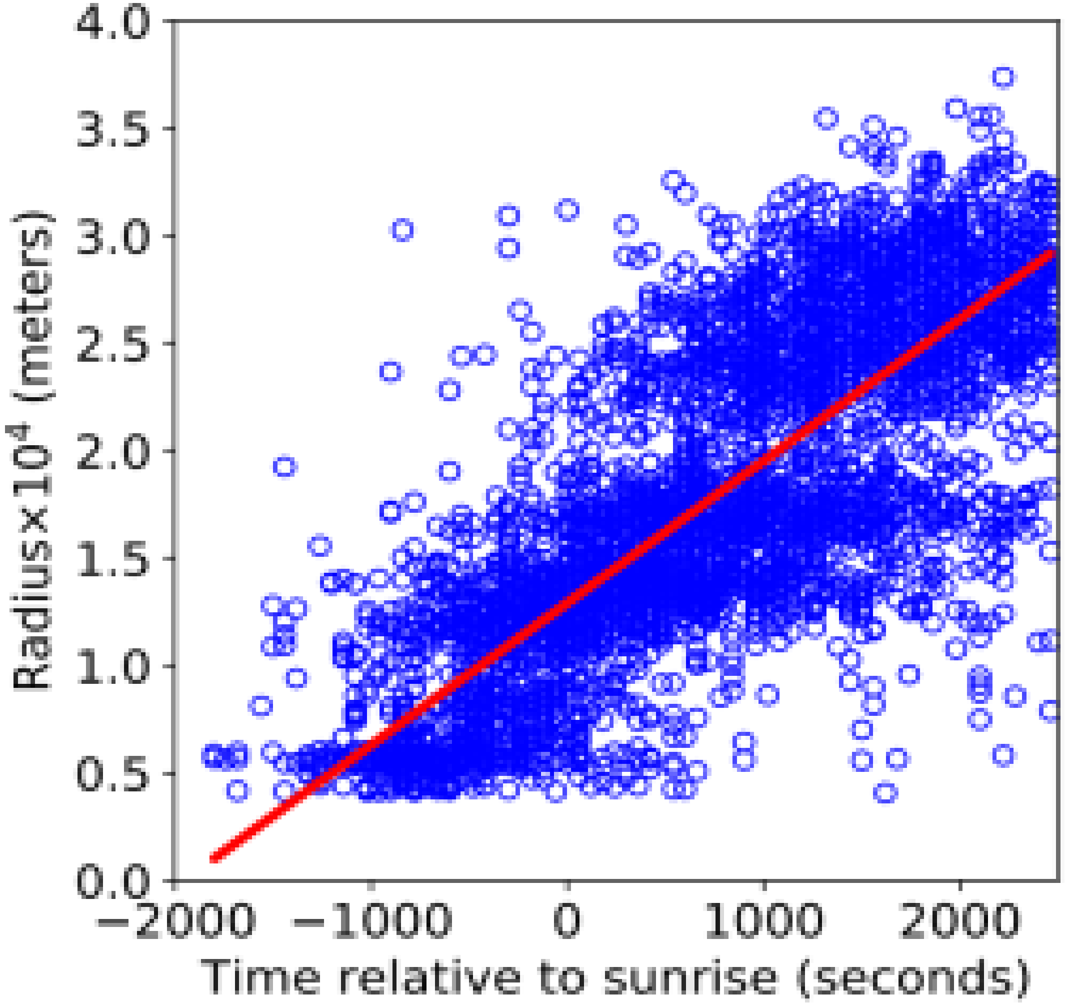
Where do the swallow roost?



Habitat: natural wetlands (e.g., cattails and phragmites) or agricultural

Widespread statistics of roost locations and habitat usage throughout a migratory season has not previously been

Roost emergence dynamics



Estimated airspeed velocity of tree swallows is 6.61 m/s (unladen)

Dark Ecology

UMassAmherst | College of Information & Computer Sciences

TheCornellLab

NSF