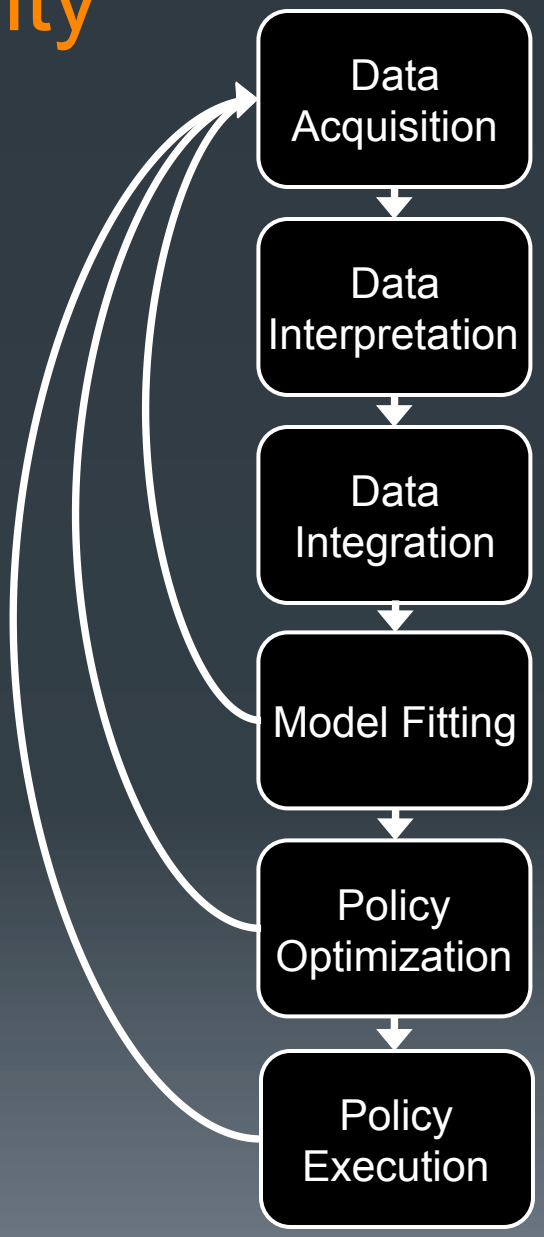
Machine Learning for Sustainable Development and Biological Conservation

Tom Dietterich
Distinguished Professor, Oregon State University
President, Association for the Advancement of Artificial Intelligence

Computational Sustainability

- The study of computational methods that can contribute to the sustainable management of the earth's ecosystems



Data Acquisition

Data Acquisition

- Africa is very poorly sensed
  - Only a few dozen weather stations reliably report data to WMO (blue points in map)
- Project TAHMO (tahmo.org)
  - TU-DELFT & Oregon State University
  - Deploy 20,000 stations across Africa
  - Provide data to farmers and to enable crop insurance industry
  - Increase agricultural productivity
- Computational Problem
  - Where to place the weather stations?
  - Krause, Singh & Guestrin, 2008

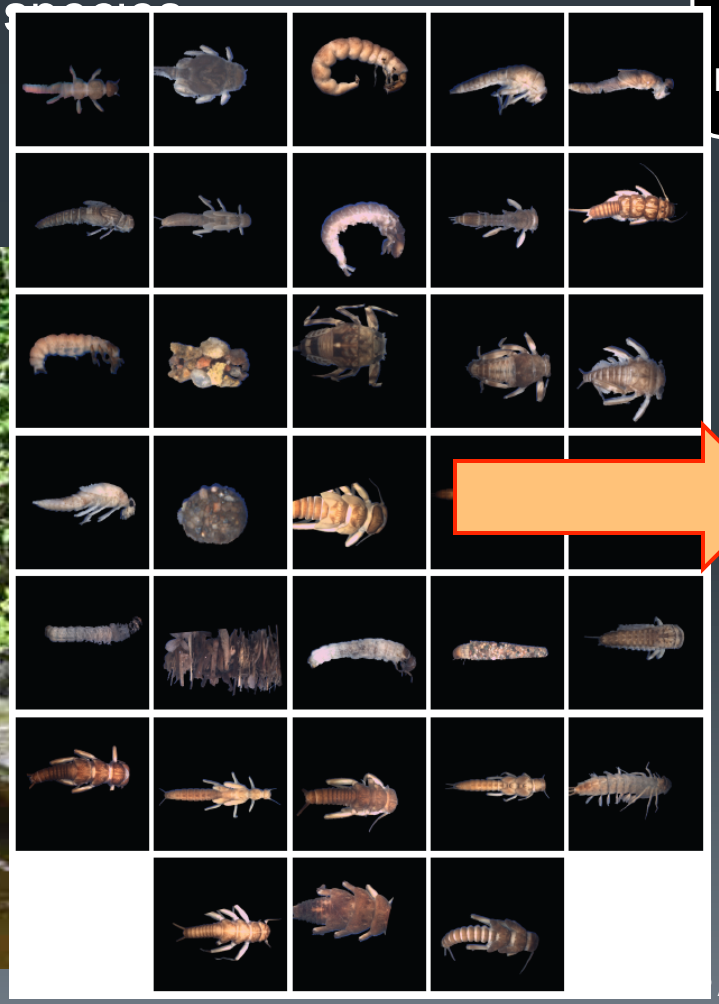
Data Interpretation

- Insect identification for population counting
- Raw data: image
- Interpreted data: Count by species
- Method: Computer Vision
- Lytle, et al., 2010

Data Acquisition

Data Interpretation

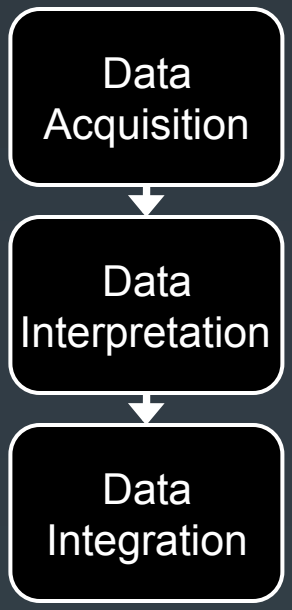
www.epa.gov

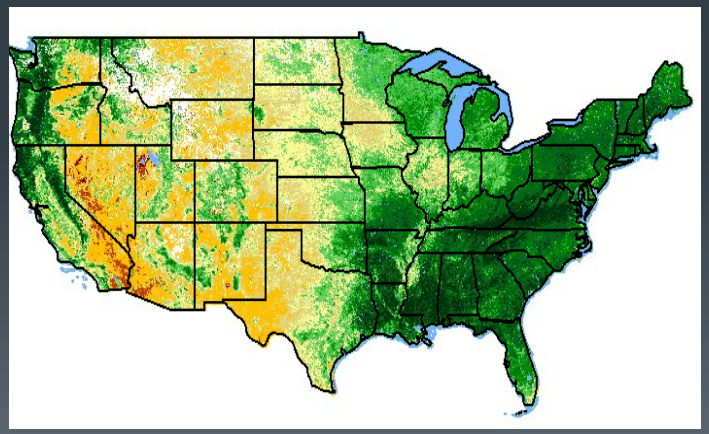


Species	Count
Limne	3
Taenm	15
Asiop	4
Epeor	25
Camel	19
Cla	12
Cerat	21

Data Integration

- Virtually all ecosystem prediction problems require integrating heterogeneous data sources
  - Landsat (30m; monthly)
    - land cover type
  - MODIS (500m; daily/weekly)
    - land cover type
  - Census (every 10 years)
    - human population density
  - Interpolated weather data (15 mins)
    - rain, snow, solar radiation, wind speed & direction, humidity

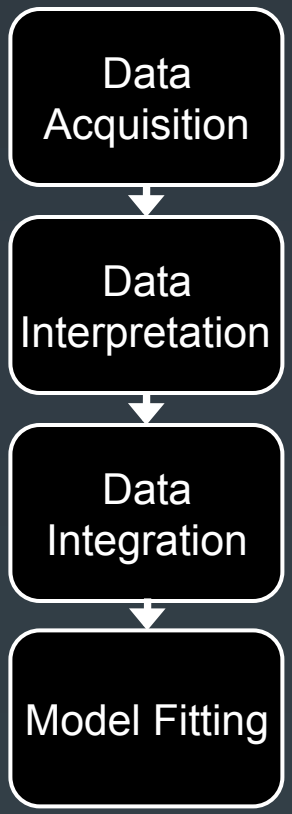




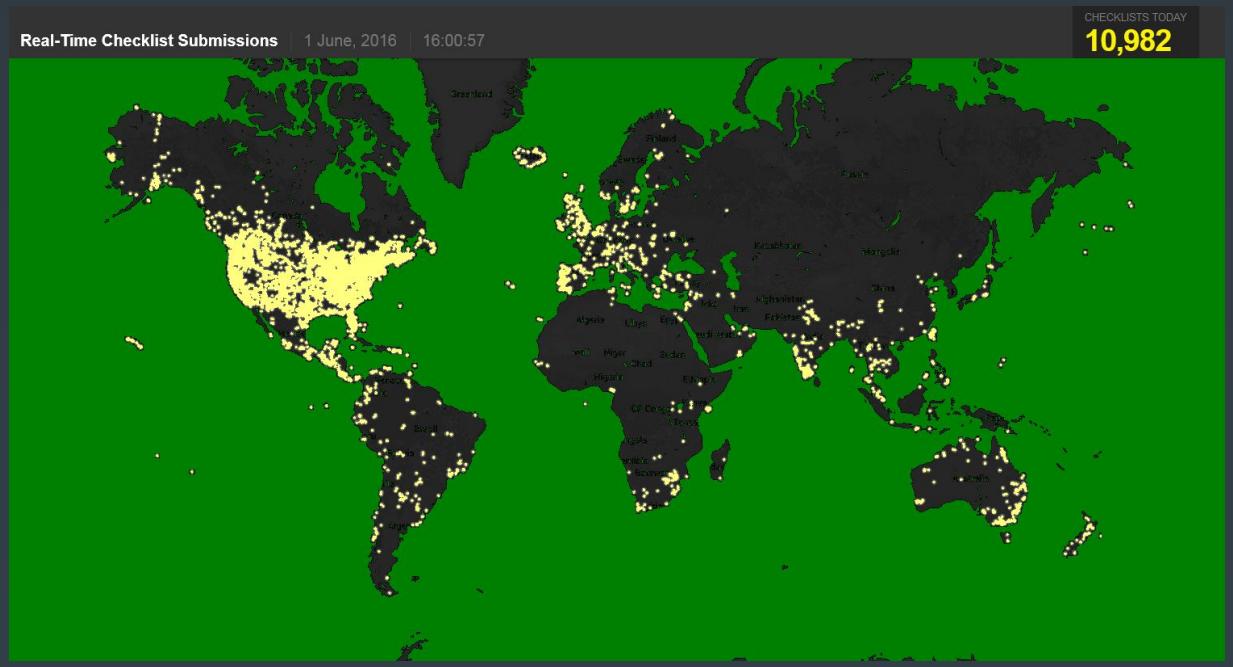
Landsat NDVI:
http://ivm.cr.usgs.gov/viewer/

Model Fitting with Machine Learning

- Species Distribution Models
  - create a map of the distribution of a species
- Migration and Dispersal Models
  - model the trajectory and timing of movement



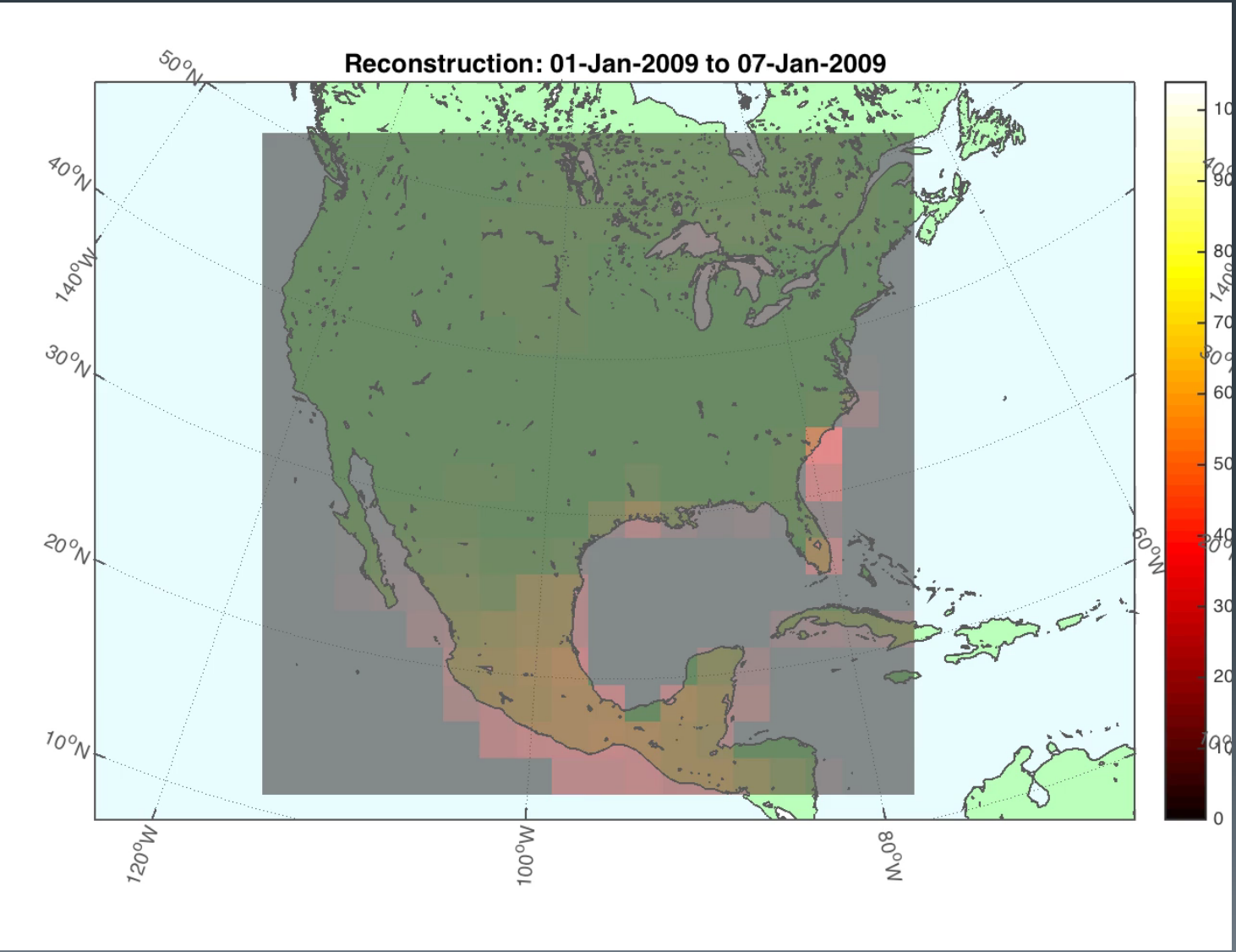
eBird Project



- Volunteer Bird Watchers
- Time, place, duration
- Species seen

- 8,000-12,000 checklists uploaded per day
- Computational Method: Collective Graphical Model (Sheldon et al., 2011)

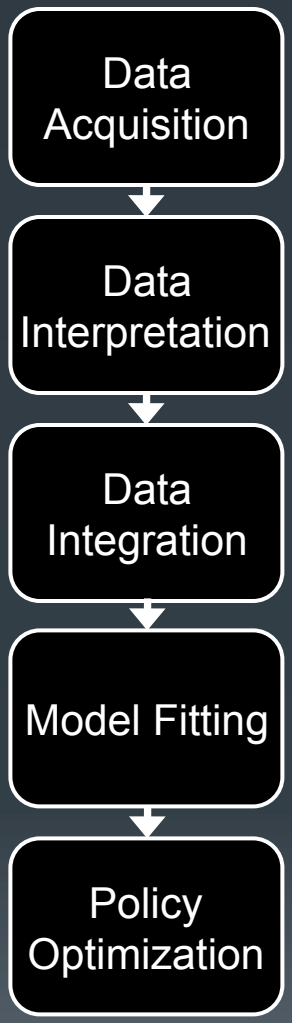
Fitted Migration Model
Ruby-Throated Humming Bird



Sheldon, Sun, Liu, Dietterich unpublished

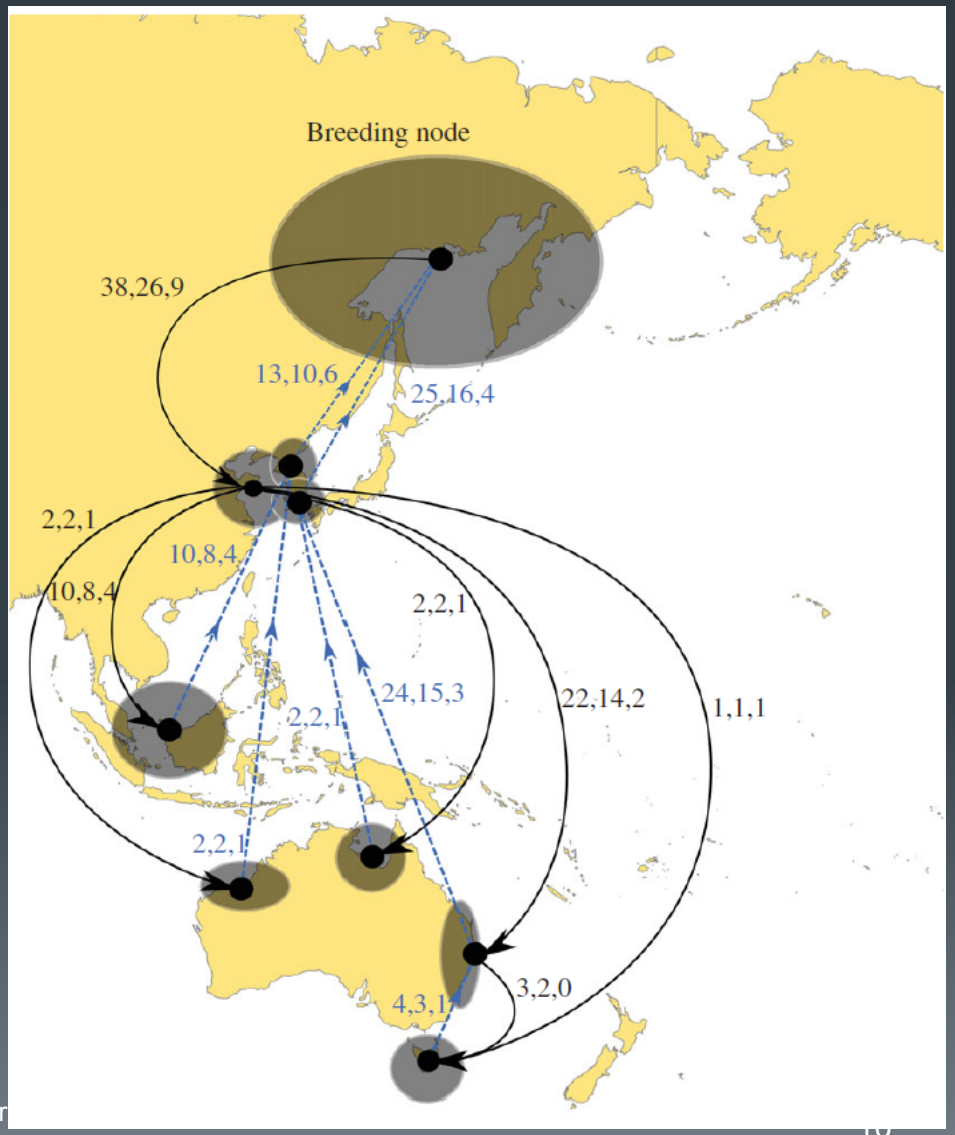
Policy Optimization

- Compute optimal policies for managing ecosystems
- Incorporate uncertainty about the future
- Computational Tools
  - MDPs (Markov Decision Problems)
  - POMDPs (Partially-Observable MDPs)
  - Point-based solvers (Pineau, 2003; Poupart, et al. 2005; Kurniawati et al, 2008)

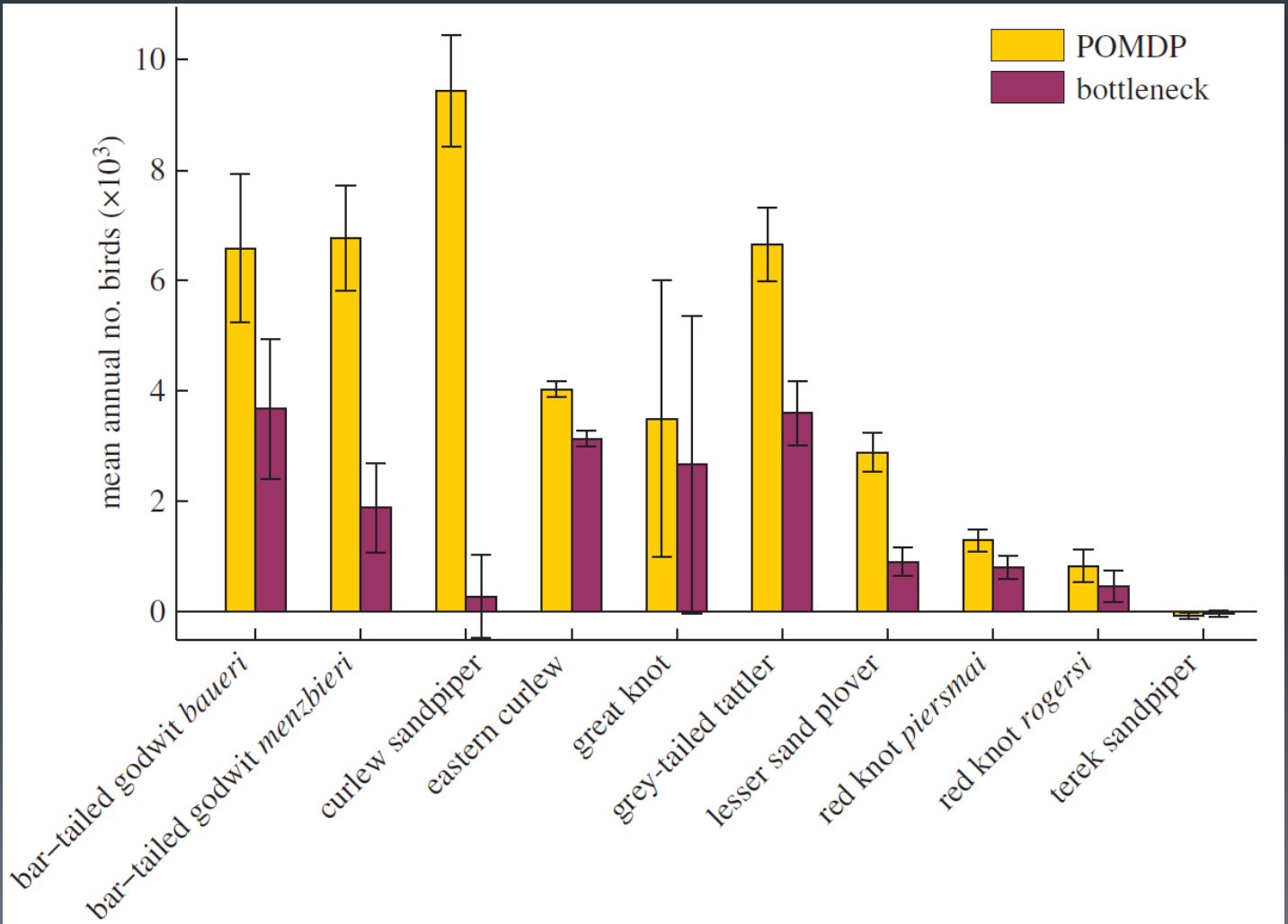


Protecting Coastal Habitat to Protect Migrating Birds from Sea Level Rise

- East Asia-Australia migratory pathways
- Sea Level Rise destroys habitat unless areas further inland have been protected
- Timing and location of protection depends on the timing of future sea level rises
- POMDP formulation
- Nicol, et al. 2015

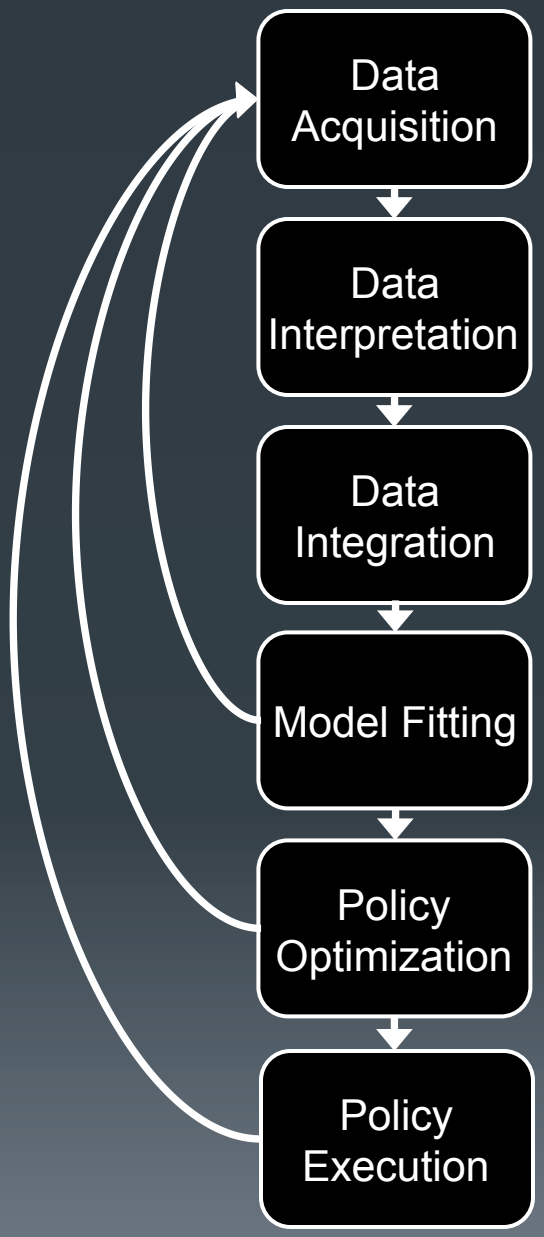


Results: Much More Successful than Existing Bottleneck Heuristic



Policy Execution

- Repeat
  - Observe Current State
  - Update Models and Re-Optimize
  - Choose and Execute Optimal Action



Summary

Locating weather stations in Africa

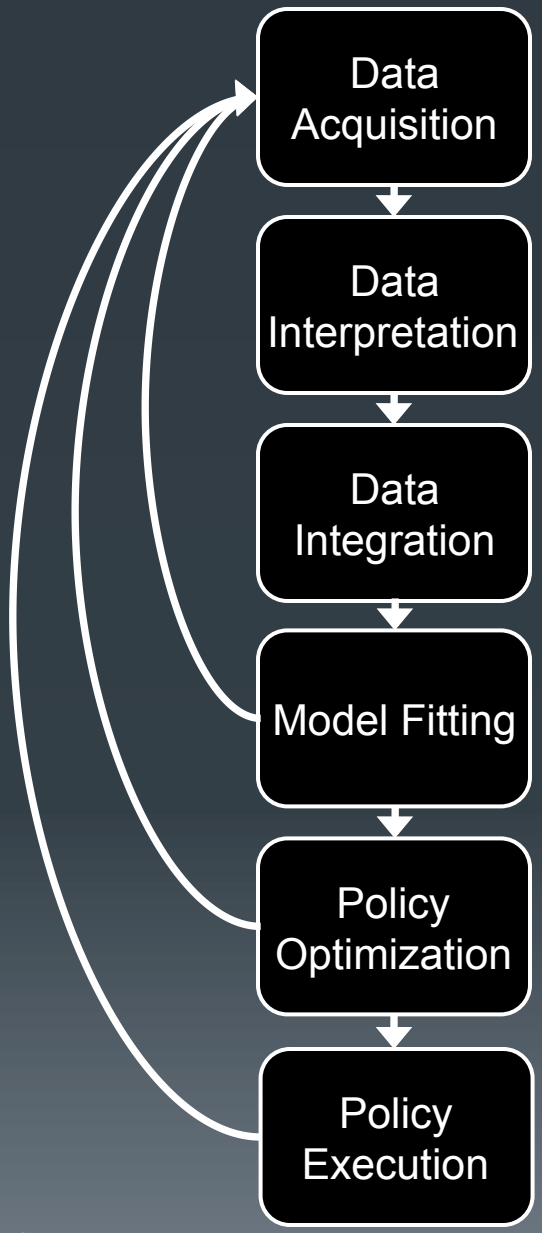
Images → Insect Species

Multiscale Data

Bird Migration Models fit to eBird Data

Where and when to purchase coastal habitat?

Action!



References

- Krause, A., Singh, A., & Guestrin, C. (2008). Near-Optimal Sensor Placements in Gaussian Processes: Theory , Efficient Algorithms and Empirical Studies. *Journal of Machine Learning Research*, *9*, 235–284.
- Lytle, D. A., Martínez-Muñoz, G., Zhang, W., Larios, N., Shapiro, L., Paasch, R., Moldenke, A., Mortensen, E. A., Todorovic, S., Dietterich, T. G. (2010). Automated processing and identification of benthic invertebrate samples. *Journal of the North American Benthological Society*, *29*(3), 867–874.
- Nicol, S., Fuller, R. A., Iwamura, T., & Chadès, I. (2015). Adapting environmental management to uncertain but inevitable change. *Proceedings Royal Society B*, *282*(1808), 20142984. http://doi.org/10.1098/rspb.2014.2984
- Pineau, J., Gordon, G., & Thrun, S. (2003). Point-based value iteration: An anytime algorithm for POMDPs. In *IJCAI International Joint Conference on Artificial Intelligence* (pp. 1025–1030).
- Sheldon, D., & Dietterich, T. G. (2011). Collective Graphical Models. In *NIPS 2011*.