Species Distribution Modeling with Remote Sensing

Amber McCullum, Juan Torres-Perez, Zach Bengtsson

Aug 19, 2021
Course Structure and Materials

• Three 1.5-hour sessions on August 12, 17, & 19
• Sessions will be presented once in English 12:00-13:30 EDT
• Webinar recordings, PowerPoint presentations, and the homework assignment can be found after each session at:
• Q&A following each lecture and/or by email at:
  – amberjean.mccullum@nasa.gov
  – juan.l.torresperez@nasa.gov
  – bengtsson@baeri.org
Webinar Agenda

Part 1: Overview of Species Distribution Models (SDMs)

Part 2: Using Wallace to Model Species Niches and Distributions

Part 3: Additional SDM Tools and Techniques, ASP Projects, and Summary
Part 3 Overview

- Mapping Application for Penguin Populations and Projected Dynamics (MAPPD)
- Wildlife Insights
- Map of Life (MOL)
- Circuitscape and Omniscape
- Fisheries and Climate Toolkit (FaCeT)
- Training Summary
- Q&A
Mapping Application for Penguin Populations and Projected Dynamics (MAPPD)
Mapping Application for Penguin Populations and Projected Dynamics

http://www.penguinmap.com/

Heather J. Lynch¹, Mathew Schwaller²
Chris Che-Castaldo¹, Grant Humphries¹, Michael Schrimpf¹

¹Stony Brook University Ecology & Evolution
²NASA Goddard
MAPPPD Overview

• Web-based, open-access, decision support tool
• Designed to assist scientists, non-governmental organizations, and policy-makers working to meet the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR) objectives
• Penguin population models
• Abundance estimates

Image Credit: Maureen Lynch via MAPPPD

Image Credit: MAPPPD
MAPPD Primary Objectives

1. A repository for submitting, vetting, and storing data on the distribution and abundance of penguin species data
2. A tool for searching the existing state of knowledge on Antarctic penguin abundance and distribution
3. A mechanism for creating and delivering checklists of all bird species at sites along the Antarctic Peninsula that are likely to be visited by humans

Drawings from del Hoyo et al. 1992
MAPPD Big Questions

1. Can we detect penguins?
2. Can we differentiate different species of penguins?
3. Can we estimate abundance?
4. Can we assess changes in abundance?
5. Can we start doing global/regional censuses?
6. Can we learn something new about penguin biology?
7. **Can we improve the decision-making process for conserving Antarctic marine living resources?**

Drawings from del Hoyo et al. 1992
MAPPPD Schematic

Satellite imagery & Field counts

Map-based search engine for current abundance and predicted dynamics

Front End

Back End

Dynamic Naive Bayesian Network Modelling

Site 1  Site 2
MAPPD Project Highlights

1 – Algorithm Development & Improvement
Develop algorithms to identify penguins and seabirds over the entire continent of Antarctica

2 – Discovery
Discovered several penguin and petrel “mega-colonies” from Landsat

> 1 million penguins discovered by Landsat
MAPPD Project Highlights

3 – Influencing Management
Danger Islands colonies were not considered high priority (blue shading) for conservation, but the Marine Protected Area (MPA) has been expanded (pink polygons) by ~ 2 million ha.

Maps taken from actual policy document being prepared by Argentina for the Antarctic Treaty Consultative Meeting.

4 – Ground Validation
Landsat-enabled exploration of previously unsurveyed territory.
MAPPD Research: Adélie Penguin Abundance

- Indicator species for ocean health
- Abundance and breeding data collected at fixed sites
- What are the mechanisms that underlie Adélie population dynamics?
  - Bayesian Population Dynamics Model
    - Includes process and observation error to all known Adélie penguin abundance data in Antarctic

Image Credit: Audubon Society

Image Credit: National Geographic
What did we find?

1) Interannual growth rates positively associated with maximum winter sea ice in year $t$ and negatively associated with maximum summer sea ice in $t - 4$.

2) Almost all of the interannual variability in growth rates remains unexplained.

Che-Castaldo et al, 2017
MAPPD Research: Adélie Penguin Abundance

- Average population growth rate multiplier: highly variable
- Periods of increased and decreased abundance
- Abundance declines in South Orkney Islands
- Steady increase in all other regions of Antarctic

Che-Castaldo et al, 2017
Random Walk of the Penguins
HOSTED BY DRIVENDATA

Penguins are among the most charismatic animals in the world and have captured the imaginations of news-makers, scientists, film producers, and the general public. Beyond their general intrinsic value, they are considered important ecosystem indicators. In other words, monitoring these beautiful species can tell us a lot about the general health of the Antarctic because penguins are important krill and fish predators, and changes (natural or anthropogenic) that influence prey abundance and environmental conditions will ultimately be detected through changes in distribution or population size.

You're not part of this competition. Yet...

Join the competition!
MAPPD Outcomes and Lessons Learned

• High-resolution remote sensing is answering questions we were not even asking 5-10 years ago (and not just for penguins!), and MAPPD is a key component of sharing this new data stream with policymakers.

• MAPPD and additional software can provide Antarctic stakeholders quasi-real time information on abundance and distribution at any user-defined spatial scale.

• MAPPD continues to be a work-in-progress.
  – “Once you get fancy, fancy gets broken.” – Morgan Spurlock

• Scientists and non-scientists have different needs, so be flexible and open to alternative workflows/software.
MAPPD Demo
Wildlife Insights
Wildlife Insights

https://www.wildlifeinsights.org

- Collection, dissemination, and analysis of camera trap data globally
- Combines field and sensor expertise, cutting edge technology, and advanced analytics to enable people everywhere to share wildlife data and better manage wildlife populations
- Upload images to website for species identification with artificial intelligence
Wildlife Insights: Model

- Artificial Intelligence Model within Google Cloud
  - Blank Image Filtering
  - Species Classification
    - Training data from multiple locations
    - Includes 837 classes with 732 species from around the world
    - Deep convolutional neural nets for classification using TensorFlow framework
Wildlife Insights: Model Performance

- Accuracy Assessment
  - View per-class performance for each species
  - [https://www.wildlifeinsights.org/about-wildlife-insights-ai](https://www.wildlifeinsights.org/about-wildlife-insights-ai)
Wildlife Insights: Get Involved

• Public Account:
  – https://app.wildlifeinsights.org/join

• Contributing and Managing Data Account:
  – https://www.wildlifeinsights.org/account-approval

• Join the Community Forum:
  – https://groups.google.com/u/0/g/wildlifeinsights?pli=1

• Check Out All the Video Tutorials:
  – https://www.wildlifeinsights.org/get-started/video-tutorials
Wildlife Insights Video Introduction
Map of Life (MOL)
Map of Life (MOL)

https://mol.org

- Provides species range information and species lists for any geographic area
- Multiple tools for exploring species habitat and trends in biodiversity
- Mobile app for discovering, identifying, and recording biodiversity
MOL Framework

- **Data Integration Goal**: Derive the best-possible probabilistic estimate of the occurrence of each species at the finest possible scale over a given temporal range, using the maximum amount of available information.
- Assembles and integrates different sources of data describing species distributions worldwide:
  - Species Range Maps
  - Species Occurrence Points
  - Ecoregions
  - Protected Areas
- Uses the power of Google Cloud to store and process data.
MOL Team

Faculty

- Walter Jetz
  Lead PI
  Yale University

- Robert Gurainick
  Associate Professor
  University of Florida

Staff

- Ajay Ranipeta
  Lead Engineer
  Yale University

- Anna Schuerkmann
  Program Manager
  Yale University

- Megan Blake
  Senior Administrative Assistant
  Yale University

- Yanna Sica
  Research Associate
  Yale University

- John Wiltshire
  Software Engineer
  Yale University

- Kate Ingeroff
  Research Associate
  Yale University

- Isabel Del Toro Mijares
  Research Assistant
  Yale University

- Jessica Vigneron
  Post Grad Associate
  Yale University

Funding for Map of Life

- NASA
- National Science Foundation (NSF)
- MacArthur Foundation

...And MANY MORE! https://mol.org/team

...And MANY MORE! https://mol.org/partnerships
Map Species

• Summary Maps
  – By species or location
• Detailed Maps
• Habitat Distribution
• Reserve Coverage
• Habitat Trends
• Projections
Species Status Information Index (SSII)

- Species occurrence data is important for understanding the distribution of biodiversity in space and time.
- For some regions and species biodiversity is poorly understood/ reflected.
  - Small number of records
  - Disproportionate data collection in time and/or space
- SSII captures how well existing data covers the species’ expected range.
Half Earth Map

https://map.half-earthproject.org

Priority places for biodiversity

The highlighted areas show places of particular importance to life on Earth. Understand these landscapes and how they need our care.

ALL MAPS
MOL Example: Global Habitat Loss

- Focus on terrestrial vertebrates
- Future land use change scenarios and impacts to biodiversity
- Substantial declines in suitable habitat for about 1,700 species of amphibians, birds, and mammals
- Identified regions and species in need of conservation planning activities

Powers and Jetz, 2019
Map of Life (MOL) Demo
Circuitscape and Omniscape
Landscape Connectivity

• The extent to which a landscape facilitates the movements of organisms and their genes
  
  – **Structural:**
    • Physical characteristics of a landscape
    • Topography, hydrology, vegetative cover, human and land use patterns, etc.
  
  – **Functional:**
    • Ecological preferences of an organism (including movement)
    • Habitat preference, dispersal ability, etc.

Areas best suited for movement of tigers and possible pinch points along them. The red and orange regions in the map show areas with least resistance to tiger movement. Image Credit: Wildlife Conservation Trust
Connectivity and Circuit Theory

- **Circuit Theory**: A process driven approach to modeling gene flow and the dispersal or movement routines of organisms.

*Image Credit: Circuitscape Team via Kim Hall*
Resistance Grid

- A grid in which each cell value reflects the landscape permeability (structural connectivity) or the energetic cost, movement difficulty, mortality risk, and/or avoidance behavior associated with species movement through that cell (functional connectivity).
Why Circuit Theory?

- **Connectivity is not just about corridors.**
  - We need to think about it more diffusely, particularly in working or dynamic landscapes. *The matrix matters.*
  - Connectivity is a dynamic process.
  - Redundancy is key - especially under changing land cover or climate.

- **Circuit theory helps to:**
  - Quantify gene flow and redundancy over complex landscapes
  - Prioritize pinch-points where connectivity might be lost sooner
  - Identify restoration opportunities and explore change scenarios
  - Provide theoretical justification for our work protecting and reconnecting landscapes

*From Dickson et al. 2019*
The Circuitscape Team

The Nature Conservancy
Kim Hall – NASA Lead (after Brad McRae)
Melissa Clark – Wall-to-Wall Circuitscape
Jim Platt – Coding/GIS/ArcGIS Plug-In
Mark Anderson – Co-PI and Applications
Carrie Schloss – Omniscape
Aaron Jones – Omniscape/Linkage Mapper

MIT & Julia Computing
Ranjan Anantharaman – Lead on Update
Viral Shah – Co-PI
Alan Edelman – Collaborator

Conservation Science Partners
Brett Dickson - Circuitscape
Dave Theobald – Resistance Grids
Vincent Landau – Circuitscape, Omniscape
What is Circuitscape?

https://circuitscape.org/

• Open-source connectivity analysis software package, currently in the Julia programming language

• Applies electrical circuit theory to questions of how genes, animals, or processes flow across a heterogenous landscape
  – Tracks where “current” flows across the resistance surface
  – Quantifies spatial patterns of current accumulation along higher or lower resistance areas

• Inputs:
  – Habitat suitability, topography, climatic factors

• Output:
  – Animal movement flow maps (current or future)
Accessing and Using Circuitscape

• The new Circuitscape is built entirely in the Julia language, a new programming language for technical computing.

• Available on GitHub: https://github.com/Circuitscape/Circuitscape.jl/blob/master/README.md#Installation
  — Anantharaman et al, 2020

• Over 500 papers in the natural sciences published in English from 2007–2020 where Circuitscape analyses used

Circuitscape is widely used, with diverse applications.

Hall et al. 2020
Omniscape

- Runs Circuitscape in a circular moving window
- Used to produce omni-directional habitat connectivity
  - Allows sources, destinations, and intensity of animal movement to be informed by continuous spatial data
- Outputs:
  - Cumulative Current Flow
  - Flow Potential
  - Normalized Current Flow
- Available on GitHub: https://github.com/Circuitscape/Omniscape.jl#installation
Migrations in Motion

https://maps.tnc.org/migrations-in-motion/#4/38.82/-103.93

• The average direction mammals, birds, and amphibians need to move to track hospitable climates as they shift across the landscape.
Circuitscape Example: Sage-Grouse Conservation

• Assisted in the development of a Sage-Grouse Conservation Partnership in Southern Oregon

• Goals:
  – Model relative accessibility of localized areas
  – Identify linkage zones of easiest movement based on landscape
  – Describe areas where movement may be constrained or fragmented

Image Credit: eBrid
Image Credit: Sage Grouse Initiative
Sage-Grouse Conservation: Study Area

[Map showing the study area with various symbols and labels]

Jones, 2015
Sage-Grouse Conservation: Data Inputs

• MANY spatial datasets, including Landsat-derived landcover map from the National Land Cover Database

• Resistance Data Types:
  – **Energy Cost/Movement Difficulty**: Includes base habitat layers
  – **Mortality Risk**: Physical footprints of anthropogenic landscape features
  – **Avoidance**: Densities/inverse Euclidean distances of anthropogenic features
Sage-Grouse Conservation: Habitat Classification

Energy Cost/Movement Difficulty

Habitat Classification

- These classes were included in the modeling of resistance due to energy cost and movement difficulty.

Jones, 2015
Sage-Grouse Conservation: Resistance

Energy Cost/Movement Difficulty

Resistance

- Higher resistance values indicate greater cost and difficulty for sage-grouse movement.

Jones, 2015
Sage-Grouse Conservation: Resistance

Mortality Risk or Avoidance

Resistance
- Higher resistance/mortality: high housing densities
- Higher resistance/avoidance: powerplants and high-voltage transmission lines

Jones, 2015
Sage-Grouse Conservation: Cumulative Resistance

**Combined Resistance Surface**
- Higher resistance values indicate greater cost and difficulty for sage-grouse movement, greater mortality risk, and/or greater avoidance behavior.

**Warmer Colors = Higher Resistance**

*Jones, 2015*
Sage-Grouse Conservation: Normalized Least-Cost Corridors (NLCCs)

- **NLCCs (Corridors):** Each defined by cumulative movement costs relative to its respective LCP.

- **‘Linkage Zone’**: Broad belts of land with relatively greater habitat continuity. (Here, NLCCs = Linkage Zones.)

- Framework for Circuitscape runs

Jones, 2015
Sage-Grouse Conservation: Linkage Pinch-Points

Pinch-Points:
• Locations of highly constricted (and thus strong) current flow

• Network severance possible with loss of small amount of movement habitat

• Potential areas for protection from habitat loss or degradation

Jones, 2015
Fisheries and Climate Toolkit (FaCeT)
Dynamic Ocean Management

- Improved understanding of the nature of aquatic species, users, and oceanographic features to guide commercial activities
  - Balance economically viable industries with ecological sustainability
- Climate change shifting species distributions and therefore fishing fleets
  - Adaptation of management and conservation strategies
- Need to address Sustainable Development Goal 14 and related targets 14.2 and 14.4
Fisheries and Climate Toolkit (FaCeT)

https://fisheriesclimatetoolkit.sdsu.edu/

Camrin Braun\textsuperscript{1,2}, Nima Farchadi\textsuperscript{3}, Kathy Mills\textsuperscript{4}, Dylan Pugh\textsuperscript{4},  
Alex Kerney\textsuperscript{4}, Riley Young Morse\textsuperscript{4}, Stephanie Brodie\textsuperscript{5}, Heather Welch\textsuperscript{5},  
Andrew Allyn\textsuperscript{4}, Nerea Lezama Ochoa\textsuperscript{5}, Steven Bograd\textsuperscript{5,6},  
Elliott Hazen\textsuperscript{5,6}, Rebecca Lewison\textsuperscript{3}

\textsuperscript{1}University of Washington, \textsuperscript{2}Woods Hole Oceanographic Institution, \textsuperscript{3}San Diego State University,  
\textsuperscript{3}Gulf of Maine Research Institute, \textsuperscript{4}NOAA SWFSC, Environmental Research Division,  
\textsuperscript{5}University of California Santa Cruz
Fisheries Needs and Applications of FaCeT

- Providing Forecasting/Predictions
- Tracking Magnitude/Velocity of Change
- Harnessing Big Data and Data Pipelines
- Climate Change Uncertainty in a Fisheries Context

Ecological Consequences

Climate-Driven Anomalies

Economic Impacts

Projected Shift in Suitable Habitat for Atlantic Cod by 2100 Under High-Emission Scenario

Image Credit: FaCeT

Wired.co
FaCeT Goals

Develop online products that enable:

- Tracking dynamic species & vessel distributions
- Project fishery interactions and “hotspots”
- Integrate climate projections to:
  - Identify fishery-relevant climate anomalies
  - Inform climate-resilient fisheries management
FaCeT Example: Shortfin Mako Shark

- Shortfin mako sharks subject to overfishing
- Slow growth rates, reproduce after many years
- Found along the East Coast of the US
- Highly migratory
- Managed through NOAA’s Consolidated Atlantic Highly Migratory Species Fishery Management Plan
- Often caught incidentally in longline fisheries targeting swordfish and tuna
- Current regulations exist
Shortfin Mako Shark Locations

ICCAT SMA conventional tags N = 7643
FaCeT Example: Dynamic Vessel Distribution Models (dVDMs)

- Shipboard AIS >70,000 fishing vessels globally
- Potential climate-driven shifts in fisher distribution
FaCeT Example: Dynamic Vessel Distribution Models (dVDMs)
FaCeT Example: Dynamic Vessel Distribution Models (dVDMs)
Summary

• Species Distribution Models (SDMs) allow us to assess the suitability of a habitat for a species.
  – SDMs primarily use environmental data and occurrence data to build a model for predictions of habitat suitability.

• **Wallace** is a user-friendly application for SDM that provides guidance towards following best-practices at each step.

• Additional projects and tools for SDM include:
  – Mapping Application for Penguin Populations and Projected Dynamics (MAPPD)
  – Wildlife Insights
  – Map of Life (MOL)
  – Circuitscape and Omniscape
  – Fisheries and Climate Toolkit (FaCeT)
Homework and Certificate

• One homework assignment:
  – Answers must be submitted via Google Form, accessed from the ARSET website.
  – Due date for homework: September 2, 2021

• A certificate of completion will be awarded to those who:
  – Attend all live webinars
  – Complete the homework assignment by the deadline
  – You will receive a certificate approximately two months after the completion of the course from marines.martins@ssaihq.com.
Contacts

• Trainers:
  – Juan Torres-Pérez: juan.l.torresperez@nasa.gov
  – Amber McCullum: amberjean.mccullum@nasa.gov
  – Zach Bengtsson: bengtsson@baeri.org

• Training Webpage:

• ARSET Website:
  – https://appliedsciences.nasa.gov/what-we-do/capacity-building/arset

Follow us on Twitter @NASAARSET
Thank You!