Bayesian data-model synthesis for biological conservation and management in Antarctica

Heather J. Lynch, Mathew Schwaller, Christian Che-Castaldo
Stony Brook University Ecology & Evolution
1 – Algorithm development & improvement:

Develop algorithms to identify penguins and seabirds over the entire continent of Antarctica. (Landsat & Sub-meter commercial)
1 – Algorithm development & improvement:
Develop algorithms to identify penguins and seabirds over the entire continent of Antarctica. (Landsat & Sub-meter commercial)

2 – Discovery: Discovered several penguin and petrel “mega-colonies” from Landsat images revealing their pinkish guano. Reshaping our understanding of seabird biogeography.

> 1 million penguins discovered by Landsat
1 – Algorithm development & improvement:
Develop algorithms to identify penguins and seabirds over the entire continent of Antarctica. (Landsat & Sub-meter commercial)

2 – Discovery: Discovered several penguin and petrel “mega-colonies” from Landsat images revealing their pinkish guano. Reshaping our understanding of seabird biogeography.

> 1 million penguins discovered by Landsat
1 – Algorithm development & improvement:
Develop algorithms to identify penguins and seabirds over the entire continent of Antarctica. (Landsat & Sub-meter commercial)

2 – Discovery: Discovered several penguin and petrel “mega-colonies” from Landsat images revealing their pinkish guano. Reshaping our understanding of seabird biogeography.

> 1 million penguins discovered by Landsat

3 – Influencing management: Danger Islands colonies were not considered high priority (blue shading) for conservation but proposed MPA has been expanded (pink polygons) by ~ 2 million ha as a direct result of discoveries made using Landsat imagery under NASA funding. Maps taken from actual policy document being prepared by Argentina for the Antarctic Treaty Consultative Meeting.
1 – Algorithm development & improvement:
Develop algorithms to identify penguins and seabirds over the entire continent of Antarctica. (Landsat & Sub-meter commercial)

2 – Discovery: Discovered several penguin and petrel “mega-colonies” from Landsat images revealing their pinkish guano. Reshaping our understanding of seabird biogeography.

3 – Influencing management: Danger Islands colonies were not considered high priority (blue shading) for conservation but proposed MPA has been expanded (pink polygons) by ~ 2 million ha as a direct result of discoveries made using Landsat imagery under NASA funding. 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.

> 1 million penguins discovered by Landsat
At the site level

Data-rich sites

Cape Crozier

Data-poor sites

Lauff Island

Litchfield Island

Cape Cornish

How does the model infer abundance when there is no data?

Shared covariates allow for a ‘best-guess’ in years with missing data.

but still, uncertainty is huge between surveys

Est. from nest counts in black
Est. from chick counts in red
Landsat pixel identified as “guano” class

High-res guano patch

30 m

Using:
• Landsat-4
• Landsat-5
• Landsat-7 (incl. SLC error era)
• Landsat-8

Using the guano stain to georegister imagery but this will not be required starting with Landsat-8.
$Y_i \sim Binom(n_i, p(\text{fraction covered}_i))$

$logit(p(\text{fraction covered}_i)) = \beta_0 + \beta_1 \times \text{fraction covered}$
We treat each pixel as its own stack; easily accommodates pixels lost to SLC error.

$Y_i \sim Binom(n_i, p(fraction\ covered_i))$

$logit(p(fraction\ covered_i)) = \beta_0 + \beta_1 * fraction\ covered$
One caveat: Total estimated abundance depends on the area of interest.

Why? Because even areas that have never been classified as guano will have some non-zero detection probability.

To the rescue: A new Landsat-8 based bare rock layer.
Does the integration of Landsat-based estimates improve model results? Yes!
Does the integration of Landsat-based estimates improve model results? Yes!

Our integration of the (statistically-downscaled) Landsat-derived abundance estimates radically changes our understanding of long-term trend and narrows our uncertainty on historical abundance.

* High resolution commercial satellite imagery not always better
MAPPPD retrospective

Successes:
• Created **reproducible workflows** for Landsat imagery interpretation
• Developed time series models that **incorporate multiple data types**
• Moved towards **open-source community development** of models that can be incorporated into ensemble model forecasts of abundance
• Created a decision support tool that is actively being used within the stakeholder community
MAPPPD retrospective

Successes:
• Created **reproducible workflows** for Landsat imagery interpretation
• Developed time series models that **incorporate multiple data types**
• Moved towards **open-source community development** of models that can be incorporated into ensemble model forecasts of abundance
• Created a decision support tool that is actively being used within the stakeholder community

Challenges/Open questions/Future directions:
• Automated image interpretation of high-resolution satellite imagery
Convolutional Neural Networks
MAPPPD retrospective

Successes:
• Created reproducible workflows for Landsat imagery interpretation
• Developed time series models that incorporate multiple data types
• Moved towards open-source community development of models that can be incorporated into ensemble model forecasts of abundance
• Created a decision support tool that is actively being used within the stakeholder community

Challenges/Open questions/Future directions:
• Automated image interpretation of high-resolution satellite imagery
• High-performance and high-throughput computing bottlenecks

↑ CNNs require GPUs
↑ Required to scale

Pleiades
ICEBERG: Imagery Cyberinfrastructure and Extensible Building Blocks to Enhance Research in the Geosciences

Heather Lynch (Stony Brook University)
Shantenu Jha (Department of Electrical and Computer Engineering, Rutgers University)
Vena Chu (Department of Geography, University of California Santa Barbara)
Mike Willis (Department of Geological Sciences, University of Colorado Boulder)
Mark Salvatore (Department of Physics and Astronomy, Northern Arizona University)
Huge thanks to Woody Turner and Cindy Schmidt for all the help over the life cycle of the MAPPPD project!