

Building Tools for Community Science Datasets to Model Bird Populations and Highly Pathogenic Avian Influenza in Aotearoa New Zealand

Richard Littauer

Kris Bubendorfer; Markus Luczak-Roesch

Pepeha

- Ko Richard Graham Littauer tōku inga.
- Nō Europe oku tupuna.
- Nō te awa o Connecticut me Winooski au.
- Kei te noho au ki Ōwhiro awa.

My goal

- Improve our understanding of taonga manu and te taiao
- Build tools and best practices for using community science data.
- Model how HPAI will affect wild bird populations here in Aotearoa New Zealand

Research questions

- How can current community science Best Practices be improved?
- How can we harmonize taxonomies?
- Can we predict HPAI effects on our taonga manu?

What is Community Science?

- No exact definition
- Amateur (non-professional) contributions
- Also called citizen science
- At scale
- Useful for research
- Skews to rich, educated, white people



Karo Felted Scale

What is Community Science?

- eBird
 - 1 billion observations
 - 100 million checklists
 - 1 million users
 - 10,000 species of birds
- iNaturalist
 - 250 million observations
 - 3 million users
 - 500k species

SCIENCE

- Silver Fern, collected by Banks and Solander
- C. 1769-1770, Aotearoa
- Specimen in Te Papa



Ponga

SCIENCE

- Silver Fern, photographed by Littauer
- Nov 14, 2022, Ohakune
- <https://www.inaturalist.org/observations/141999533>



Ponga

SCIENCE

eBird New Zealand CHECKLIST S122400012

Mon 14 Nov 2022 7:06 PM

Edit date and effort

Tongariro NP -- Maungawhero Forest Walk  Ruapehu District County, Manawatu-Wanganui, New Zealand

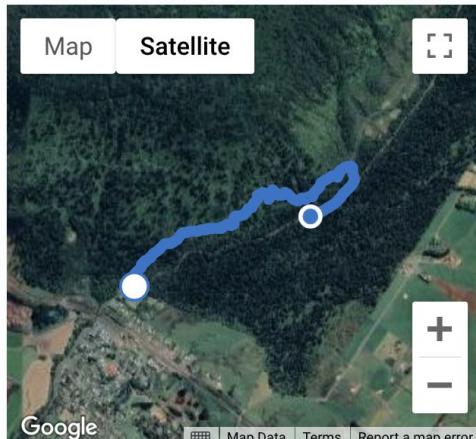
Edit location

Richard Littauer

Share

Traveling Incomplete 

2 1 hr, 3 min 1.33 mi



 eBird Mobile Tracks

Submitted from eBird for iOS, version 2.13.20

Edit comments



8

Species Observed 16 individuals

- 1 New Zealand Pigeon *Hemiphaga novaeseelandiae*
- 1 Long-tailed Koel *Urodynamis taitensis*
- 1 Tui *Prosthemadera novaeseelandiae*
- 2 Gray Gerygone *Gerygone igata*
- 5 Tomtit *Petroica macrocephala*
- 3 Silvereye *Zosterops lateralis*
- 1 Song Thrush *Turdus philomelos* *
- 2 Eurasian Blackbird *Turdus merula* *

eBird

- Birds.
- Different protocols for birding
- Audio and photo uploads
- For science.



Takahē

eBird Stats WORLD

10,826

Species

97,753,170

Complete Checklists

1,034,384

eBirders

eBird

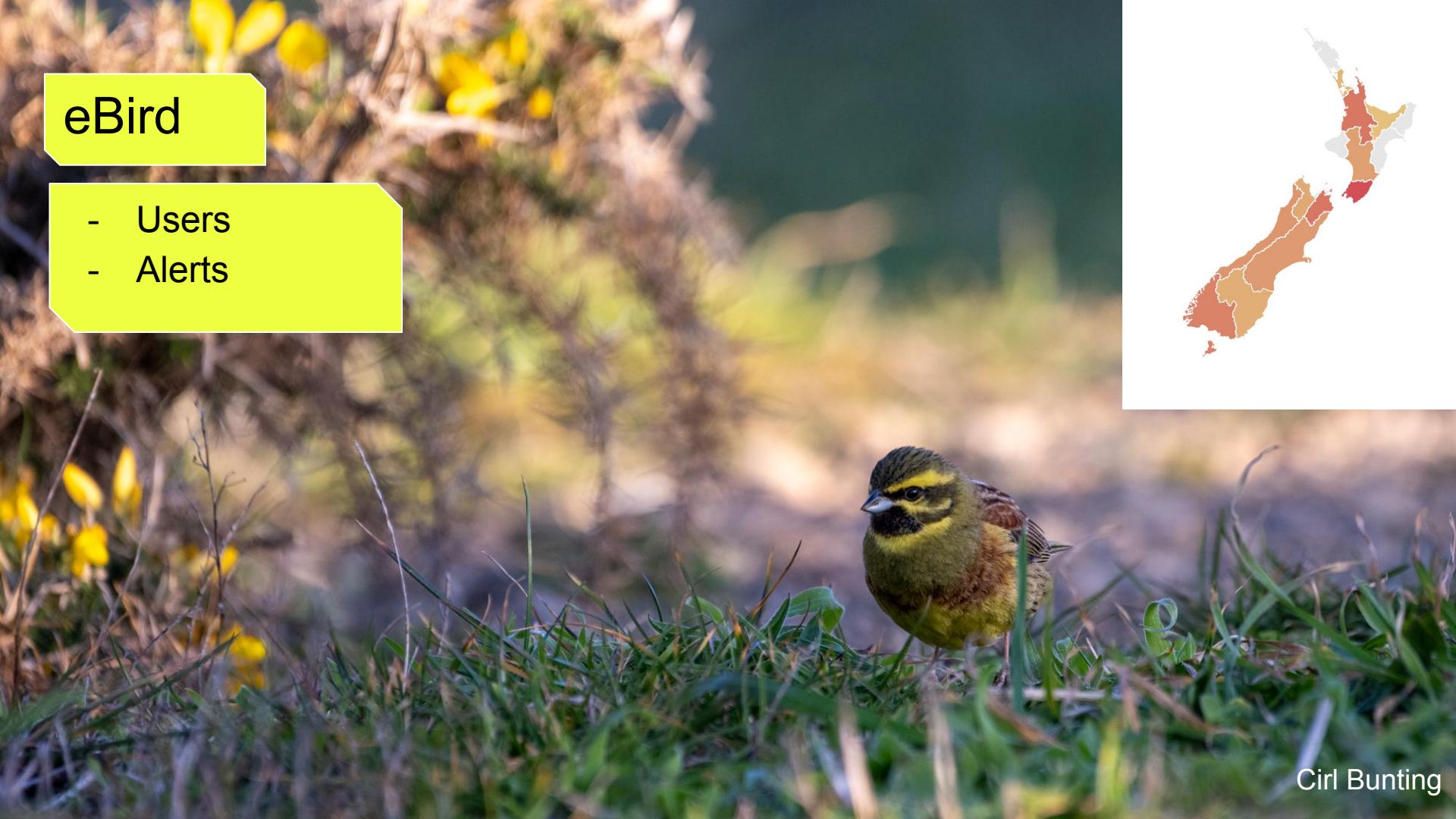
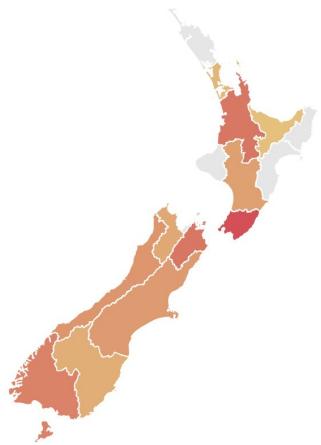
- Regions
- Hotspots



Te Whanga-nui-a-Tara

eBird

- Users
- Alerts



Cirl Bunting

eBird

- Reviewers



Pōpokotea

iNaturalist

- All momo
- Photos or audio
- "Research grade"
- Comments



Korimako, Manuka, & Rō

[Change Region](#)

New Zealand



363
Species

812.7K
Checklists

13,726
eBirders

My Stats 132 Checklists 1243 eBirders 101 Trip Reports 27

Overview

[Partners](#)[News & Resources](#)[Bird List](#)[Recent Checklists](#)[Trip Reports](#)[Subregions](#)[Hotspots](#)[eBirders](#)[Illustrated Checklist](#)[VIEW MY...](#)[My eBird](#)[Life List](#)[Target Species](#)

eBird New Zealand

[Learn more about regional partners](#)

New Zealand News



Easier, more accurate distance tracking in eBird Mobile

17 Jun 2025 | By Aotearoa NZ eBird Team

Resources:

[Contact the Aotearoa NZ eBird Team](#)[Best Practices](#)[Recommended eBird Settings \(App/Online\)](#)[Documenting Rarities](#)[Young Birders NZ](#)[Are you really seeing Pārera | Grey Ducks?](#)

Observations



Species

Location

Go

Filters

New Zealand

3,062,893 OBSERVATIONS | 20,282 SPECIES | 18,290 IDENTIFIERS | 47,996 OBSERVERS

Caledonia

Map Grid List Places of Interest Redo search in map

Brisbane Norfolk Island Caledonia Auckland Wellington Hobart Sydney Tasman Sea

Dacrycarpus dacrydioides (Kahikatea) South Island, Chri... Today 1 1min

Gallirallus australis (Weka) Milford Track, Sou... Dec 30, 2013 1 6m

Lepidothamnus laxifolius (Pygmy Pine) 7875, New Zealand Yesterday 1 9m

Halocarpus bidwillii (Bog Pine) 7875, New Zealand Yesterday 1 9m

Celmisia armstrongii (Armstrong's Mountain Daisy)

Map Legend ▾

The map displays a high density of orange observation points across the North Island of New Zealand, with a few scattered points on the South Island. The interface includes a search bar, location filters, and a sidebar for recent species sightings.

The problem

- Protocols not standardized between platforms
- Skill varies wildly
- Vetting process different per platform
- Bias in many areas:
 - Bigger birds are easier to see
 - Birds near human habitation are easier to count
 - People mainly bird on the weekends or in Zealandia
- And, generally, few best practices exist for cleaning data

The problem

The screenshot shows a Zenodo dataset page. At the top, there's a blue header bar with the Zenodo logo, a search bar, navigation links for 'Communities' and 'My dashboard', and buttons for 'Log in' and 'Sign up'. A message about a planned intervention on July 22nd is displayed. Below the header, the dataset details are shown: 'Published January 21, 2020 | Version v1.0'. The title is 'Best Practices for Using eBird Data v1.0'. The authors listed are Strimas-Mackey, Matthew¹; Hochachka, Wesley M.¹; Ruiz-Gutierrez, Viviana¹; Robinson, Orin J.¹; Miller, Eliot T.¹; Auer, Tom¹; Kelling, Steve¹; Fink, Daniel¹; Johnston, Alison¹. There are buttons for 'Book' and 'Open'. To the right, there are statistics: 1K views and 140 downloads, with a link to 'Show more details'. Below this, there's a 'Versions' section showing 'Version v1.0' (Jan 21, 2020) with DOI 10.5281/zenodo.3620739. A note explains that you can cite all versions by using the DOI. Further down, there's an 'External resources' section with a link to 'CornellLabofOrnithology/ebird-best-practices'.

zenodo

Search records...

Communities My dashboard

Log in Sign up

Planned intervention: On Tuesday, July 22nd 05:00 UTC Zenodo will be unavailable for 15-30 minutes to perform a storage cluster upgrade.

Published January 21, 2020 | Version v1.0

Book Open

1K VIEWS 140 DOWNLOADS

Show more details

Best Practices for Using eBird Data v1.0

Strimas-Mackey, Matthew¹; Hochachka, Wesley M.¹; Ruiz-Gutierrez, Viviana¹; Robinson, Orin J.¹; Miller, Eliot T.¹; Auer, Tom¹; Kelling, Steve¹; Fink, Daniel¹; Johnston, Alison¹

Show affiliations

View the book Best Practices for Using eBird Data at <https://cornelllabofornithology.github.io/ebird-best-practices/>

Files

CornellLabofOrnithology/ebird-best-practices-v1.0.zip

CornellLabofOrnithology/ebird-best-practices-v1.0.zip

CornellLabofOrnithology-ebird-best-practices-6180df5

.Rprofile 19 Bytes

.gitignore 148 Bytes

.nojekyll 0 Bytes

01_introduction.Rmd 21.1 kB

02_ebird-data.Rmd 36.7 kB

Versions

Version v1.0 Jan 21, 2020

10.5281/zenodo.3620739

Cite all versions? You can cite all versions by using the DOI 10.5281/zenodo.3620739. This DOI represents all versions, and will always resolve to the latest one. [Read more](#).

External resources

Available in

CornellLabofOrnithology/ebird-best-practices Release: v1.0

Indexed in

The problem

- eBird best practices aren't immediately transferable to iNaturalist
- There's nothing related to seabird data
- Reviewer effort isn't standardized
- Cleaning can be species or location specific
- There's no automatic tool for comparing taxonomies across datasets

Use Case

- Highly pathogenic avian influenza (bird flu)
- Has killed millions of birds in the rest of the world
- Particularly decimates seabird colonies
- Hasn't arrived in Oceania yet

... Can we model it?

Use Case: HPAI in NZ

- Vectors: Brown Skua, Kelp Gull, Giant Petrels, Snowy Sheathbills, Wandering Albatross
- Probable entry point: the Subantarctic Island of New Zealand.
- Spread: Seabirds > ducks & land birds & seals > poultry

Use Case: HPAI in NZ

- Show what bird flu has done elsewhere
 - Import and harmonize multiple community science datasets
 - Extrapolate what it might do here
-
- Document what issues we find during that process
 - Work on those.

Harmonizing Community Science Datasets to Model Highly Pathogenic Avian Influenza (HPAI) in Birds in the Subantarctic

1st Richard Littauer

Te Kura Mātai Pūkaha, Pūrorohiko

School of Engineering and Computer Science

Te Herenga Waka Victoria University of Wellington

Te Whanganui-a-Tara, Aotearoa

Wellington, New Zealand

ORCID: 0000-0001-5428-7535

2nd Kris Bubendorfer

Te Kura Mātai Pūkaha, Pūrorohiko

School of Engineering and Computer Science

Te Herenga Waka Victoria University of Wellington

Te Whanganui-a-Tara, Aotearoa

Wellington, New Zealand

ORCID: 0000-0003-4315-8337

Novel dataset

TABLE II
SUMMARY OF POPULATION STATISTICS FROM SGSSI

Species	Population	Mortalities	Mortality in %
Brown Skua	5,333	1,000	18.75
Kelp Gull	4,000	100	2.50
King Penguin	900,000	100	0.01
Wandering Albatross	3,425	50	1.46

TABLE III
SUMMARY OF POPULATION STATISTICS FROM MARION AND PEI.
MORTALITY RATES ARE ONLY KNOWN FOR MARION.

Species	Population	Mortalities	Mortality in %
Wandering Albatross	10,184	150	1.47
Brown Skua	1,225	80	6.53
King Penguin	134,000	120	0.09

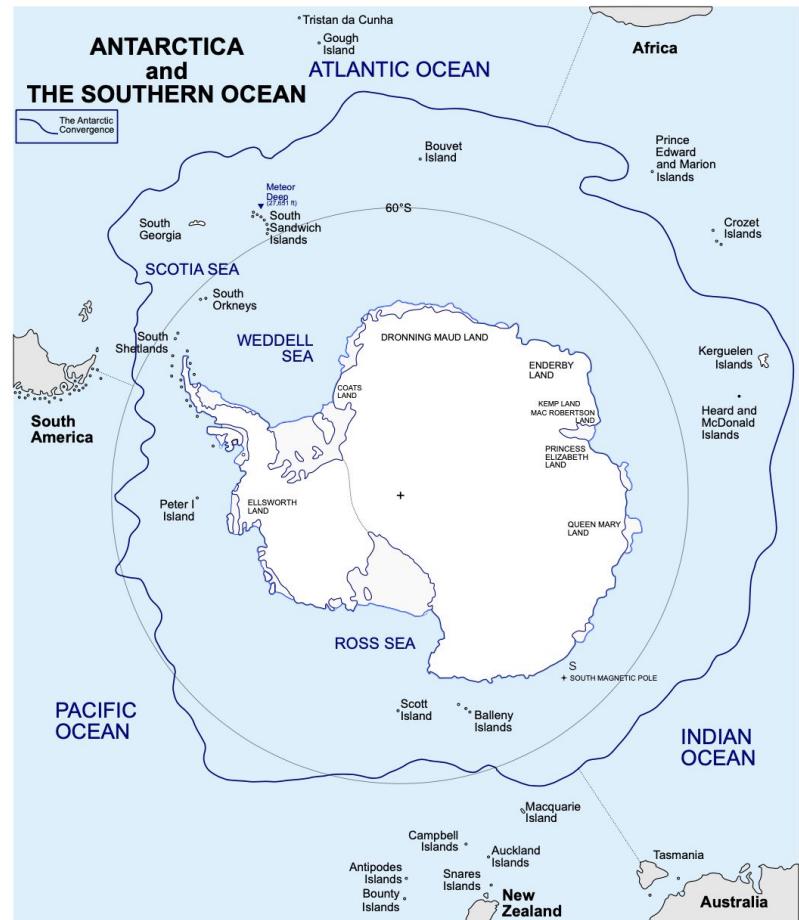


Fig. 1. The Subantarctic Islands.

Community science datasets

- eBird (API, regions)
- iNaturalist (Online interface)
- GBIF
 - API; 3b records, 100k datasets
- Royal Naval Birdwatching Society (RNBWS)
 - Entire db, 30k records
- At-Sea Observations of Seabirds 1969 to 1990, Tasman Sea, New Zealand and Australian waters (ASObs)
 - Entire db, 50k records

Community science datasets

TABLE I

SUMMARY OF DATASETS USED IN THIS STUDY. * AUCKLAND, CAMPBELL, THE SNARES (** ∩ SUBANTARCTIC IS.), AND MACQUARIE ISLANDS ARE TREATED AS GEOGRAPHIC SUBSETS OF THE SUBANTARCTIC ISLANDS REGION.

	All (Distinct)	Crozet	Falkland Is.	SGSSI	Kerguelen I	Marion & PEI	Auckland Is.*	Campbell I.*	The Snares**	Macquarie I.*	Subantarctic Is.
Latitude	-	-46.33	-51.80	-54.42	-49.24	-46.90	-50.74	-52.55	-48.02	-54.63	-54.50
Longitude	-	51.55	-59.30	-36.56	69.45	37.77	166.12	169.15	166.60	158.70	164.00
Radius	-	500	300	500	250	500	250	250	250	250	800
Observations											
eBird	319,652	1,132	182,225	68,393	5,644	11,226	15,668	9,075	10,558	13,580	51,032
iNat	27,184	460	10,392	6,088	1,209	806	3,340	1,809	2,210	1,039	8,229
GBIF	6,498	1,481	405	2,495	1,640	14	51	98	30	200	463
RNBWS	2,171	19	603	1,121	9	73	80	64	80	70	346
ASObs	215	0	0	0	0	0	0	0	0	61	215
All Records	355,720	3,092	193,625	78,097	8,502	12,119	19,139	11,046	12,878	14,950	60,285

Pipeline

- Download
- Clean
 - OpenRefine
 - Python
 - Normalize scientific names
 - Remove extraneous data
- GIS filtering
 - Reproject, filter by area
- Taxonomic reconciliation
- Filter to only target species
- Get population metrics from literature

Model

- Given a known population and mortality metrics in South Georgia or Marion, and given an assumed similar observation frequency in community science data, use that frequency to project population sizes and mortality rates in the Subantarctic Islands of New Zealand.
 - If we have population sizes for SINZ, use those instead.

Model 1 - eBird Only

- Brown Skua:
 - Known population of 500-1000 pairs
 - Predicted population of 3545
 - Predicted mortality of 326 (16.3%)
- Other birds: Minimal effect
 - King Penguin: Too small of a mortality rate to have any effect, population overestimation by 4x
 - Wandering Albatross: mortality rates of 325~ for SINZ
 - Black-browed Albatross: 400~ birds from a 40k population
 - Kelp Gull: On MacQuarrie, 150 birds, 4 deaths. SINZ: 3133 pop., 78 deaths.

Model 2 - eBird + iNaturalist

- Brown Skua:
 - Slightly higher: 3,565 birds, 632 mortality rate
- King Penguin:
 - Population estimates closer to reality on Crozets and Kerguelen
- Albatrosses:
 - Extrapolated populations much worse (as high as 1.4m on Campbell)
- Kelp Gull:
 - Almost identical numbers: 3106 pop., 78 deaths.

Model 3 - all of the datasets

- Generally unrealistic results.
- Some datasets too sparse (Kelp Gull, Skua, Penguin)
- Wandering Albatross: 7.5k on SINZ, mortality rate of 1.46%
- Black-browed Albatross: Overestimation of population by 10x

Discussion

- Skua population projections - especially Model 2 - was similar to known surveys.
- For known populations of birds, this system works well.
- Time may be a factor: 2353 out of 3565 Skuas may die from HPAI, if Great Skua numbers are used.

Discussion

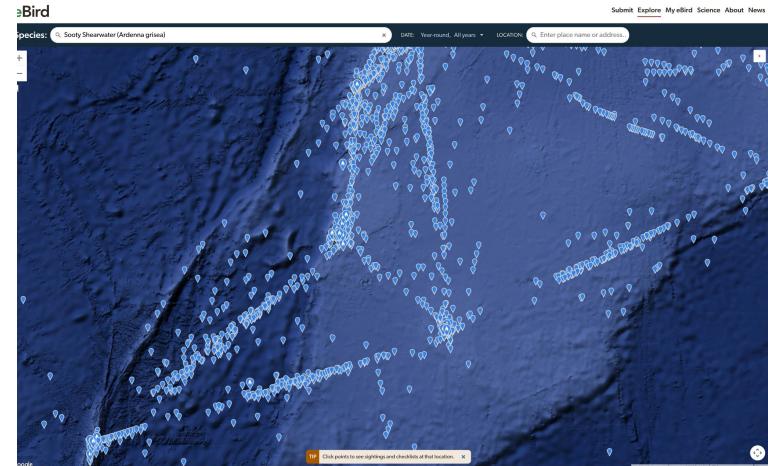
- For other birds, better data or data cleaning may be needed.
- ASObs and RNBWS are pelagic; eBird and iNat aren't.
- Sparse data projects are not robust.
- Biases should be accounted for:
 - Charismatic species
 - Island differences
 - At sea vs. at land
 - Species differences (lifespan of albatross vs. gulls; breeding differences...)

What's next?

- Better data cleaning
 - Account for known errors in eBird (am/pm, traveling/stationary, etc).
 - Use automatic clustering tools for checklists
 - Compare across community science platforms
- Integrate species-specific data into the cleaning process
- Build a specific cleaning pipeline in R and Python

What's next?

- Naval data
 - Point-to-point counts for known shipping routes
 - Provide better buffer areas for islands and coasts
 - Compare frequencies throughout the year
 - Allow arbitrary polygon downloads from eBird
 - Look for Ashmole's Halo in community science data



What's next?

- Taxonomic ontology reconciliation
 - Cleaning tool for dirty datasets (what names are missing?)
 - Comparison tool across taxonomies (NZ Birds Checklist vs. eBird)
 - Integrate with OpenRefine and Wikidata through Avibase IDs

What's next?

- Quantitative metastudy on community science data usage
- Focus on cleaning and data pipelines
- Focus on iNaturalist and eBird within Aotearoa New Zealand
- AFAIK: hasn't been done anywhere.

What's Next?

- Include terrestrial species in an HPAI model
- Integrate with the NZ Bird Atlas data

Immediate plan

- See if this paper can be resubmitted elsewhere
- Best practices for pelagic bird species using community science data
- Taxonomic resolution for BirdsNZ datasets and AviList
- Publish on RNBWS, ASObs, and Beach Patrol Scheme data

Connections and activities so far

- BirdsNZ Checklist Committee
- Nomenclature Committee for AviList
- eBird reviewer for Wellington
- Gave talk at BirdsNZ Conference in Auckland
- Organizing Kiwi Python, on the board for Python New Zealand
- MRSNZ: ECR for Royal Society Te Apārangi
- 2 grants from Wikimedia Aotearoa New Zealand
- 1 grant from Wellington Botanical Society
- Collaboration with: RNBWS and BirdsNZ Beach Patrol Scheme
- Connected with: Manaaki Whenua Landcare Research, AgResearch, NIWA, Te Papa, REANNZ, Genomics Aotearoa, National Library, and Pacific Seabirds

Papers in 2025

- Richard Littauer (2025). *Renaming a bill type for Tundra Swan (*Cygnus columbianus*)*. **Wilson Journal of Ornithology**, 1–3. <https://doi.org/10.1080/15594491.2025.2498780>
- Richard Littauer (2025). *Corrections for agreement in gender in Notonectidae (Hemiptera: Heteroptera) under Article 31 of the Code*. Zootaxa 5618(2): 287-288. <https://doi.org/10.11646/zootaxa.5618.2.8>
- Richard Littauer (2025). *A nomenclatural issue related to Lopdells' penguin *Archaeospheniscus lopdelli**. **Notornis** 72(1): 49–50. <https://doi.org/10.63172/777144pbjhsf>
- Richard Littauer (2025). *First revision of karoro Larus dominicanus antipodum (Bruch, 1853)*. **Notornis** 72(1): 51–53. <https://doi.org/10.63172/093633lrrwbq>
- Richard Littauer (2025). *Tristan Moorhen (*Gallinula nesiotis*)*, version 2.0. In G. M. Kirwan and N. D. Sly eds., Birds of the World. Cornell Lab of Ornithology, Ithaca, NY, USA. <https://doi.org/10.2173/bow.trimoo2.02>
- Richard Littauer (2025). *Corrections for grammatical agreement in Ophiidermatidae*. Zootaxa 5569(2): 400-400. <https://doi.org/10.11646/zootaxa.5569.2.11>

Other work

- 4 papers in press (1 ornithological, three nomenclatural)
- 11 other papers in review.
- Active member of SNAP; hosted two workshops
- Instructor for the Carpentries

Questions?