

Can We Trust the Power of the Crowd?

A Look at Citizen Science Data Quality from NOAA Case Studies

Laura Oremland
NOAA Fisheries

ESIP 2020 Winter Meeting
January 8, 2020



With Special Thanks to:

- Chris Bowser (Hudson River Eel Project)
- Amy Fritz, Bryant Korzeniewski, Matthew Menne (Cooperative Observer Program)
- Ken Knapp (Cyclone Center)
- Manoj Nair (CrowdMag)
- Noah Newman (CoCoRaHS)
- Lisa Natanson (Cooperative Shark Tagging Program)
- Katie Sweeney (Steller Watch)

Outline

Overview of NOAA Citizen Science

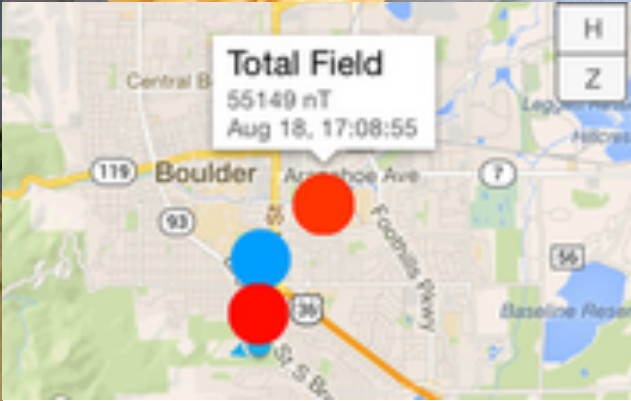
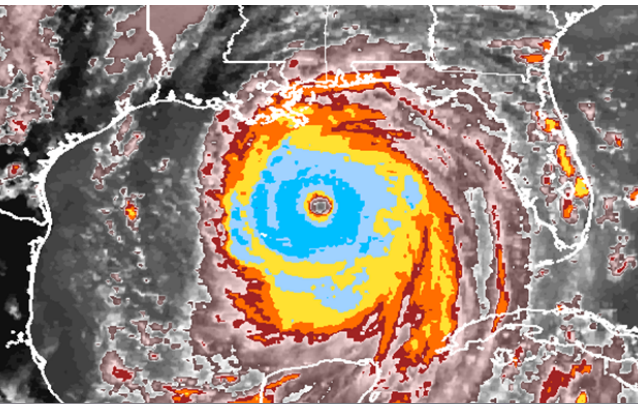
Case Studies

Common Themes

Recommendations

Discussion/Questions

NOAA Citizen Science




Case Studies

Unknown Participants

- 
1. Steller Watch
 2. Cyclone Center
 3. CrowdMag

Known Participants

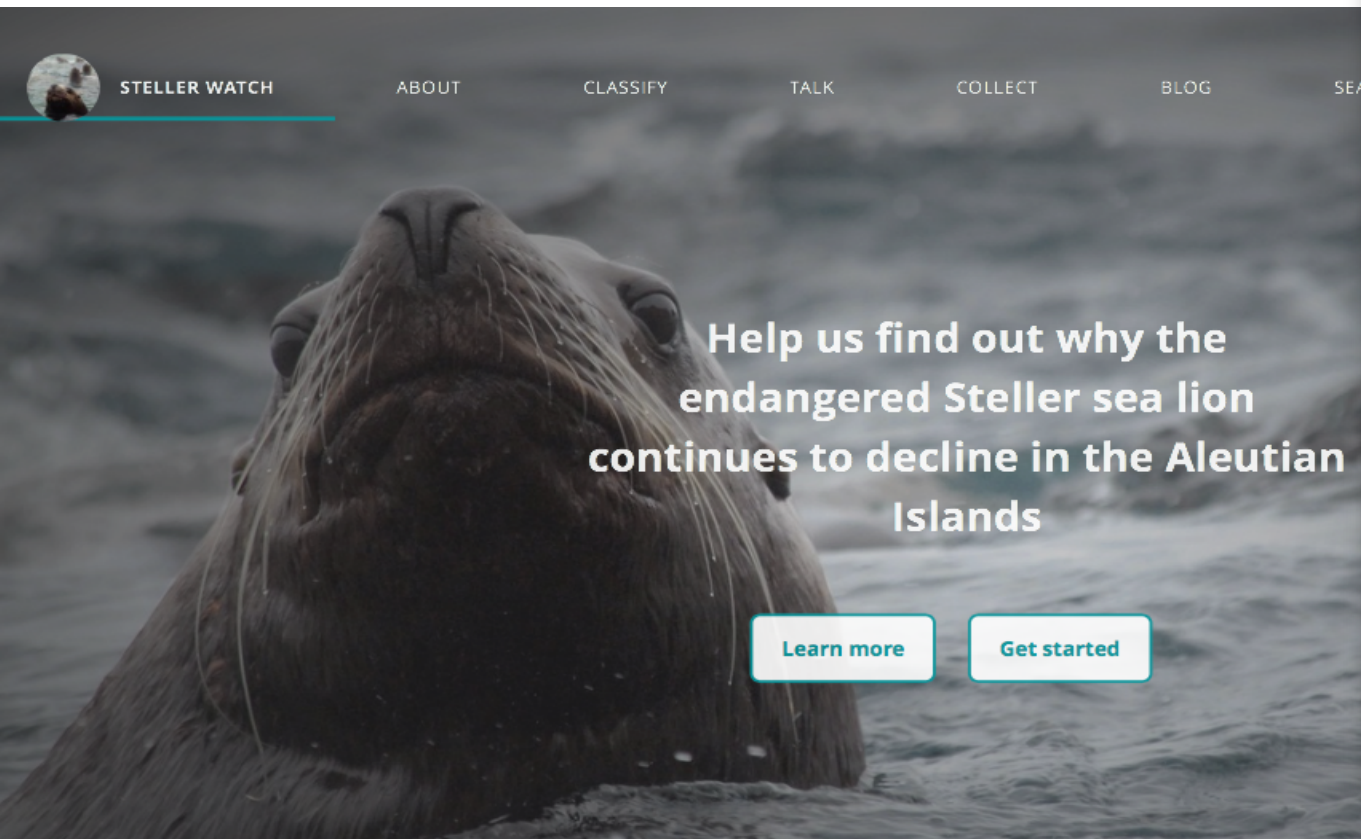
- 
1. CoCoRaHS
 2. Cooperative Observer Program
 3. Cooperative Shark Tagging Program
 4. Hudson River Eel Program

Unknown Participants

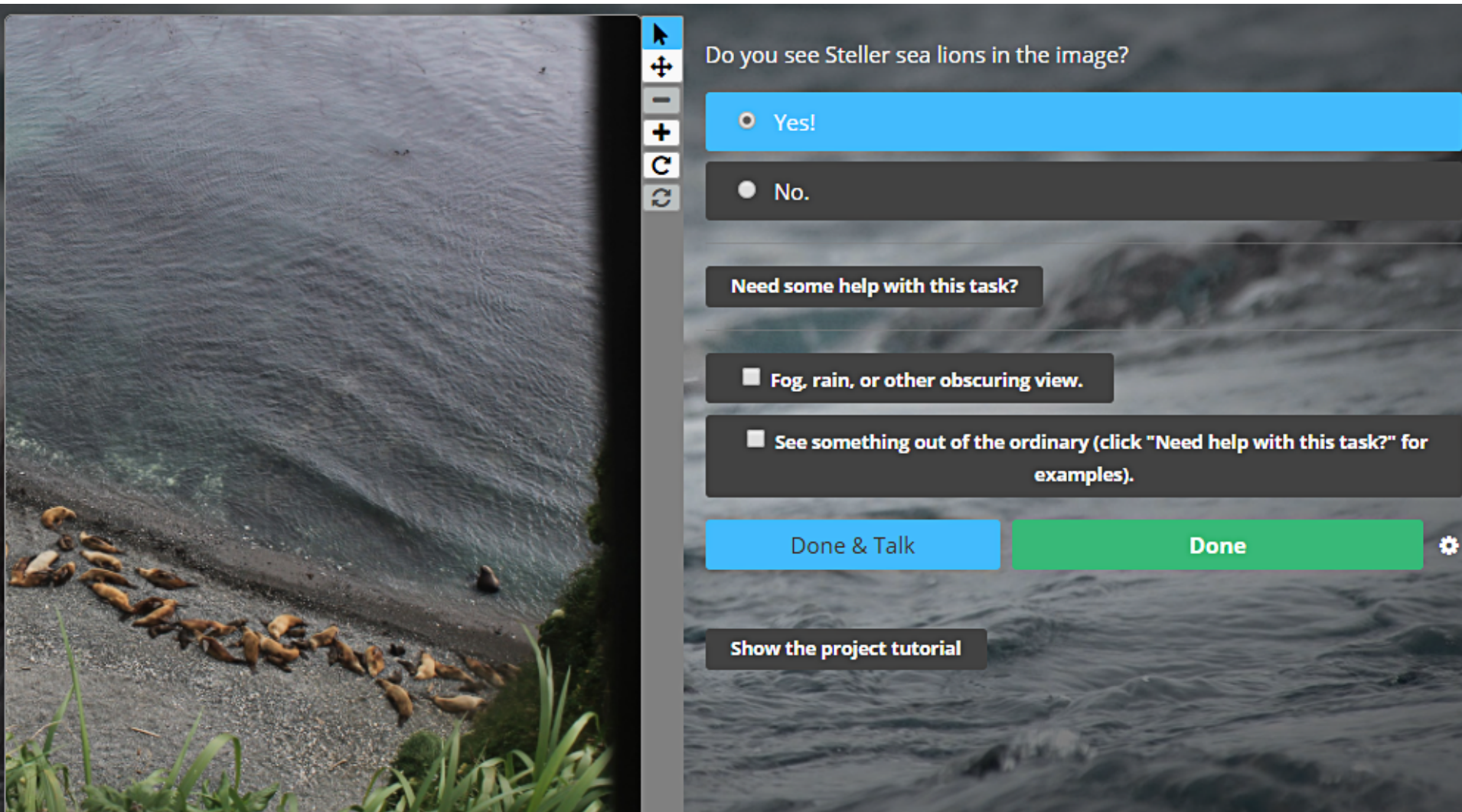


Steller Watch

- Participants help with photo ID research
- Uses Zooniverse platform



Binary Workflow



The screenshot shows a web-based interface for a binary workflow. On the left is a large image of a rocky beach with many Steller sea lions. To the right of the image is a vertical toolbar with icons for zooming in (+), zooming out (-), zooming to fit (C), and a refresh icon. The main panel on the right contains the following elements:

- A question: "Do you see Steller sea lions in the image?"
- Two radio button options: "Yes!" (selected) and "No.".
- A button: "Need some help with this task?"
- Two checkbox options:
 - ☐ "Fog, rain, or other obscuring view."
 - ☐ "See something out of the ordinary (click 'Need help with this task?' for examples)."
- Two buttons: "Done & Talk" (blue) and "Done" (green).
- A button: "Show the project tutorial" (dark grey).
- A settings gear icon in the bottom right corner.

3 Option Workflow



TASK

TUTORIAL

Can you read marks on any individual(s)? **You will not need to report any marked animal sighting with the Done & Talk button or Talk forum.**

Yes! I **can** read the letter/symbol or at least one number on any individual(s).

Yes, but I **cannot** read the letter/symbol or any numbers on any individual(s).

No, I see no marked sea lions.

NEED SOME HELP WITH THIS TASK?

Done & Talk

Done

Data Quality



Methods

- Pilot group test responses
- Binary Workflow:
7 reviewers/image (5 if all in agreement)
- 3 Options Workflow:
13 reviewers/image

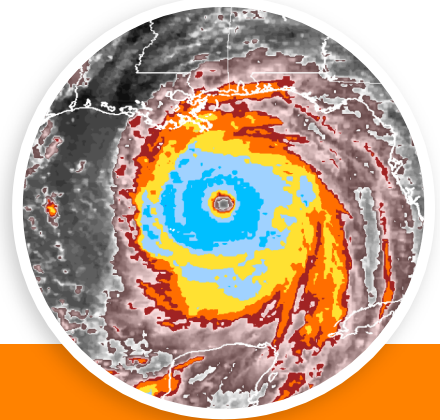
Lessons Learned

- Binary (yes/no) workflows best;
Saves analysis time
- Pilot group helpful
- Is cit sci best approach?



Cyclone Center

- Participants answer questions about satellite imagery of tropical cyclones
- Uses Zooniverse platform



Welcome to

Cyclone Center

Tropical cyclones are still a mystery.
We need your help to decipher them.

What is Cyclone Center?

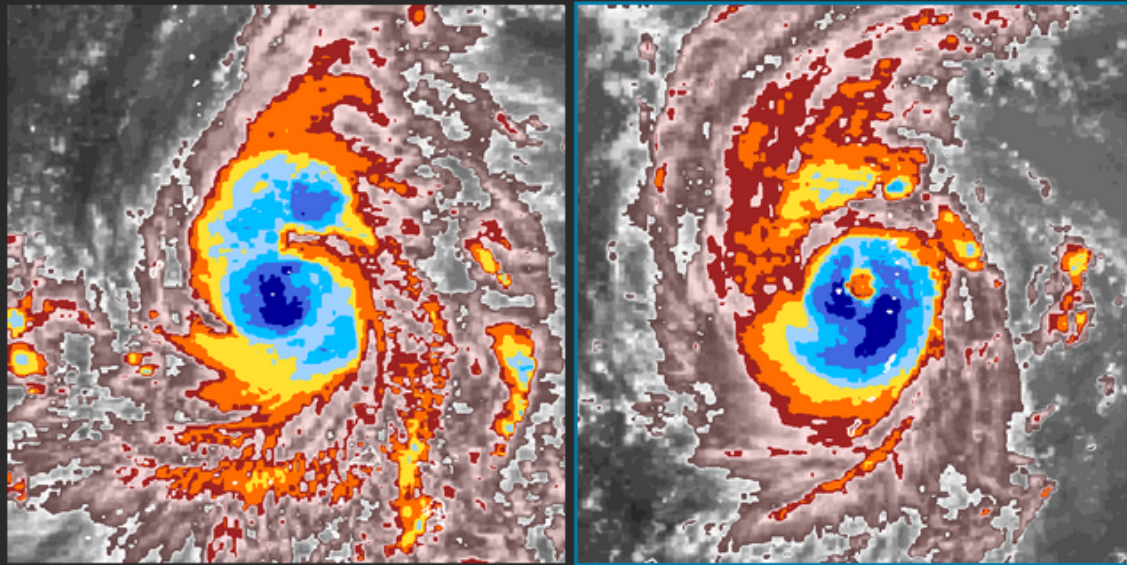
The climatology of tropical cyclones is limited by uncertainties in the historical record. Patterns in storms imagery are best recognized by the human eye, so we need your help analyzing these storms.

Are you ready to start investigating?

[Get Started](#)[Learn More](#)

Sample Question 1

Investigations—Classify the cyclone by answering the questions below.



Choose the storm image that appears **stronger**.

Left

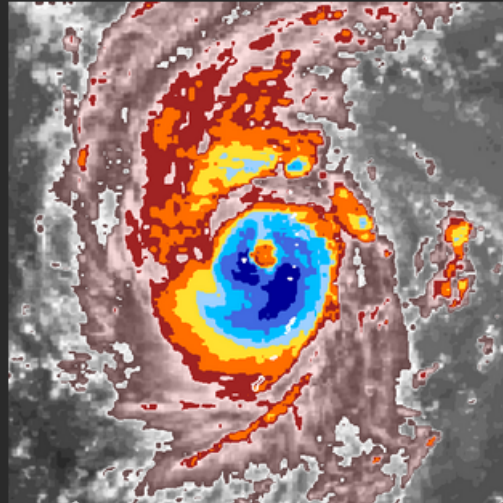
They're about the same

Right

Continue

Sample Question 2

Investigations—Classify the cyclone by answering the questions below.



Pick the cyclone type, then choose the closest match.

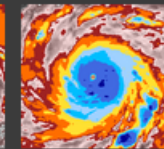
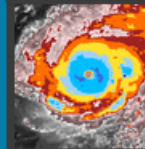
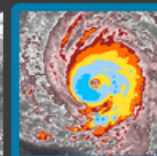
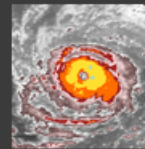
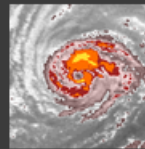
Eye

Embedded

Curved

Shear

Other



Data Quality

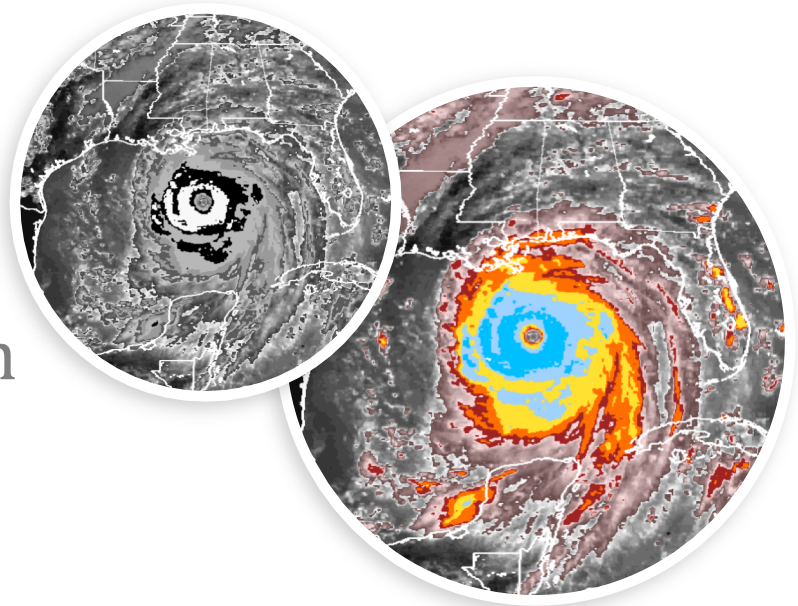


Methods

- Reviewers/image: 30 to 10
- Expectation-Maximization (EM) Algorithm

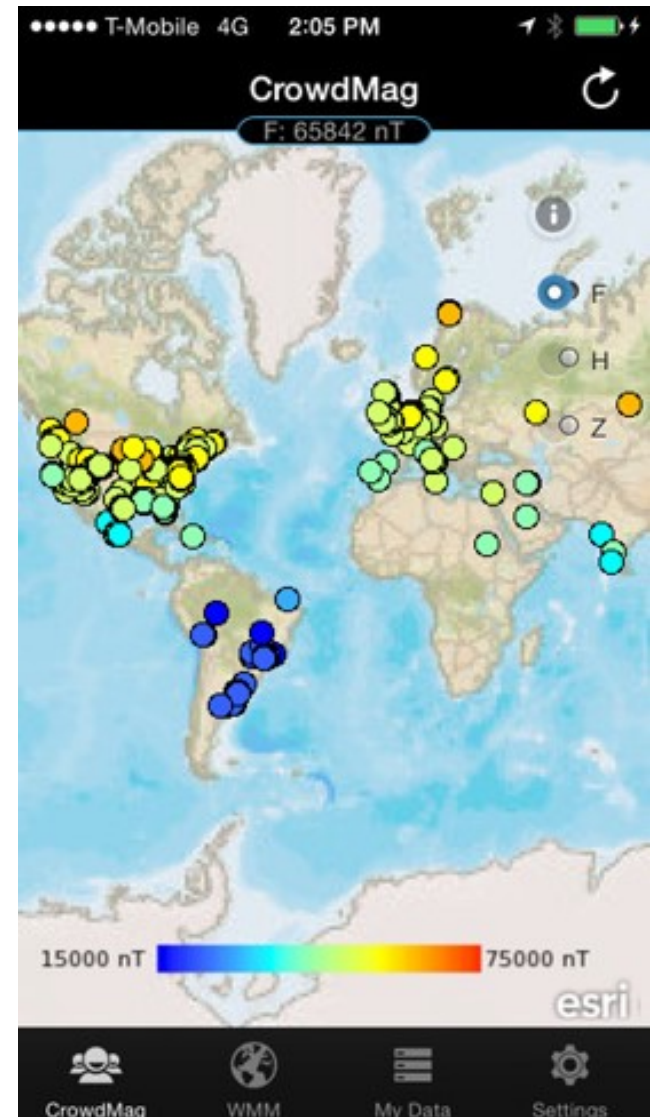
Lessons Learned

- Do analysis early to maximize your “clicks”
- Short-term goals
- Simple works best
- Is cit sci best method?
- Discovery vs classification
- Invest in good tutorials



CrowdMag

- Participants provide magnetic field measurements via smart phones
- Help fill gaps in geomagnetic data coverage, improve models of Earth's magnetic field



Data Quality



Methods

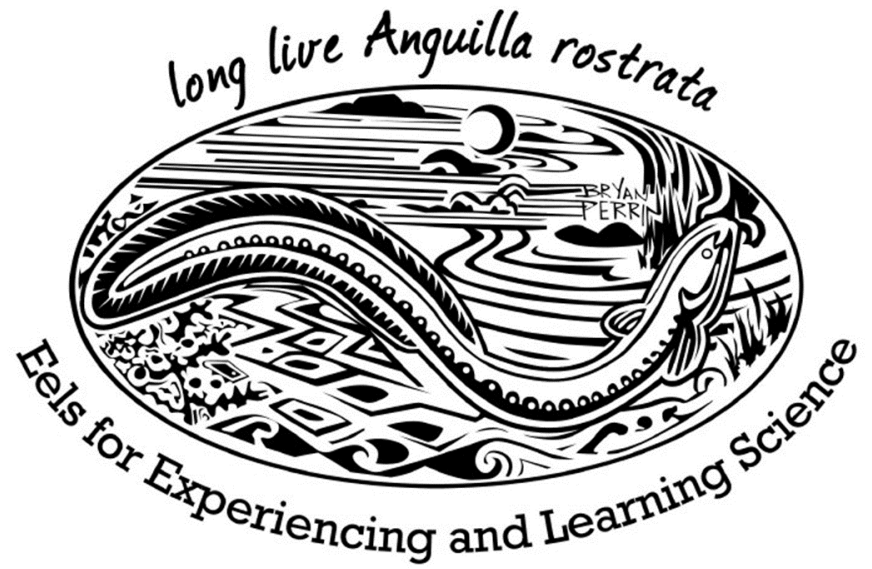
- High volume data collection
- Use median
- Compare data to model
- Compare surrounding data

Lessons Learned

- Data quality improved with large numbers
- Can generate localized maps with greater resolution
- Community engagement



Known Participants



CoCoRaHS

- Participants measure and map precipitation
- Provides high quality precipitation data for use by NWS, emergency managers, hydrologists, many others



COMMUNITY COLLABORATIVE RAIN, HAIL & SNOW NETWORK
"Because every drop counts"

Welcome to CoCoRaHS! "Volunteers working together to measure precipitation across the nations."

Data Quality



Methods

- Data entry (range checks)
- Post entry (manual, automated)
- Ticketing system
- Training, video

Lessons Learned

- Most errors: reporting, not measuring
- Volunteers appreciate being contacted
- Expect mistakes, especially typos
- Value of metadata
- Future, AI/machine learning

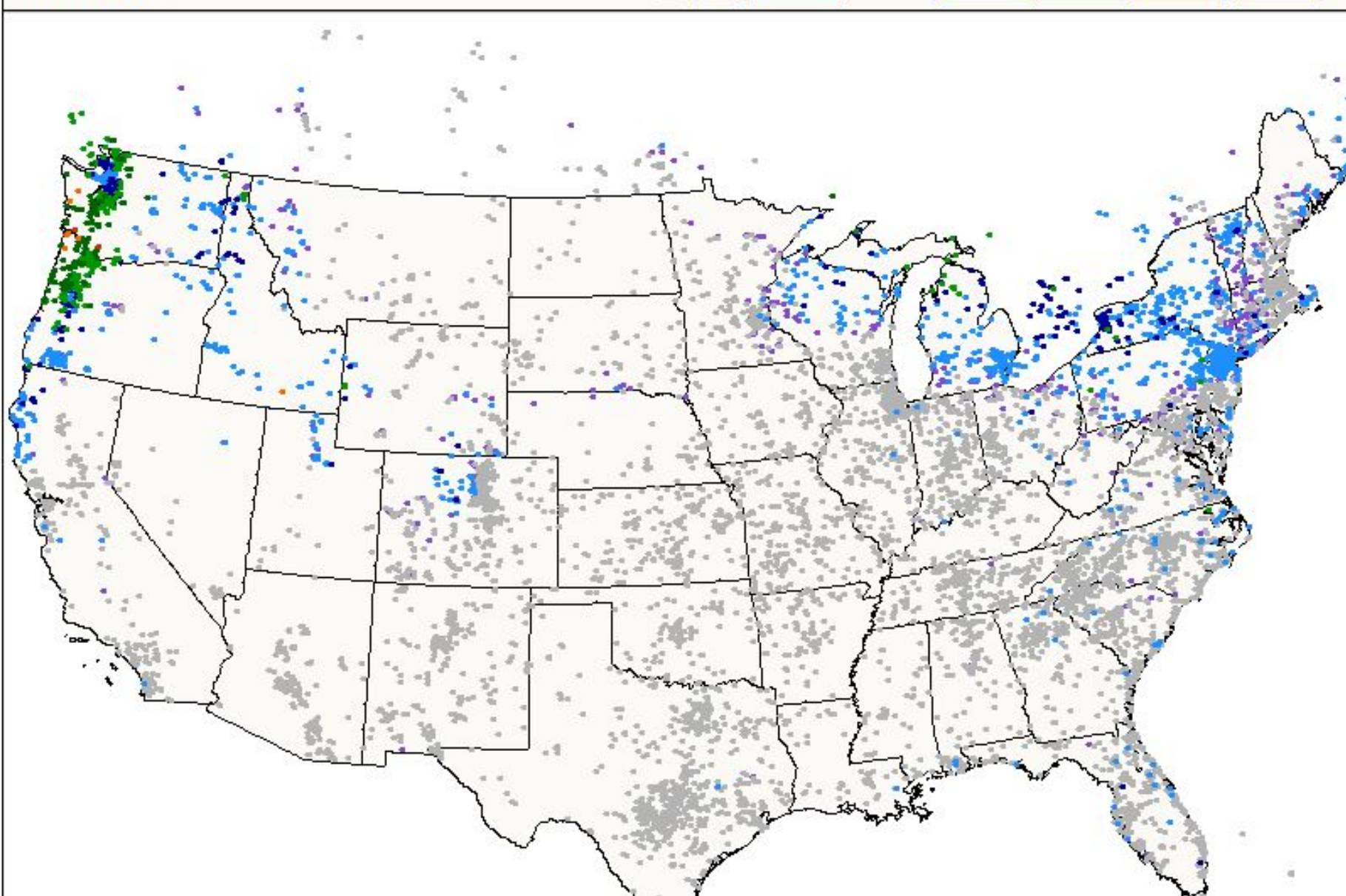


Map Type		Map Location		Date	Colors	
Precipitation ▼	National ▼	No State Selected ▼	1/6/2020	Standard ▼	Get Map	

Daily Precipitation (inches x.xx), for the 24 hour period ending ~7:00 am

USA 1/6/2020

0.0 Trsce 0.01 - 0.15 0.16 - 0.30 0.31 - 0.75 0.76 - 1.80 1.81 - 2.70 2.71 - 2.99



Cooperative Observer Program

- Participants measure daily air temps, precip
- Supports U.S. climate records and weather forecasts, models, and warnings



Cooperative Weather Observer

Data Quality



Methods

- Data entry (WxCoder3)
- Monthly closeout reviews (observer, WFO)
- NCEI automated checks (spatial, climatological, etc)
- Ticketing system

Lessons Learned

- More extremes make QA/QC more difficult
- >100 years supporting forecast validation, climate, drought monitoring



Cooperative Shark Tagging Program

- Participants (mostly fisherman) tag sharks
- Provides information on shark biology, movements, migrations, abundance, age and growth, mortality, and behavior



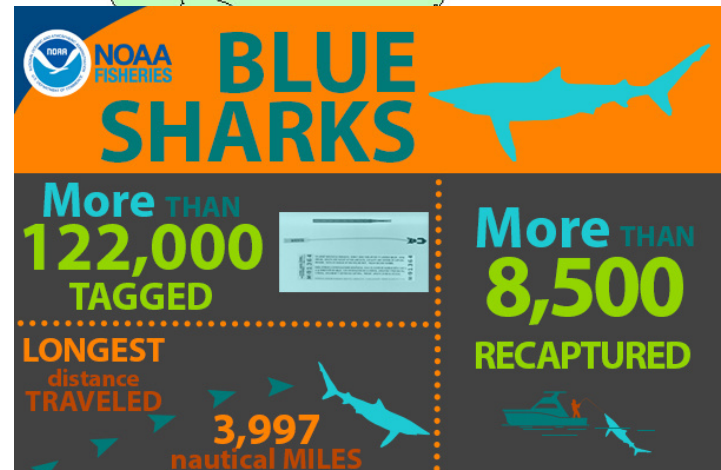
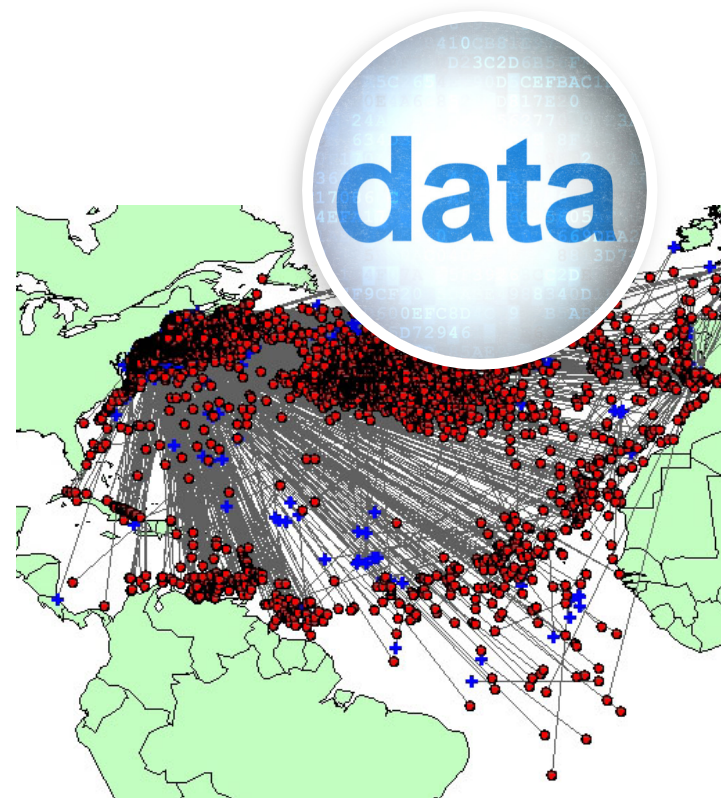
Data Quality

Methods

- 1st Step: Follow-up with participants (rapid, pictures)
- 2nd Step: Database quality controls (data entry, maintenance)

Lessons Learned

- Volunteers happy to be contacted
- Integrated Mark/Recapture Database System (I-MARK) improved QA/QC
- Volunteer incentives



Hudson River Eel Project

- Participants catch, count, release American eels, a species in decline
- Provides data on eel population, biology



Data Quality

Methods

- Quality Assurance Project Plan
- Simple, seasonal participation
- Volunteer training, oversight
- Staff manual data review (x2)

Lessons Learned

- Need paper, electronic data records (value of metadata)
- Simple procedures for volunteers improve data quality
- Volunteer participation (seasonality, internal incentives)



Common Themes



Simplicity enhances data quality

Value of metadata

Paired manual and automated checks

AI/Machine learning and citizen science

Known vs. Unknown Participants

Unknown Participants

- Pilot test
- The value of simple questions
- Is citizen science the best approach
- Higher data volume and data quality

Known Participants

- Value and benefits to following up with volunteers on questions
- Reporting errors most common

Recommendations

Keep it simple

Pilot test early

Combine manual and automated checks

Engage volunteers with questions (soon)

Ask if citizen science is the best approach

Questions?

Laura.Oremland@noaa.gov

