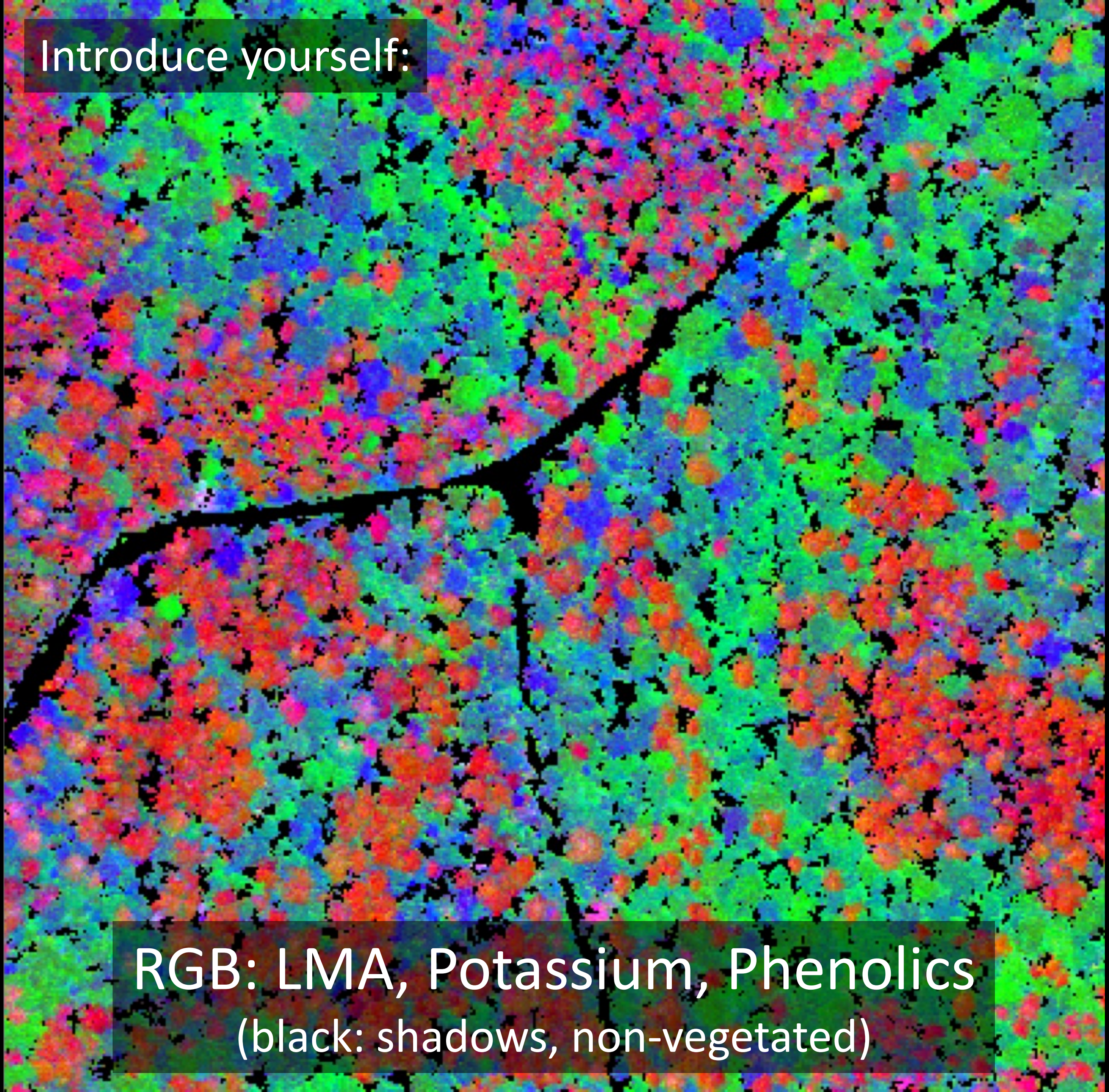
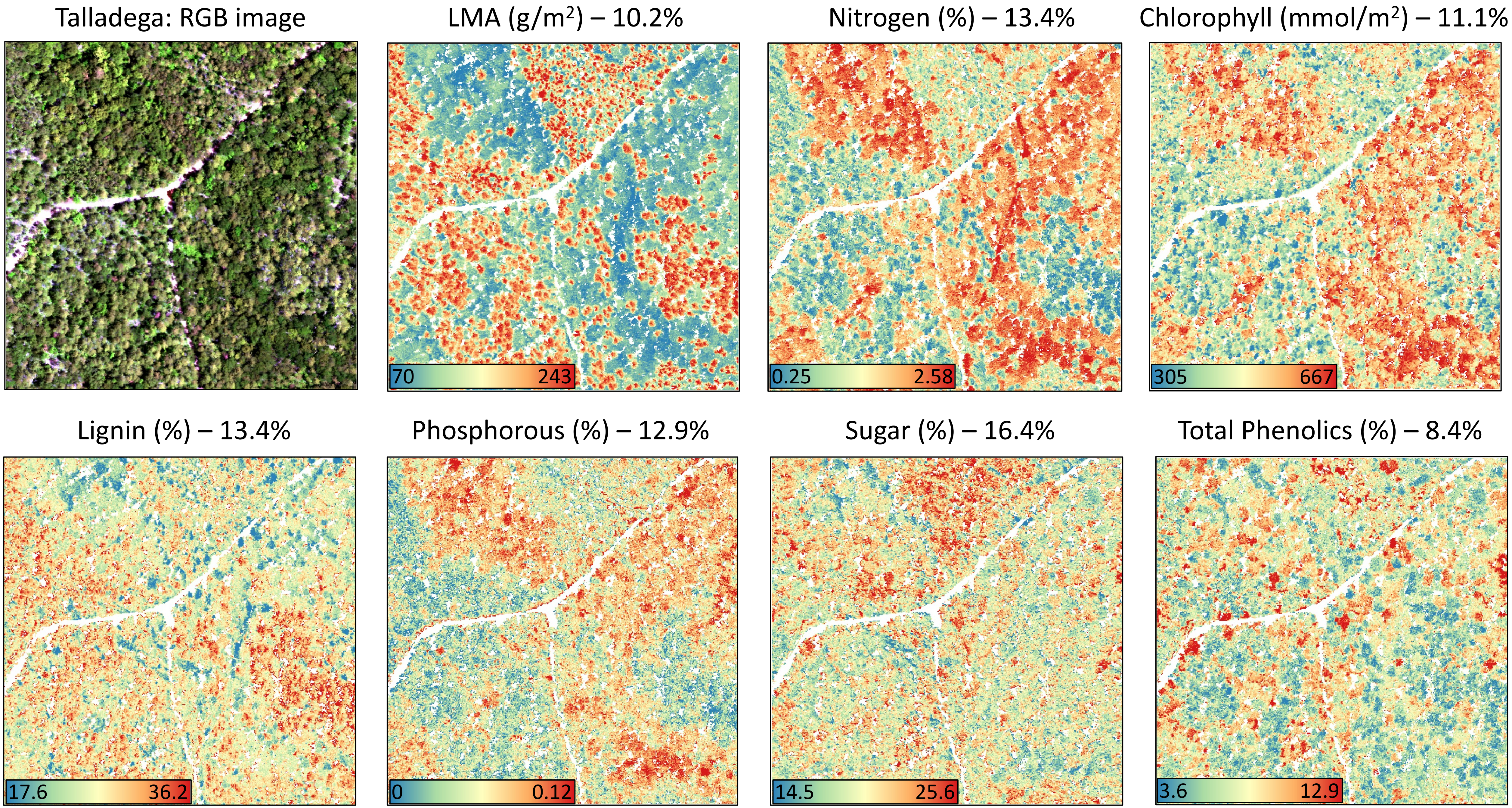
NASA and NEON Airborne Imaging Spectroscopy:
Characterizing Fine-Scale Vegetation Function at the Continental Scale

Phil Townsend
University of Wisconsin-Madison

EnSpec
Environmental Spectroscopy Laboratory



Traits	R^2 (val)	% RMSE
% Carbon	0.63	13.06
Aluminum	0.55	10.4
Carotenoids (area)	0.57	11.92
Carotenoids (mass)	0.59	13.3
Cellulose	0.6	12.11
Chlorophyll (area)	0.64	11.14
Chlorophyll (mass)	0.58	13.85
Isotopic δ13C	0.63	13.12
EWT	0.62	14.13
Flavonoids	0.53	14.5
Iron	0.51	15.03
Lignin	0.57	13.37
Leaf Mass per Area	0.78	10.2
Manganese	0.62	12.99
Nitrogen	0.55	13.44
Phosphorous	0.5	12.93
Potassium	0.66	10.74
Starch	0.61	12.73
Sugar	0.45	16.37
Phenolics	0.82	8.4
Water (percent)	0.58	11.86
Zinc	0.54	16.12



Talladega: RGB image
LMA (g/m²) – 10.2%
70
243
Nitrogen (%) – 13.4%
0.25
2.58
Chlorophyll (mmol/m²) – 11.1%
305
667
Lignin (%) – 13.4%
17.6
36.2
Phosphorous (%) – 12.9%
0
0.12
Sugar (%) – 16.4%
14.5
25.6
Total Phenolics (%) – 8.4%
3.6
12.9

ABoVE

Arctic-Boreal Vulnerability Experiment

AVIRIS-NG imaging + ground sampling

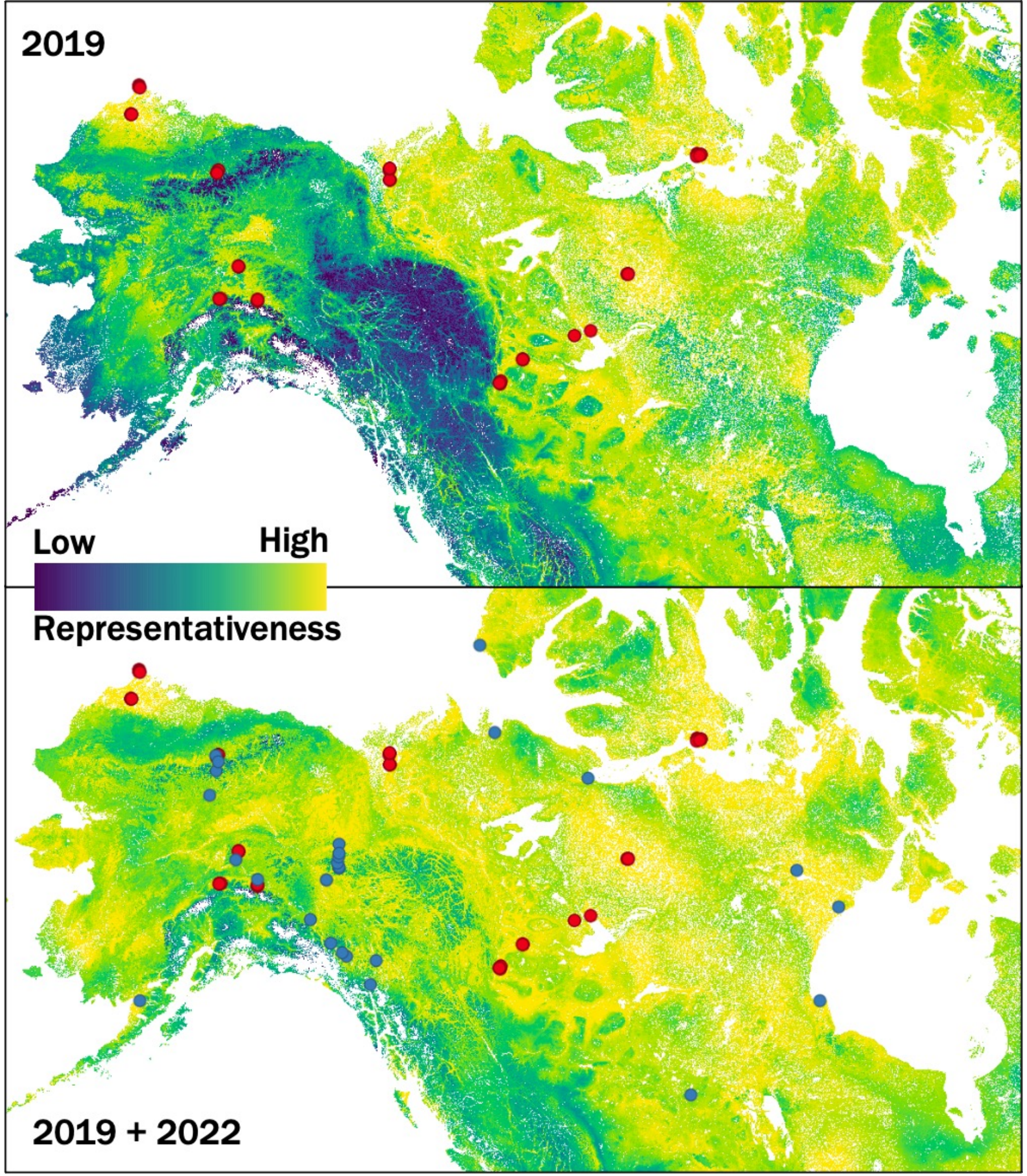


Figure: Ryan Pavlick, Morgan Dean and Kyle Kovach

Which airborne data sets do you use?

- AVIRIS-NG
- AVIRIS-Classic
- MASTER
- NEON AOP
- My lab generates its own imagery from NASA-funded research:
  - Cessna HySpex
  - UAV HySpex Mjolnir (heavy-lift)
  - UAV lidar
  - Various UAV multispectral and FLIR
- We also generate products that we wish to distribute (or are funded to):
  - ABoVE: BRDF corrections, trait maps

How do you access this data?

- In the past:
  - Receive hard drives
  - Direct download (AVIRIS from JPL)
- Currently:
  - Via API (NEON)
- Our own data are stored on local servers or in the cloud
- Issue with all of this:
  - Data storage costs
  - We largely work in code and only store final products
- We have not yet provided data to a DAAC to distribute
  - JPL connection

How are you using this data?

- We link field and image data to calibrate and validate models/maps of foliar functional traits (chemical, physiological, morphological), functional diversity, and other properties, e.g., related to stress, pests and pathogens.
- We have an extensive Python image processing workflow (HyTools) that we use to generate products (e.g., BRDF corrections, traits):
  - https://github.com/EnSpec/hytools
  - Has been adopted by quite a number of people for processing AVIRIS-NG imagery
- Airborne products to baseline/validate satellite products
  - Global scale functional trait mapping (SBG)

What works well for you

- We tend to work with “all” images
- Accessing data through API (NEON)
- Direct downloading is also fine
- High-throughput computing on our end
- Downloading datasets with the least amount of authentication, and without any GUI (batch processing)

What pain points do you find with data access or data use?

- The data volume is huge. It is difficult to download only a small spatial subset from a huge dataset. Sometimes, the product we need is bundled with other products, maybe for better organization. That makes us download many datasets we don’t need.
  - Need to download the entire tarred radiance data to get OBS/ORT files (which are small)
  - Data subsetting when accessing (cloud/on-demand processing would solve)
- Consistent data structure is preferred (all JPL stuff, may not be a DAAC issue)
  - Data type
  - Interleaving
  - Float vs integer
  - Nodata values
- Updates to data (versioning, notifications on recent or forthcoming updates)
  - Versioning issues: new processing implemented for one data set but older data not reprocessed
- Tracking processing workflows and provenance
  - Citations are not always complete
  - Code usually not open-source and therefore cannot be replicated
  - Not a DAAC problem, but data storage (hot, warm, cold) is costly with large volumes of data
  - Cloud storage
- On-demand processing / cloud computing options

What do you wish you could do but can’t?

- Inexpensive cloud storage
- On-demand processing
- Cloud computing
- We are well set up to process large volumes of data: my group has options
  - Others are not – there are barriers to access

Do you have any suggestions for improvement?

- Address the pain points 2 slides back
- Easily discovered documentation, especially for derived products
- Quality control information front and center (so we don’t waste our time, and exclude it when automating processing)
- We need to develop tools that lower barriers to access for non-expert users.
- Specifically, we need to distribute products that are more developed/mature for the non-experts (example BRDF)

Have you tried to use data in the cloud?

- Yes

Have you tried to use data in the cloud?

- Yes
- Well, no, not me personally in any meaningful way, but my lab does extensively
- Most of the data we use does not have this option
  - GEE gets modest use
- Note: development of GeoSPEC for on-demand processing

What support do you need from ADMG?

- I do not know what support you provide!
- Distribution of data we generate in support of NASA funded research
- Tutorials for data handling
  - Code repositories, jupyter notebooks
  - Readthedocs
  - Etc.

Thank you for asking!