

The background of the slide is a dramatic, dark blue scene of a massive wooden ark floating on a turbulent sea. The ark is illuminated from within, with a row of warm, yellow lights visible along its upper edge. The sky is dark and filled with rain, with several bright, jagged lightning bolts striking down. The overall mood is one of impending disaster and urgency.

# Biogeoscience observing system *deluge:*

## Are we ready for the flood?

Ankur Desai, University of Wisconsin-Madison

B41B-02

AGU Fall 2019, Monday 1610-1620, MW 2003

Image: <https://insider.pureflix.com/>

# Acknowledge

- PEcAn project (NSF DBI-1457897)
  - Mike Dietze, Istem Fer, Bailey Murphy, Katie Zarada
- CHEESEHEAD (NSF AGS-1822420)
  - Stefan Metzger, Sreenath Paleri, Dave Durden, Ke Xu
  - Cyverse: Tyler Swetnam
- DOE Ameriflux
  - Kim Novick, George Burba
- Ben Bond-Lamberty, Jack Williams, many conversations from many of you, especially students...





Photo: A Desai



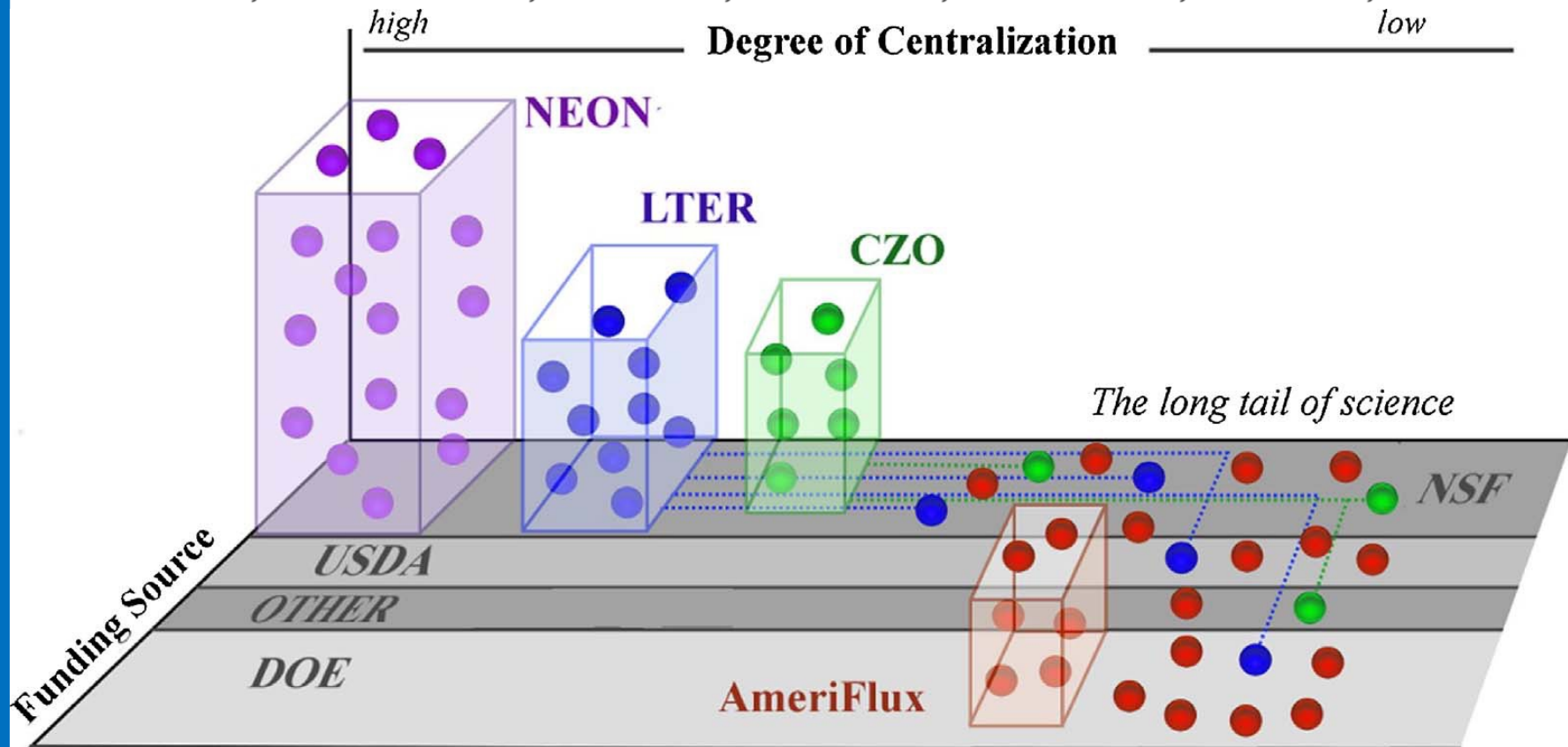
# Huge Ecology?

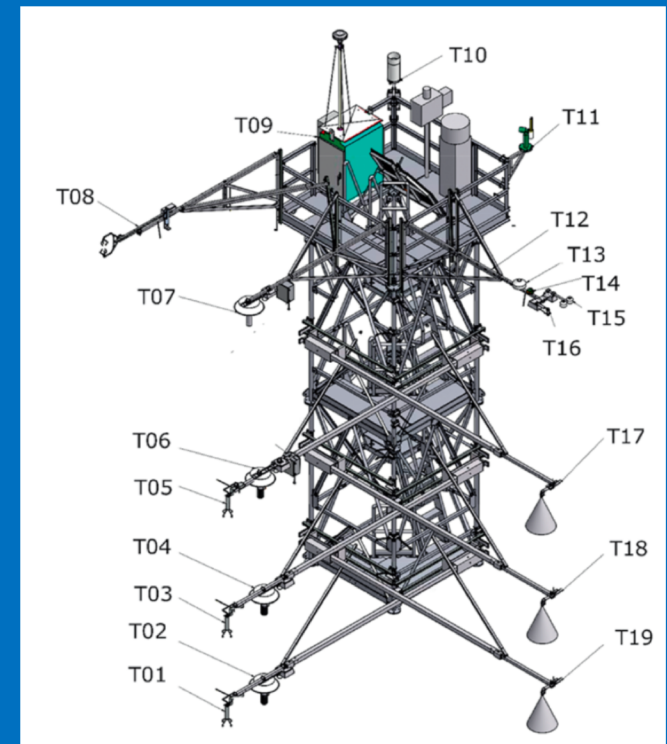
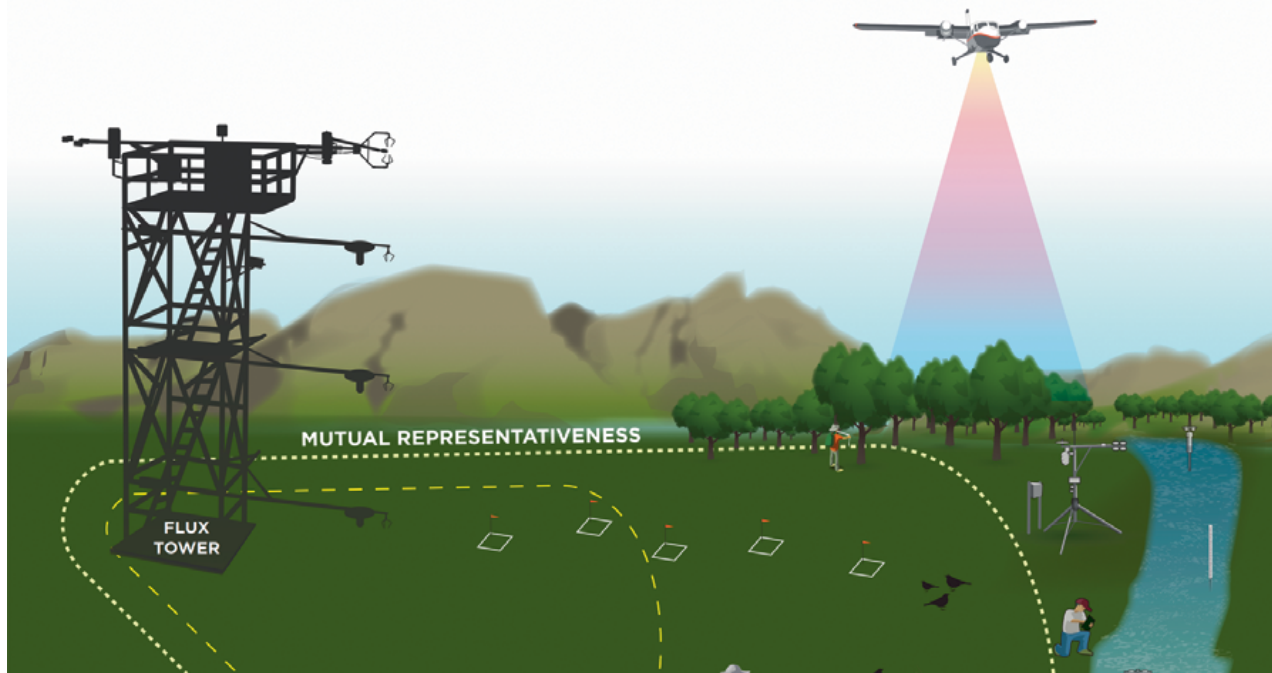




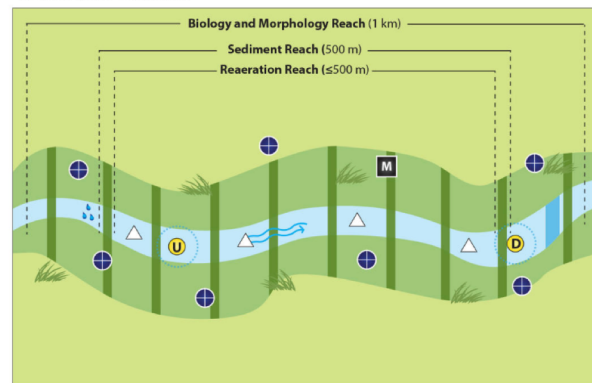
## The AmeriFlux network: A coalition of the willing

K.A. Novick<sup>a,\*</sup>, J.A. Biederman<sup>b</sup>, A.R. Desai<sup>c</sup>, M.E. Litvak<sup>d</sup>, D.J.P. Moore<sup>e</sup>, R.L. Scott<sup>b</sup>, M.S. Torn<sup>f</sup>





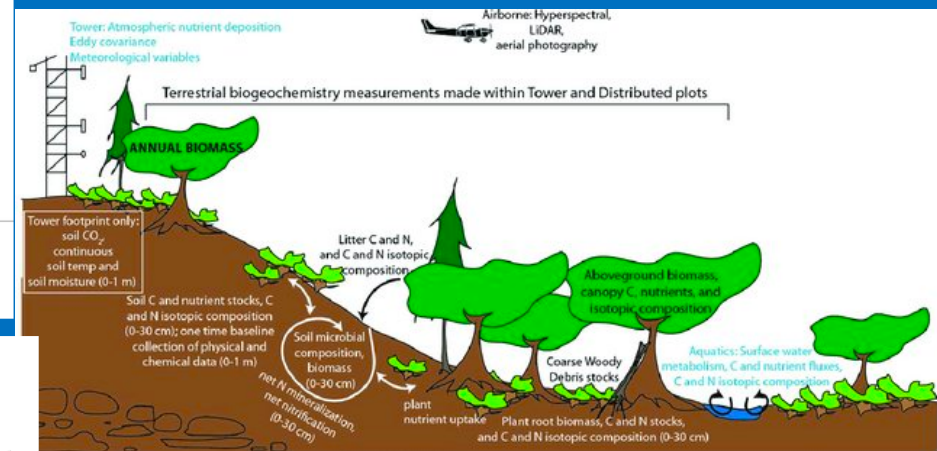
#### Wadeable Stream



#### Legend

- Sensor Station
- Water Chemistry Sampling
- ⊕ Groundwater Well
- M Meteorological Station
- Riparian Assessment
- ⬆ Reaeration Drip
- △ Reaeration Sampling

Note: Fish, sediments, macroinvertebrates, plants, and macroalgae are sampled based on site-specific habitats and are not identified in the figures.



> 62,000 samples

> 400 taxa



Distribution of samples by collection type.



Distribution of samples by top 5 determined taxa.

Metzger et al., 2019, Bulletin AMS; Hinckley et al., 2016, Ecosphere; neonscience.org

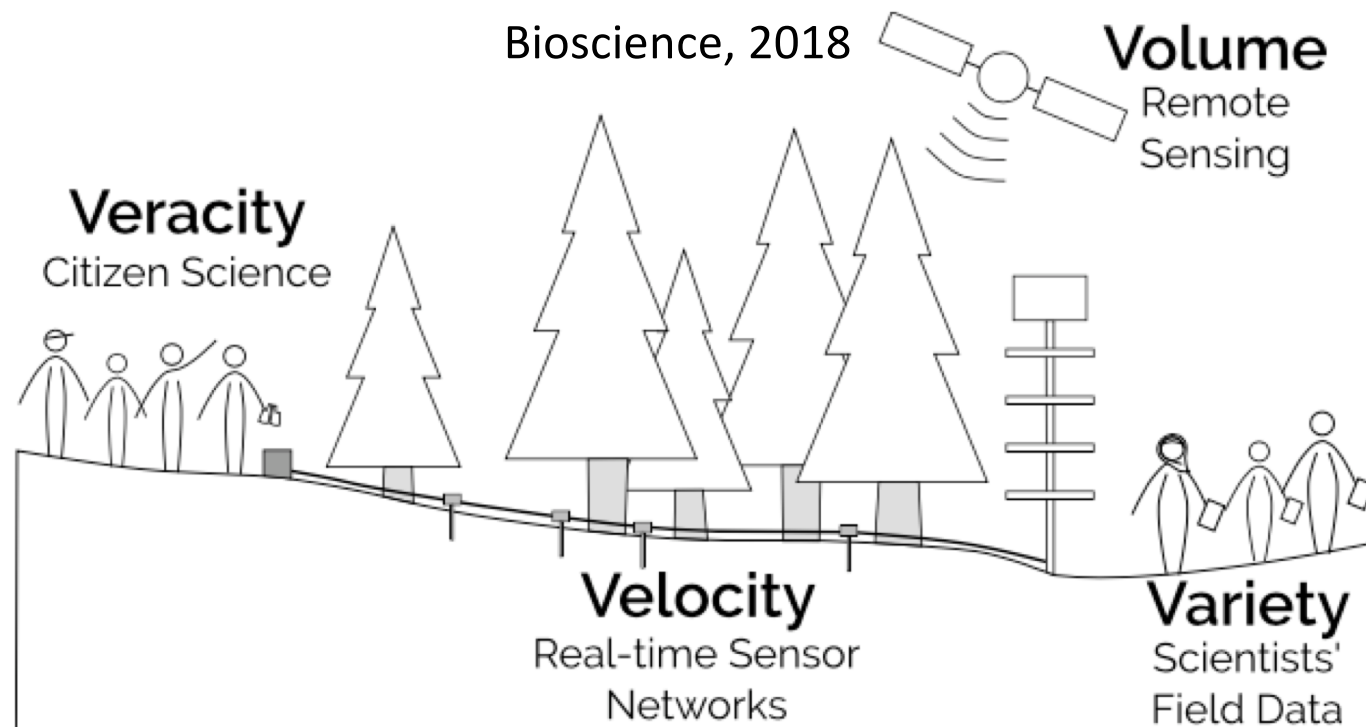


# Huge Ecology!

- The past 100 years of ecology went from its birth as a discipline, with theory tested primarily from small experiments in one's “backyard”, to a modern science that sits at the center of key biogeoscience challenges on land, freshwater, and ocean management, global climate change, biodiversity and habitat loss, restoration, ...
- Its complexity and data volume will only accelerate in next 100...

# Situating Ecology as a Big-Data Science: Current Advances, Challenges, and Solutions

SCOTT S. FARLEY, ANDRIA DAWSON, SIMON J. GORING AND JOHN W. WILLIAMS





# Huge Ecology!

- What elements of Biogeosciences must be nurtured to make sure we are capable?
- I will talk about 3 (in 5 minutes!):
  - Promoting an open science culture
  - Enabling scalable and reproducible analysis
  - Removing the wall between models and observations

# Promoting an Open Science Culture



CHEESEHEAD photos: <https://groups.google.com/forum/#!forum/cheesehead>



# What is open science?

- Open, FAIR data (and models!)
- Citable code
- Transparency in methods (“pre-registration”)

<https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Data-for-Authors>

**AGU100** ADVANCING EARTH AND SPACE SCIENCE

SITE



# Data for publication

Guidelines to support author compliance with open data standards

**WHAT IS NEEDED**

The data that supports your research and the visualizations in your paper must be deposited in a trusted repository that practices t for your data, please consider the following prioritization. We recommend a repository that specializes in the data for your scientific will be interoperable and reusable. If that is not available for your data type, next is your institutional repository, your computing ce

Data availability statements

Data citations

Repository selection

Modeling data

**Data availability statement examples**



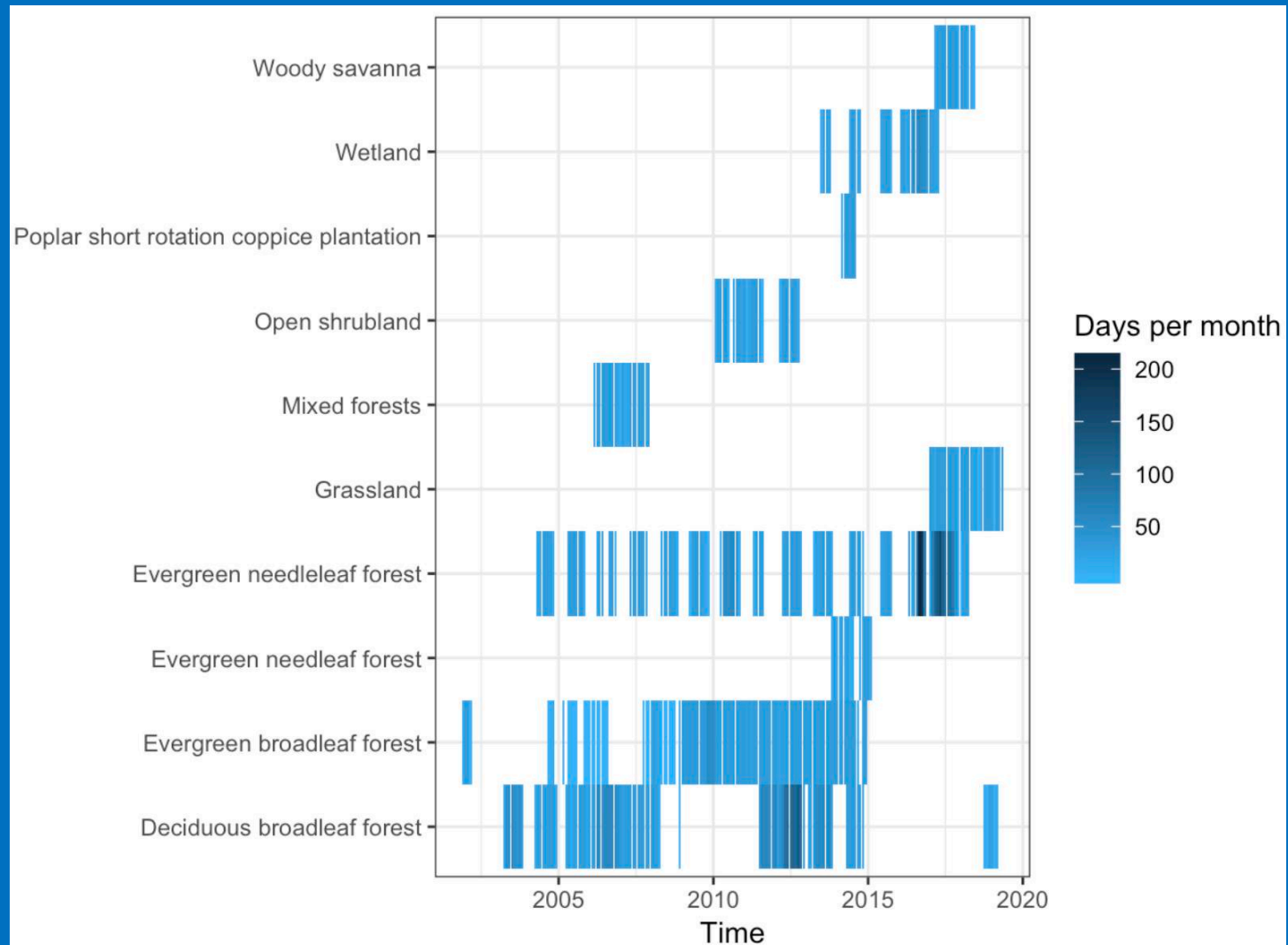
# Citable code

- <https://guides.github.com/activities/citable-code/>
- Code licenses matter!
  - <https://choosealicense.com/licenses/>

# #COSORE (Bond-Lamberty)

<https://github.com/bpbond/cosore>

Open Science: Lamberty et al., 2016



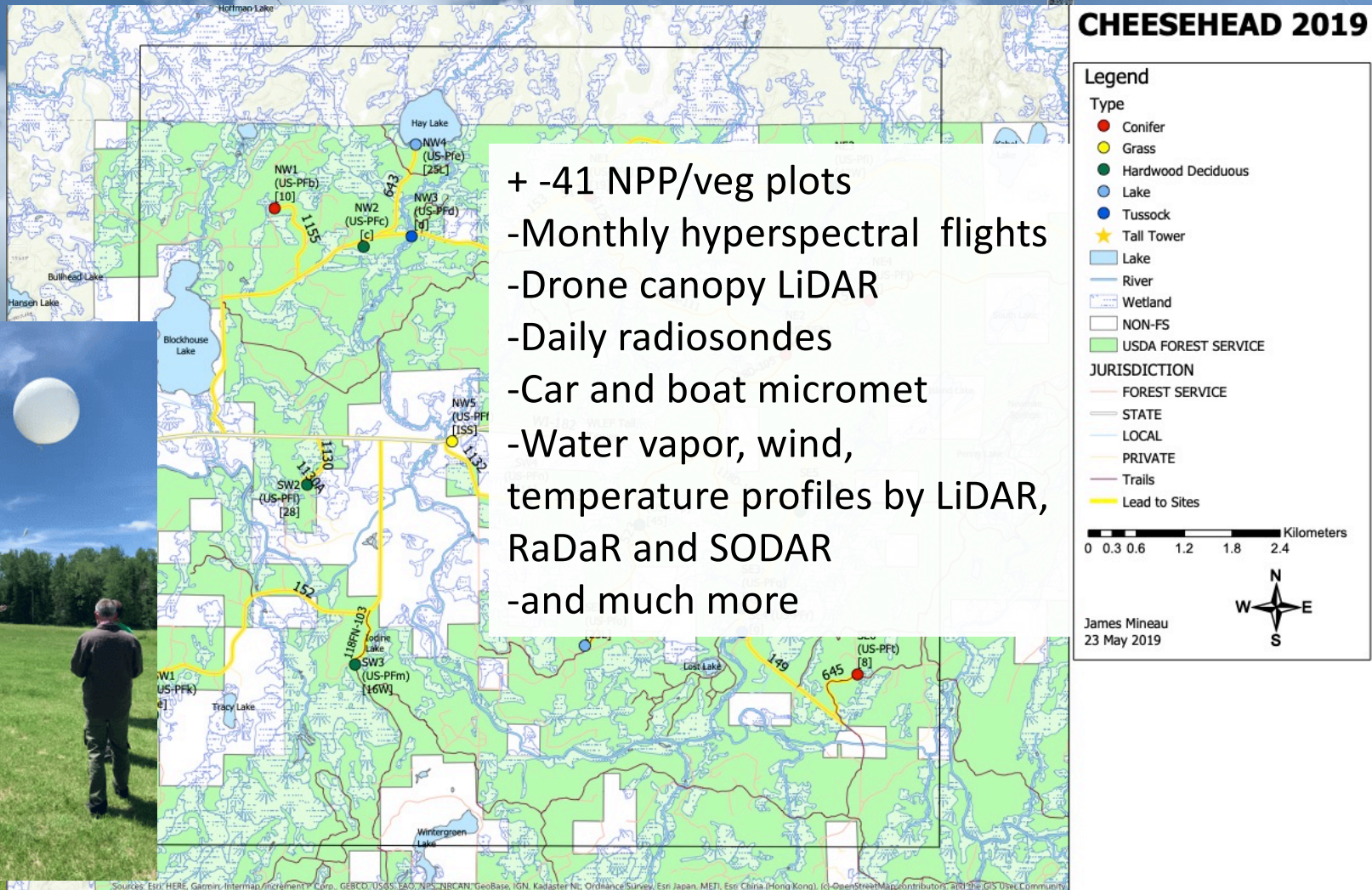
# Enabling scalable and reproducible analysis



Metzer et al., 2019, BAMS



# “Small” field projects...



[https://www.eol.ucar.edu/field\\_projects/cheesehead](https://www.eol.ucar.edu/field_projects/cheesehead)



# Your eyes should now glaze over..

Data source	Data provider	POC	Location(s)	Measurement domain	Funded by NSF?	Collection period	Archive location
Ameriflux/NOAA very tall tower (US-PFA/WLEF)	University of Wisconsin-Madison, Atmospheric and Oceanic Sciences	Ankur Desai <desai@aos.wisc.edu>	WLEF	Surface	Y	Continuous	<a href="http://ameriflux.lbl.gov">http://ameriflux.lbl.gov</a> and <a href="https://flux.aos.wisc.edu">https://flux.aos.wisc.edu</a>
NOAA tall tower greenhouse gases	NOAA ESRL	Arlyn Andrews <Arlyn.Andrews@noaa.gov>	WLEF	Surface		Continuous	<a href="https://www.esrl.noaa.gov/gmd/ccgg/obspack/">https://www.esrl.noaa.gov/gmd/ccgg/obspack/</a>
NOAA Airborne greenhouse gas profiles	NOAA ESRL	Colm Sweeney <colm.sweeney@noaa.gov>	WLEF	Airborne Profiling		Biweekly	<a href="https://www.esrl.noaa.gov/gmd/ccgg/aircraft/data/">https://www.esrl.noaa.gov/gmd/ccgg/aircraft/data/</a>
ChEAS Ameriflux tower network (US-WCr/US-Los/US-Sy)	University of Wisconsin-Madison, Atmospheric and Oceanic Sciences	Ankur Desai <desai@aos.wisc.edu>	Ameriflux sites (4)	Surface		Continuous	<a href="http://ameriflux.lbl.gov">http://ameriflux.lbl.gov</a> and <a href="https://flux.aos.wisc.edu">https://flux.aos.wisc.edu</a>
Total Column Carbon Observing Network (TCCON)	CalTech	Coleen Roehl <coleen@gps.caltech.edu>	WLEF	In-Situ Profiling		Continuous	<a href="https://tcconda.org/">https://tcconda.org/</a>
Ground-based vegetation/phenology sampling	University of Wisconsin-Milwaukee, Geography	Mark Schwartz <mds@uwm.edu>	10x10 domain, 27 plots	Surface	Y	Sept-Oct 2019	NCAR EOL
Vegetation inventory, root growth, NPP, Biometry, leaf spec	University of Wisconsin-Madison, Forest and Wildlife Ecology	Eric Kruger <elkruger@wisc.edu>	10x10 domain, 41 plots	Surface		June-Oct, 2019	NCAR EOL
ISFS Eddy covariance towers: CO2/H2O/H flux, netrad, n	NCAR EOL Integrated Surface Flux System (ISFS)	Steve Oncley <oncleyst@ucar.edu>, Gary C	10x10 domain, 17 locations	Surface	Y	June-Oct, 2019	<a href="https://www.eol.ucar.edu/field_projects/cheesehead/">https://www.eol.ucar.edu/field_projects/cheesehead/</a>
MSU Eddy covariance towers	University of Wisconsin-Madison, Biological Systems Engineering	Paul Stoy <pctstoy@wisc.edu>	NW5 (ISS) and SE1	Surface		June-Oct, 2019	NCAR EOL
449 MHz modular wind profiler	NCAR EOL Integrated Sounding System (ISS)	Bill Brown <wbrown@ucar.edu>	ISS/Flambeau field	In-Situ Profiling	Y	July-Oct, 2019	<a href="https://www.eol.ucar.edu/field_projects/cheesehead/">https://www.eol.ucar.edu/field_projects/cheesehead/</a>
RASS/SODAR	NCAR EOL Integrated Sounding System (ISS)	Bill Brown <wbrown@ucar.edu>	ISS/Flambeau field	In-Situ Profiling	Y	July-Oct, 2019	<a href="https://www.eol.ucar.edu/field_projects/cheesehead/">https://www.eol.ucar.edu/field_projects/cheesehead/</a>
Ceillometer	NCAR EOL Integrated Sounding System (ISS)	Bill Brown <wbrown@ucar.edu>	ISS/Flambeau field	In-Situ Profiling	Y	July-Oct, 2019	<a href="https://www.eol.ucar.edu/field_projects/cheesehead/">https://www.eol.ucar.edu/field_projects/cheesehead/</a>
All sky camera and web camera	NCAR EOL Integrated Sounding System (ISS)	Bill Brown <wbrown@ucar.edu>	ISS/Flambeau field	Surface	Y	July-Oct, 2019	<a href="https://www.eol.ucar.edu/field_projects/cheesehead/">https://www.eol.ucar.edu/field_projects/cheesehead/</a>
Surface tower met (T,RH,P,winds,precip,winds)	NCAR EOL Integrated Sounding System (ISS)	Bill Brown <wbrown@ucar.edu>	ISS/Flambeau field	Surface	Y	July-Oct, 2019	<a href="https://www.eol.ucar.edu/field_projects/cheesehead/">https://www.eol.ucar.edu/field_projects/cheesehead/</a>
Daily radiosonde	NCAR EOL Integrated Sounding System (ISS)	Bill Brown <wbrown@ucar.edu>	ISS/Flambeau field	In-Situ Profiling	Y	July-Oct, 2019	<a href="https://www.eol.ucar.edu/field_projects/cheesehead/">https://www.eol.ucar.edu/field_projects/cheesehead/</a>
High-resolution PBL sondes during IOPs	University of Wisconsin-Madison, Atmospheric and Oceanic Sciences	Ankur Desai <desai@aos.wisc.edu>	10x10 domain, ISS field	In-Situ Profiling	Y	IOPs	<a href="https://www.eol.ucar.edu/field_projects/cheesehead/">https://www.eol.ucar.edu/field_projects/cheesehead/</a>
3-hrly daytime sondes (4-5 per day, 5 days per IOP)	NCAR EOL Integrated Sounding System (ISS)	Bill Brown <wbrown@ucar.edu>	ISS/Flambeau field	In-Situ Profiling	Y	IOPs	<a href="https://www.eol.ucar.edu/field_projects/cheesehead/">https://www.eol.ucar.edu/field_projects/cheesehead/</a>
Atmospheric Emitted Radiance Interferometer (AERI)	UW Space Science and Engineering Center Portable Atmospheric Research Center (SPARC)	Timothy J Wagner <tim.wagner@ssec.wisc.edu>	WLEF	In-Situ Profiling	Y	July-Oct, 2019	<a href="http://webmads.ssec.wisc.edu/#SPARC/dash">http://webmads.ssec.wisc.edu/#SPARC/dash</a>
HALO Photonics Streamline scanning wind LIDAR (1)	UW Space Science and Engineering Center Portable Atmospheric Research Center (SPARC)	Timothy J Wagner <tim.wagner@ssec.wisc.edu>	WLEF	In-Situ Profiling	Y	July-Oct, 2019	
High-Spectral Resolution Lidar (HSRL)	UW Space Science and Engineering Center Portable Atmospheric Research Center (SPARC)	Ed Eloranta <eloranta@wisc.edu>, Timoth	WLEF	In-Situ Profiling	Y	July-Oct, 2019	<a href="http://hsrl.ssec.wisc.edu/">http://hsrl.ssec.wisc.edu/</a>
Microrain radar (MRR)	UW SSEC	Claire Pettersen <claire.pettersen@ssec.wisc.edu>	WLEF	In-Situ Profiling		July-Oct, 2019	
Precipitation Imaging Probe (PIP)	UW SSEC	Claire Pettersen <claire.pettersen@ssec.wisc.edu>	WLEF	In-Situ Profiling		July-Oct, 2019	
COSMOS Soil Moisture	University of Arizona	Marek Zreda <marek@hwr.arizona.edu>	WLEF	Surface		July-Oct, 2019	<a href="http://cosmos.hwr.arizona.edu/Probes/StationData/042">http://cosmos.hwr.arizona.edu/Probes/StationData/042</a>
DIAL/Raman Lidar	Karlsruhe Institute for Technology, IMK IFU	Hannes Vogelmann <vogelmann@kit.edu>	WLEF	In-Situ Profiling		July-Oct, 2019	
HALO Photonics Streamline scanning wind LIDAR (2,3)	Karlsruhe Institute for Technology, IMK IFU	Hannes Vogelmann <vogelmann@kit.edu>	WLEF	In-Situ Profiling		July-Oct, 2019	
CAMSIS GOES-simulator vegetation camera	NASA Goddard	Joel McCorkel <joel.mccorkel@nasa.gov>	WLEF	Surface		July-Oct, 2019	
Airborne eddy-covariance (25 Hz)	University of Wyoming King Air	David Plummer <dplumme1@uwyo.edu>	30x30 domain, 24 flights	Airborne	Y	IOPs	<a href="http://flights.uwyo.edu/projects/cheesehead19/">http://flights.uwyo.edu/projects/cheesehead19/</a>
Airborne meteorology and radiation (1 Hz)	University of Wyoming King Air	David Plummer <dplumme1@uwyo.edu>	30x30 domain, 24 flights	Airborne	Y	IOPs	<a href="http://flights.uwyo.edu/projects/cheesehead19/">http://flights.uwyo.edu/projects/cheesehead19/</a>
Compact Raman LIDAR (CRL)	University of Wyoming King Air	Zhien Wang <Zhien.Wang@colorado.edu>	30x30 domain, 24 flights	Airborne	Y	IOPs	<a href="http://flights.uwyo.edu/projects/cheesehead19/">http://flights.uwyo.edu/projects/cheesehead19/</a>
Wyoming cloud LIDAR (WCL)	University of Wyoming King Air	David Plummer <dplumme1@uwyo.edu>	30x30 domain, 24 flights	Airborne	Y	IOPs	<a href="http://flights.uwyo.edu/projects/cheesehead19/">http://flights.uwyo.edu/projects/cheesehead19/</a>
Hyperspectral high-resolution maps (Hyspex)	University of Wisconsin Hyperspectral Imager	Erin Wagner <ephokanson@wisc.edu>, Pi	10x10 domain, 4 flights?	Airborne	Y	IOPs	<a href="http://flights.uwyo.edu/projects/cheesehead19/">http://flights.uwyo.edu/projects/cheesehead19/</a>
Boundary-layer heat and water budget of domain	University of Wisconsin-Ultralight	Grant Petty <gwpetty@wisc.edu>	10x10 domain	Airborne	Y	IOPs	<a href="http://flights.uwyo.edu/projects/cheesehead19/">http://flights.uwyo.edu/projects/cheesehead19/</a>
UAVs (fixed-wing and quadcopter) - payloads TBD	University of Wisconsin-Madison, Atmospheric and Oceanic Sciences	Grant Petty <gwpetty@wisc.edu>	ISS	Airborne	Y	IOPs	
NOAA SURFRAD Station and TWST (downwelling and upwelling)	NOAA ESRL GMD	Kathy Lantz <kathy.o.lantz@noaa.gov>	ISS field	Surface		July-Oct, 2019	<a href="https://esrl.noaa.gov/gmd/grad/tools/cheesehead/">https://esrl.noaa.gov/gmd/grad/tools/cheesehead/</a>
NOAA RadSys Stations, Ceillometer (total/direct/diffuse S)	NOAA ESRL GMD	Kathy Lantz <kathy.o.lantz@noaa.gov>	Prentice Airport, Arbor Vitae	Surface		July-Oct, 2019	<a href="https://esrl.noaa.gov/gmd/grad/tools/cheesehead/">https://esrl.noaa.gov/gmd/grad/tools/cheesehead/</a>
CLAMPS (Microwave radiometer, AERI, Doppler wind lidar)	NOAA NSCL	Dave Turner <dave.turner@noaa.gov>	Prentice Airport, Arbor Vitae	Surface		Sept-Oct 2019	<a href="https://www.nssl.noaa.gov/users/turner/public_html/">https://www.nssl.noaa.gov/users/turner/public_html/</a>
Mobile vehicle transects	Jackson State University	Loren White <Loren.D.White@sums.edu>	10x10 domain, trails	Surface	Y	IOPs	
Pedestrian/boat mobile transects	Jackson State University	Loren White <Loren.D.White@sums.edu>	10x10 domain, trails and Ha	Surface	Y	IOPs	
NOAA ATDD Drones (2, Metex and IR temp)	NOAA ATDD	Temple Lee <temple.lee@noaa.gov>, Edw	WLEF and SW2	Airborne Profiling		IOPs	
UWEC Ozone Profile	UW-Eau Claire	Patricia Cleary <clearypa@uwec.edu>	WLEF	Airborne Profiling		IOP1	
Ozone flux profile	UW-Madison Chemistry	Michael Vermeuel <mvermeuel@wisc.edu>	WLEF	Surface		IOP1	
Soil and water chemistry	Butternut Schools	Laurie Fox <lafdx85@gmail.com>, Rosa	5 sites	Surface	Y	July 2019	
Tree temperature	Class ACT Charter, Chequamegon High School	Travis Augustine <taugustine@csdk12.net>	5 sites, 10 trees	Surface	Y	Oct 2019	
Soil carbon, nitrogen	University of Wisconsin-Madison AOS	Bailey Murphy <bamurphy5@wisc.edu>	16 tower sites	Surface		Oct 2019	
Soil bulk density and heat capacity	NCAR	Steve Oncley <oncleyst@ucar.edu>	17 tower sites	Surface	Y	July-Oct, 2019	
Drone canopy height LIDAR leaf-on	University of Wisconsin-Geography	Christian Andresen <candresen@wisc.edu>	8 1x1 km plots	Surface	Y	June 2019	
QL2 leaf-off LIDAR	United States Forest Service	Mark Farina <mark.farina@usda.gov>	30x30 domain	Airborne		Fall 2018	<a href="https://www.sco.wisc.edu/data/elevationlidar/">https://www.sco.wisc.edu/data/elevationlidar/</a>
NASA ECOSTRESS, GEDI, OCO3 overpasses	NASA JPL, target mode planned for October 8	Annmarie Eldering <annmarie.eldering@jpl.nasa.gov>	30x30 domain	Satellite		July-Oct, 2019	
Earth observing landcover datasets, vegetation maps	Globalview, WISCLAND, Landsat, MODIS, Suomi, Sentinel, GOES, and other land cover overpa	Varies	Varies	Satellite and Maps		IOPs	
G-LiHT thermal and LIDAR 2012	NASA Goddard	Bruce Cook <bruce.cook@nasa.gov>		Airborne		June 2012	<a href="https://gltihsdata.gsfc.nasa.gov/files/G-LiHT/">https://gltihsdata.gsfc.nasa.gov/files/G-LiHT/</a>
NWS model forecasts and soundings	NOAA NWS via NCAR (HRRR, NAM, GFS, Soundings, local weather stations)	Varies	Varies	Model + Other Obs	Y	N/A	
Large Eddy Simulation (LES)	University of Hanover	Matthias Sührling <suehring@muk.uni-hannover.de>		Model		N/A	
Large Eddy Simulation (LES)	University of Wisconsin	Sreenath Paleri <paleri@wisc.edu>		Model		N/A	
Large Eddy Simulation (LES)	KIT IMK IFU	Matthias Mauder <mthias.mauder@kit.edu>		Model		N/A	
Photos and videos	UW-Madison and partners, camera, drones	Brian Butterworth <butterworth@wisc.edu>	30x30 domain	Photos	Y	July-Oct, 2019	<a href="https://photos.google.com/share/AF1QipN58-BZ777B6">https://photos.google.com/share/AF1QipN58-BZ777B6</a>
Field notes, logs, documents	All	Brian Butterworth <butterworth@wisc.edu>		Documents	Y	July-Oct, 2019	<a href="https://drive.google.com/drive/u/0/folders/1yMuVDBW">https://drive.google.com/drive/u/0/folders/1yMuVDBW</a>
NEON TOS data (biomass, biodiversity, small mammals, ...)	National Ecological Observatory Network (NEON)	Stefan Metzger <smetzger@battelleecolog	UNDE, TREE, STEI	Surface	Y	Continuous	<a href="https://data-portal.neon.org/">Data Portal - NEON Data</a>

# Analysis and coding are now central to ecology

- When data volume is in terabytes, data types number in the dozens, and collaborator lists are in the tens to hundreds, traditional “cottage-industry” style analysis is impractical and inefficient
- Bringing code to the data is the new mantra
  - Example: Cyverse



# CyVerse

(NSF, U Arizona)

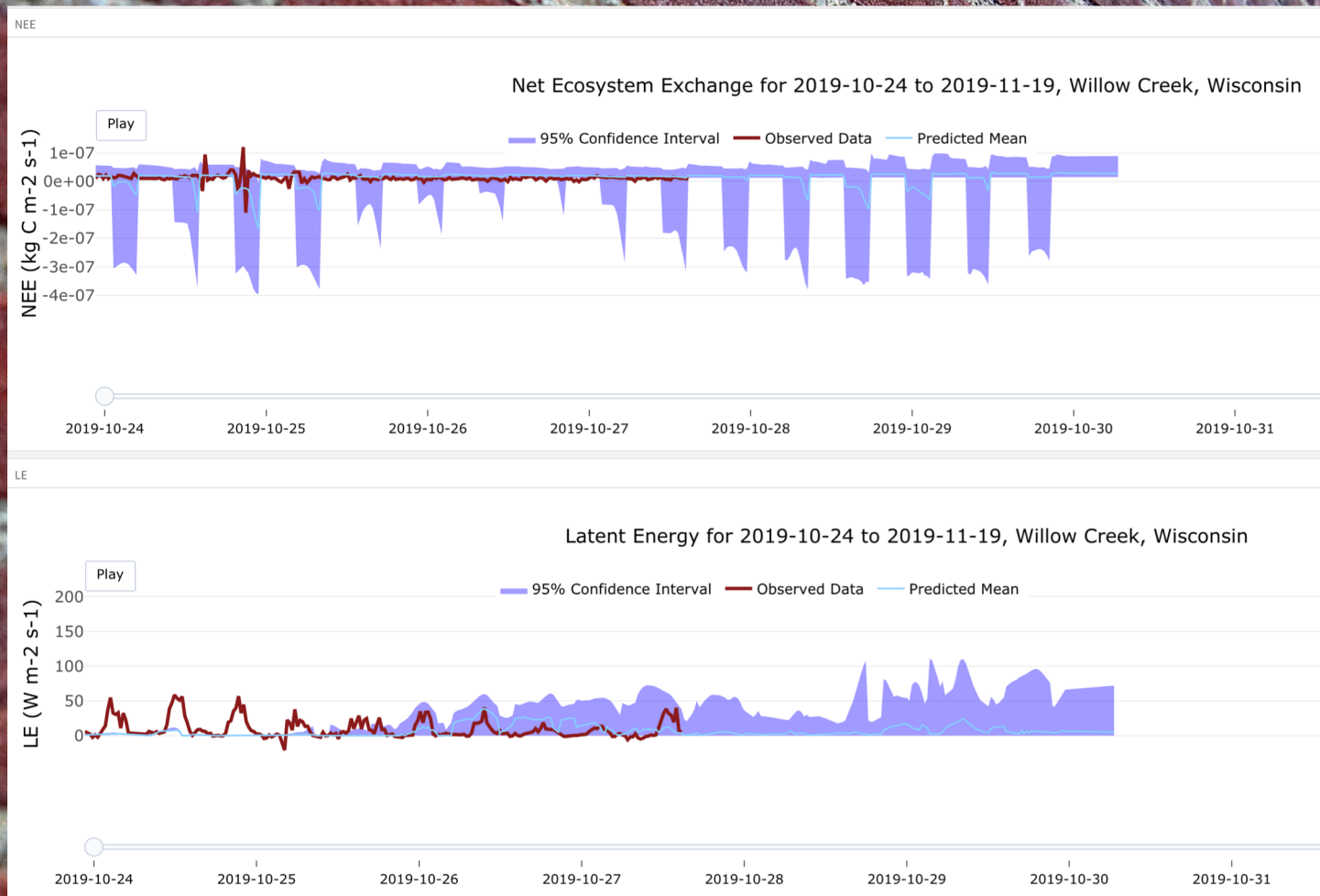
- <https://www.cyverse.org/atmosphere> <https://de.cyverse.org/de/>

The screenshot displays the CyVerse Discovery Environment interface. The top navigation bar includes the CyVerse logo, the text "CyVerse Discovery Environment", and user icons for notifications, profile, and help. The main interface is divided into several panels:

- Left Panel:** Contains icons for "Data" and "Apps". Below these is a "Data: inp" section with a navigation tree showing a hierarchy of folders including "ankurdesai", "Community Data", "Astrolabe", "BioViz", "Brassica\_dwarf", "CAP\_TCN", "CHEESEHEAD", "derv", "inp", "EuPathDB-Data", "G-OnRamp\_hubs", "GoreLab", "KBase\_staging", "Legume\_Federation", "NASA\_GeneLab\_data", "NEON\_Pilot", "NEON\_data\_institute\_20", "SEPA\_microbiome\_2016", and "TCIA".
- Apps Panel:** Features a "Categories" sidebar with "My Apps", "Topic", "Operation", and "HPC". The main area shows "Shared with me" with a list of apps, including "eddy4r-rstudio-deve" by "Tyson Swetnam".
- Data: inp Panel:** Shows a file browser view of the "/iplant/home/shared/CHEESEHEAD/inp" directory. It lists files and folders such as "ISFS", "metzger-et-al-2013", "name-of-observing-", "paleri-2019\_pre-iop", "suhring-et-al-2018", "uwka", "wiscland\_landcover", and "xu-et-al-2017\_tall-t".
- Analyses Panel:** Displays a table of analyses. The table has columns for Name, Owner, App, Start Date, End Date, and Status. One analysis is listed: "rstudio-3.5.0\_analysis1" by "ankurdesai" using "rstudio-3.5.0", which was "Canceled".

At the bottom right, there is a chat icon.

# Removing the wall between models and observations



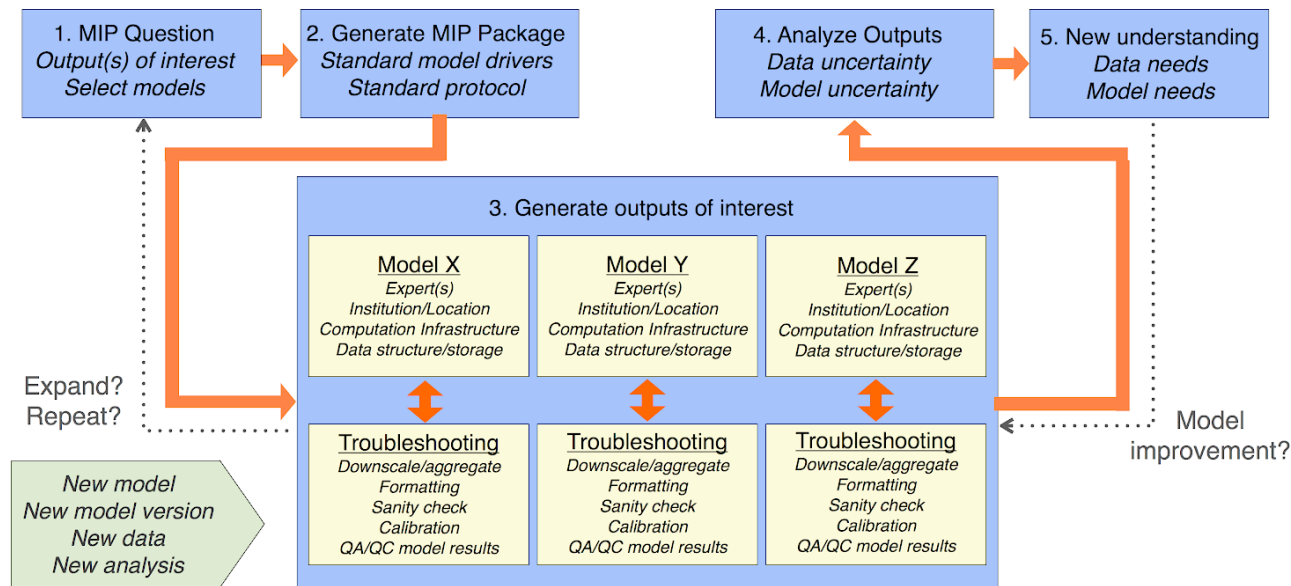
Katie Zarada, Mike Dietze, PEcAn project

# No one should be called a “modeler” or an “experimentalist”

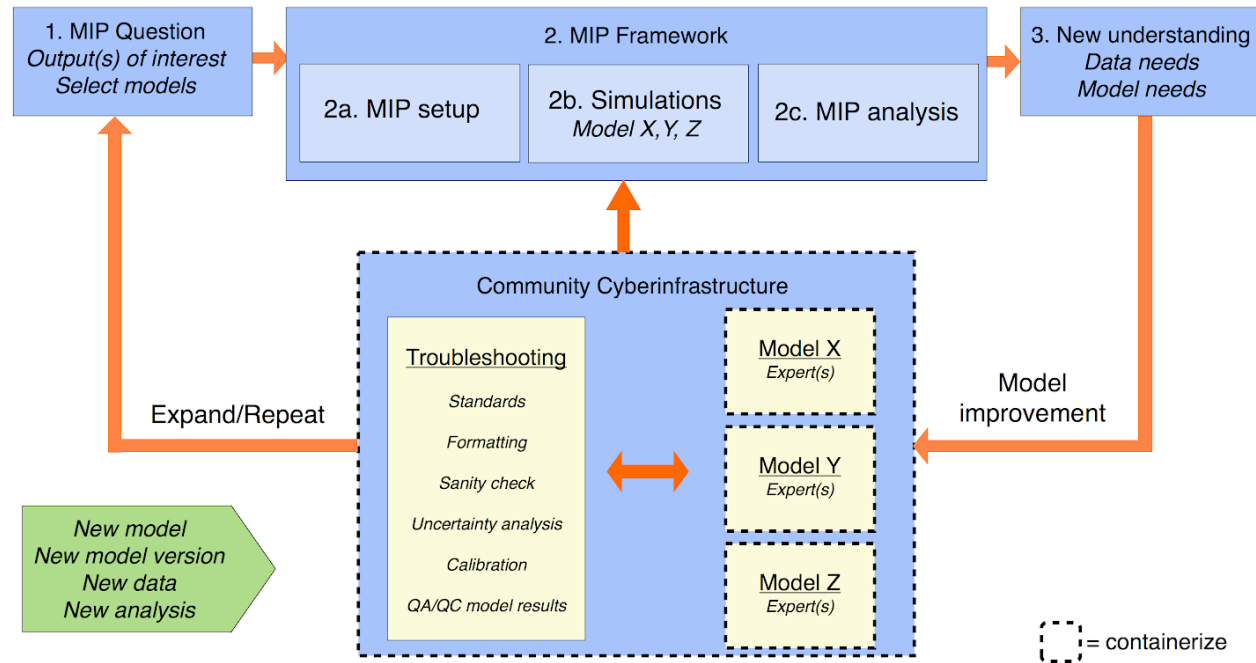
- Models are how we represent complex non-linear interactions among our theories
- No single type of observation is sufficient to evaluate the hypotheses embedded in models
- Model intercomparison and benchmarking projects can be made more routine



## Traditional Model Intercomparison Project (MIP) Framework

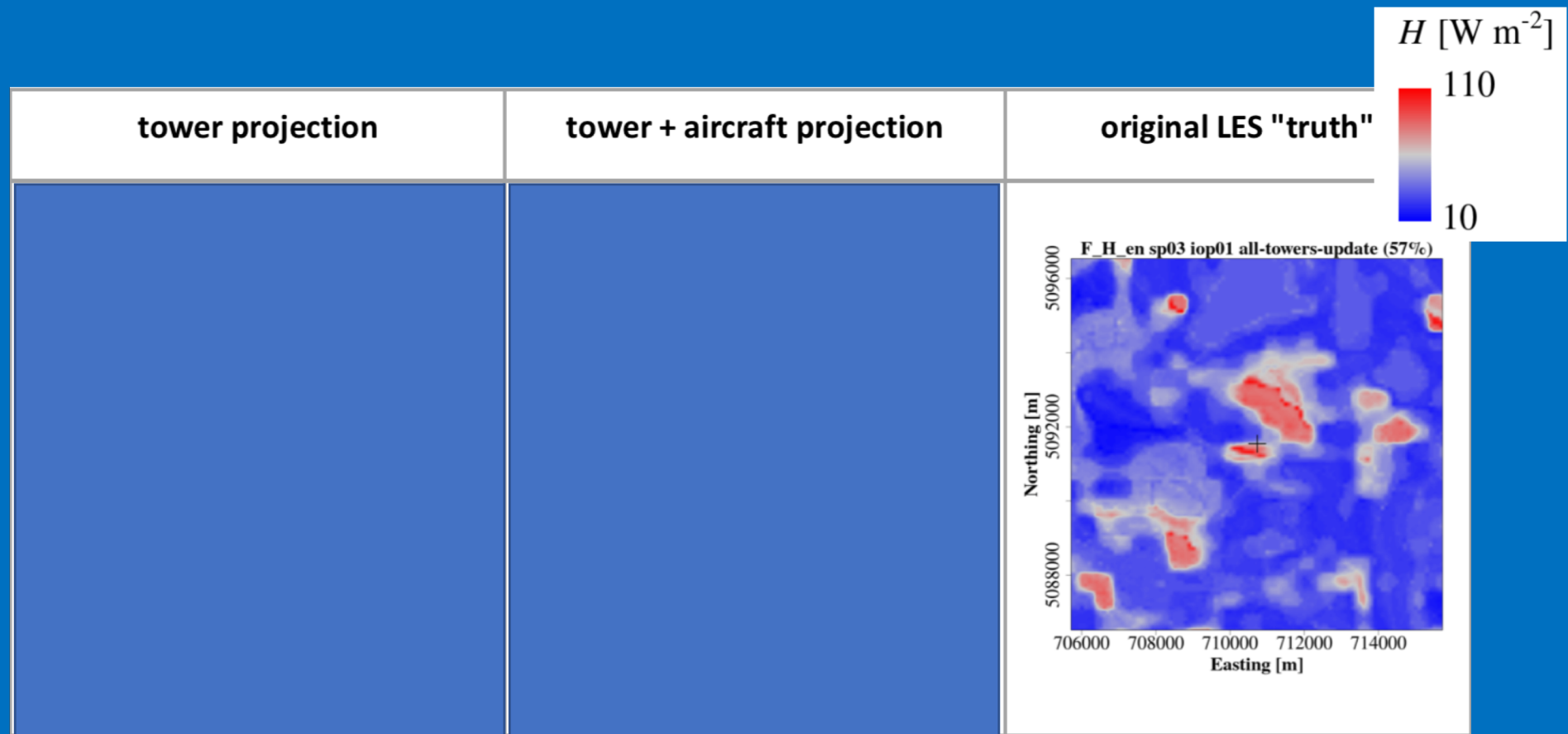


## MIP Framework with a Community Cyberinfrastructure



Learn more at B22D-01 Metzger TU 1020 MW3007 and B21C-06 Paleri TU 915-930 MW3007

# Observing System Simulation Experiments (OSSEs) flip the script on how models are used in experiments



Metzger *et al.*, 2013; Xu *et al.*, 2017, 2018;

S. Metzger, S. Paleri, K. Xu

# Huge Ecology is possible

- A diverse community makes global change ecology possible
- Diverse solutions are needed to confront the increasing volume in data and complexity in theory and models
- There are challenges that need to be solved mostly by students and early career scientists
  - But supported by community, funders, mentors
- The next few decades will be consequential for our discipline for a number of reasons



A wide-angle landscape photograph showing a vast forest of green trees under a blue sky with scattered white and grey clouds. A faint rainbow is visible in the sky, arching from the left side towards the horizon. The foreground shows the dense canopy of the forest, while the background extends to a distant horizon line.

# Thank you!

Ankur Desai

[desai@aos.wisc.edu](mailto:desai@aos.wisc.edu)

[@profdesai](https://flux.aos.wisc.edu)

Photo: A. Desai