



# SOils DAta Harmonization Tools (SoDaH)

Will Wieder, Kate Lajtha, Stevan Earl, Derek Pierson  
*& LTER Soil Synthesis Working Group*



# Contributions Welcome!



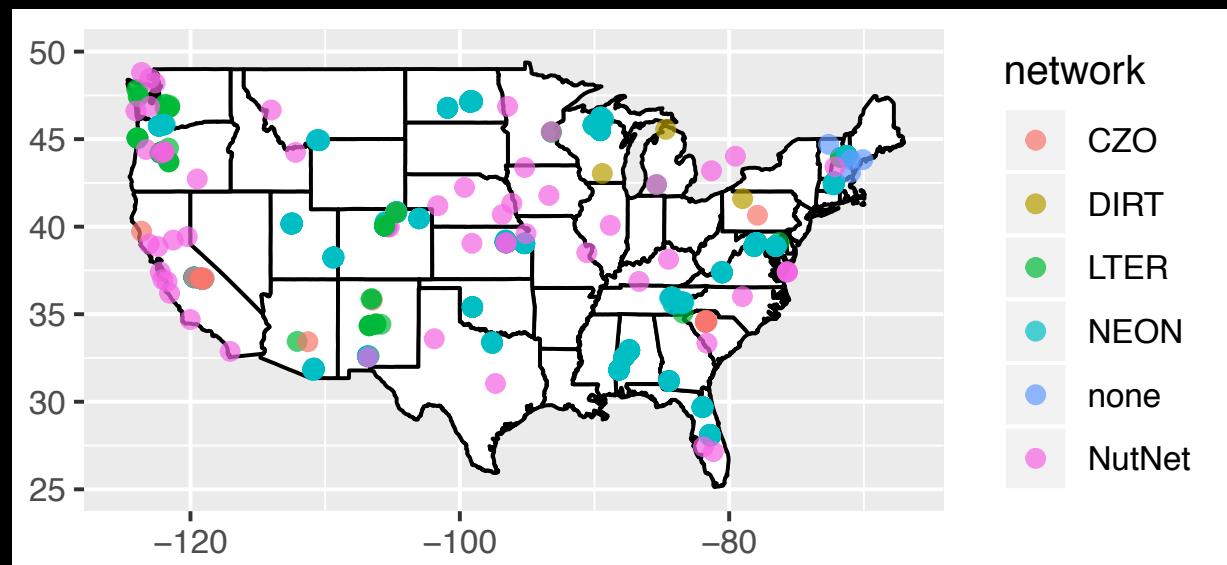
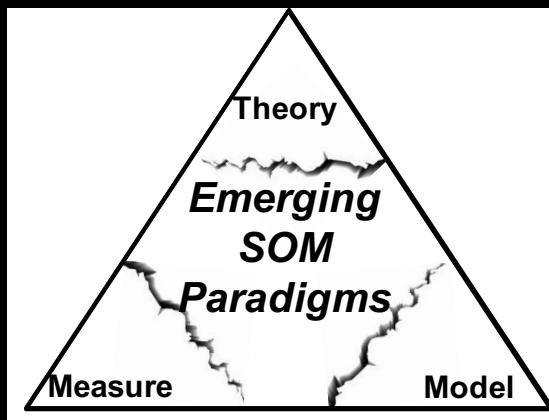
**Website**  
**Shiny app**  
**GitHub**

<https://lter.github.io/som-website/index.html>

<https://cosima.nceas.ucsb.edu/lter-som/>

<https://github.com/lter/lterwg-som>

# *Can research networks fill in data gaps?*





Broad data  
coverage +  
Standardization

Ecological  
dynamics +  
Disturbance

Deep soils +  
Deep time

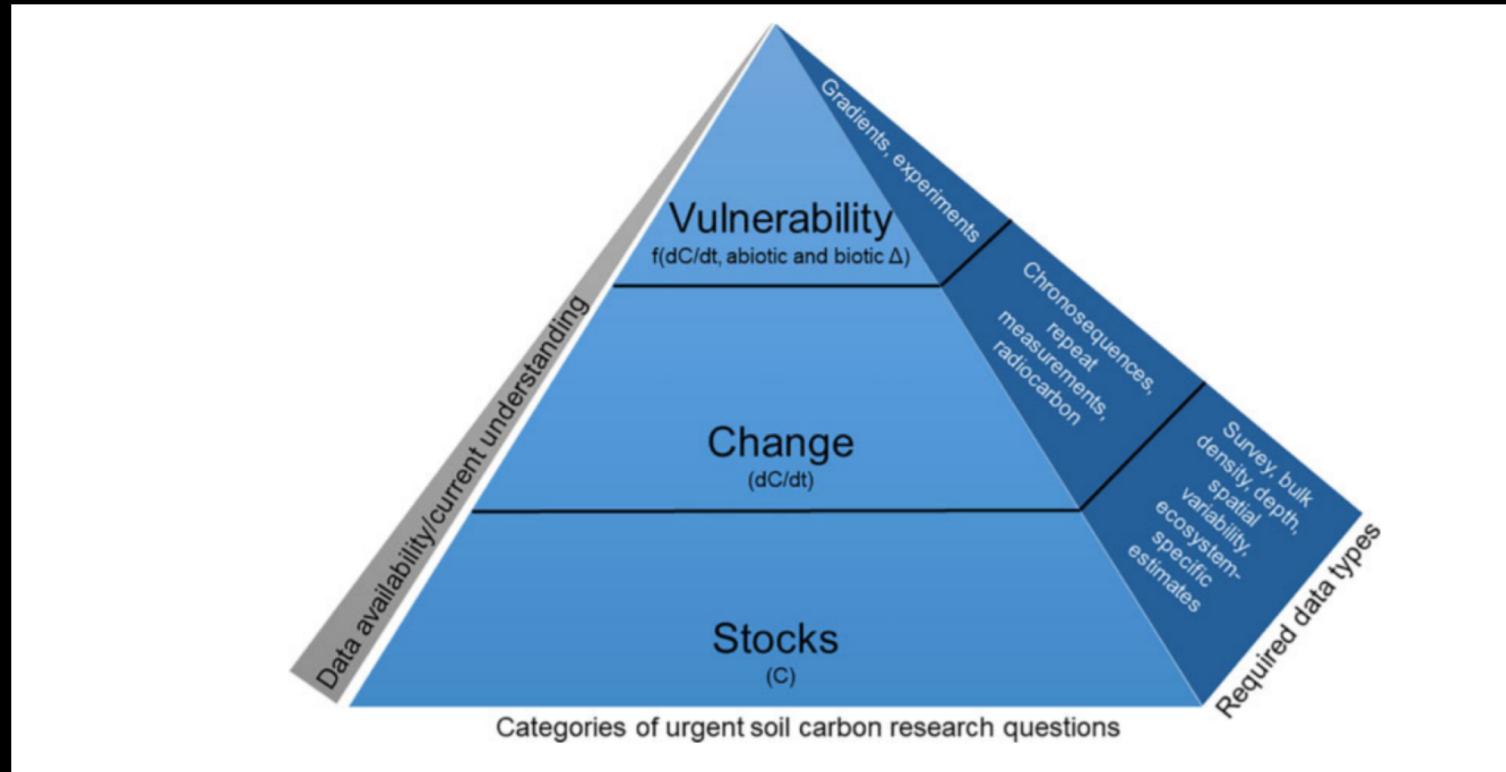
Modified from Weintraub et al *JGR-B* 2019



**DIRT**

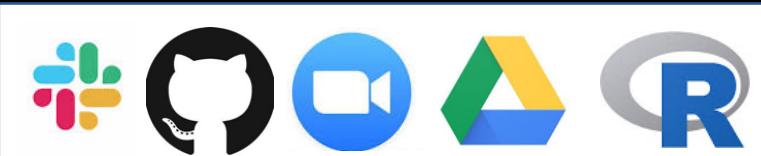


What is the landscape of soils data?  
What are the strengths of each database?



Malhotra, Avni, et al. (2019) The landscape of soil carbon data: emerging questions, synergies and databases. *Progress in Physical Geography: Earth and Environment* doi:10.1177/0309133319873309

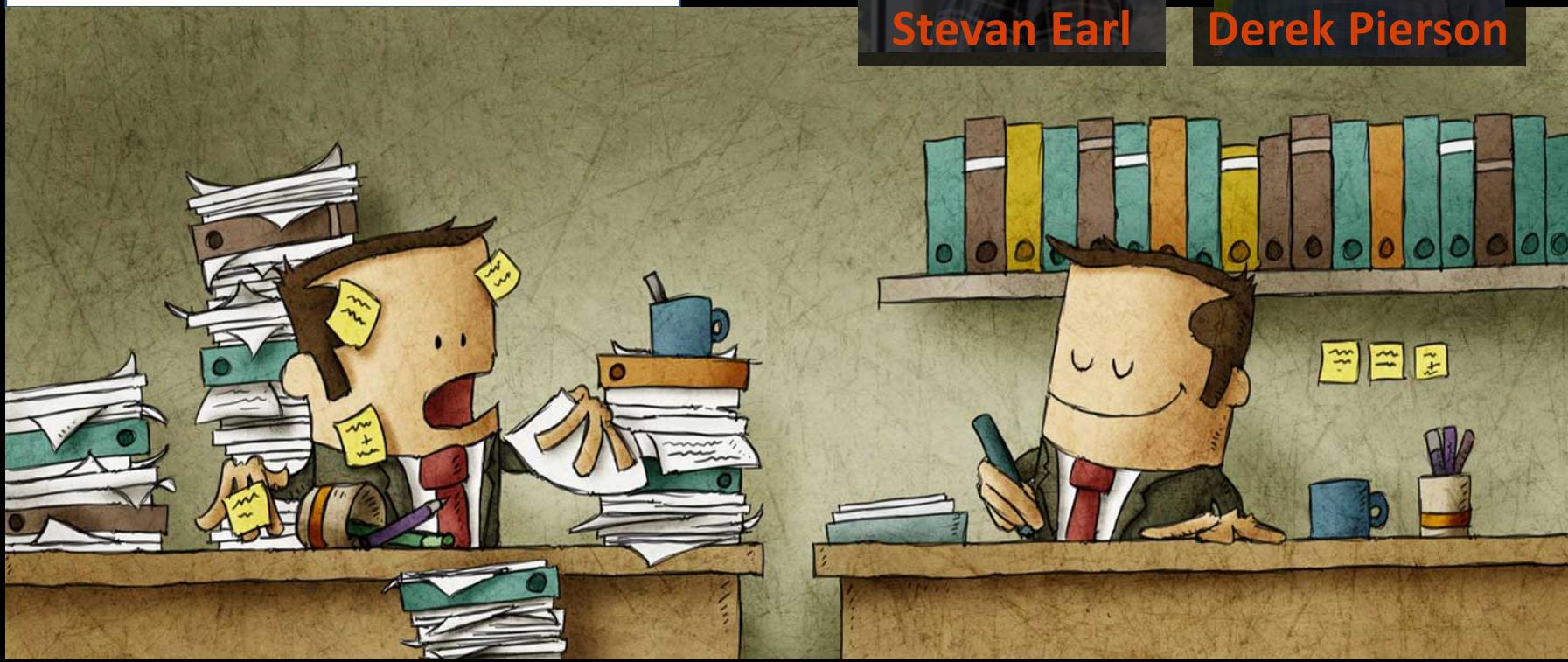
# *Data synthesis challenges*



Stevan Earl



Derek Pierson



# Data Synthesis Challenges

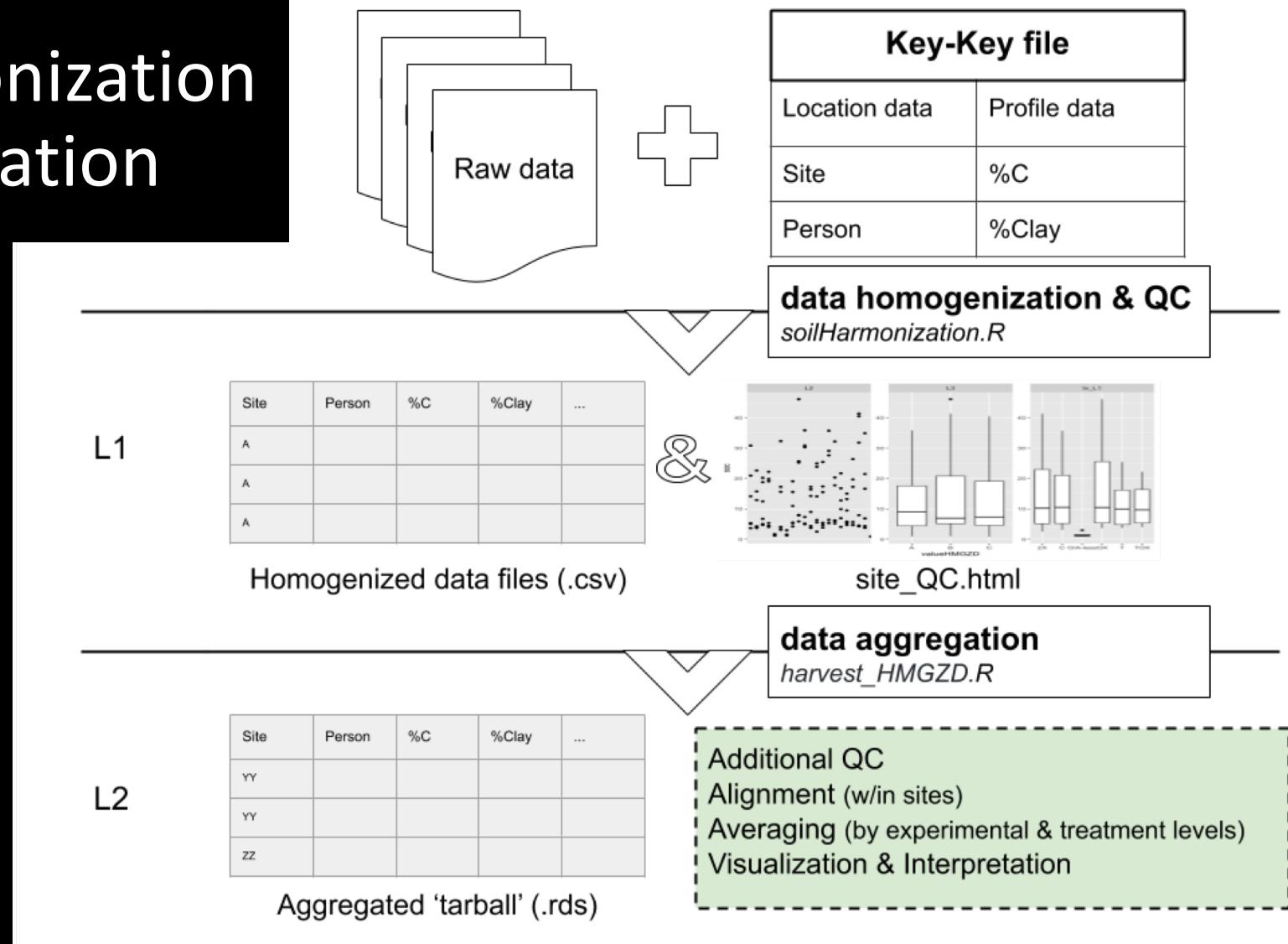
Identification

Harmonization & Aggregation

Visualization

Analyses

# Data Harmonization and Aggregation



# Data Visualization

~ 190 Sites

~ 400 Unique locations

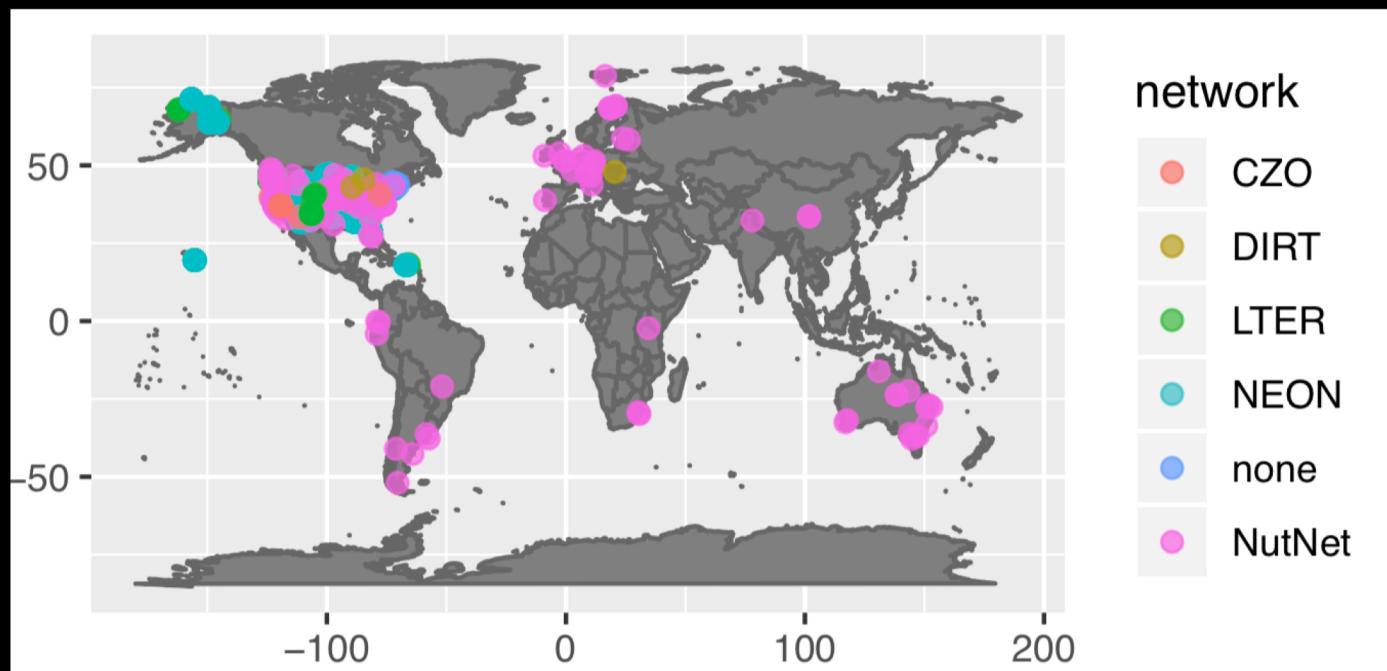
> 360,000 observations (rows)

## Shiny app

<https://cosima.nceas.ucsb.edu/lter-som/>

Shiny Tutorial:

<https://drive.google.com/file/d/1qAUFprGVZy4LGyaVtU2z6sS3Eo9C76H4/view?usp=sharing>

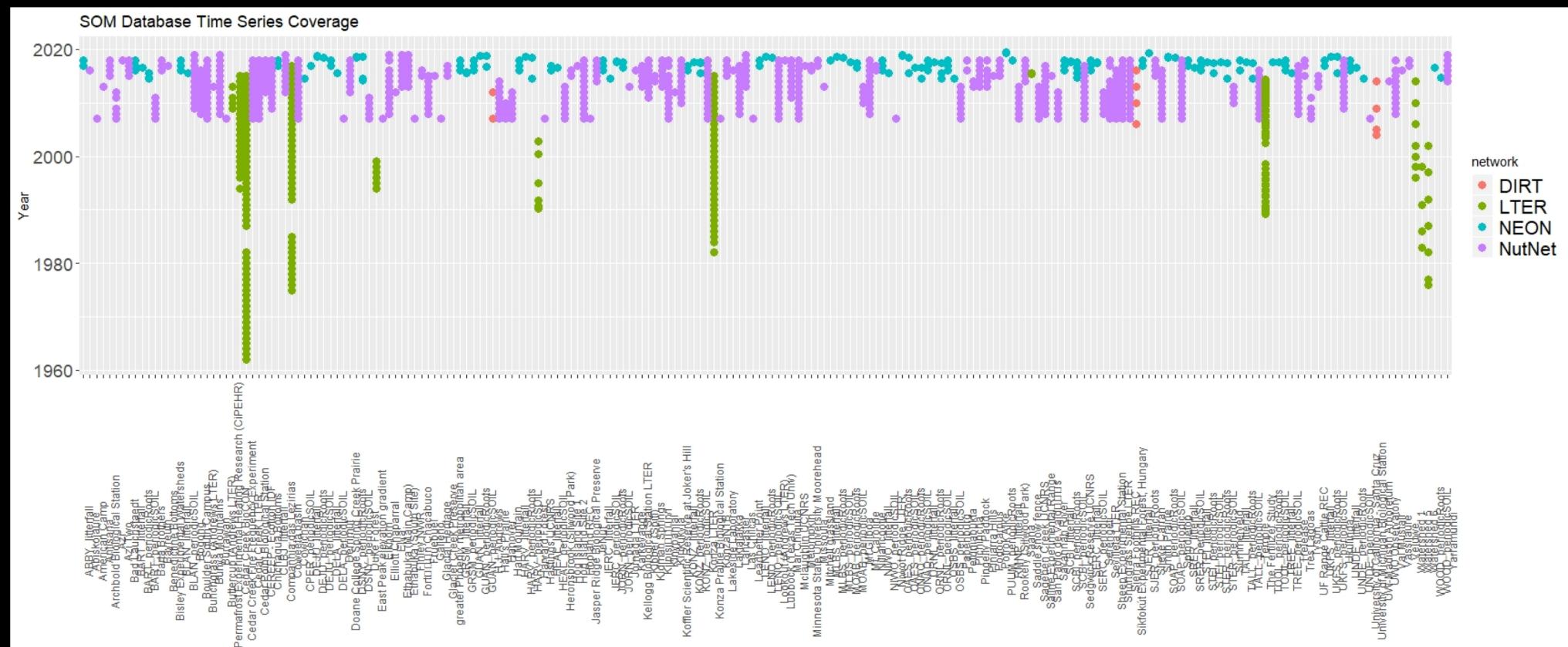




## Next steps:

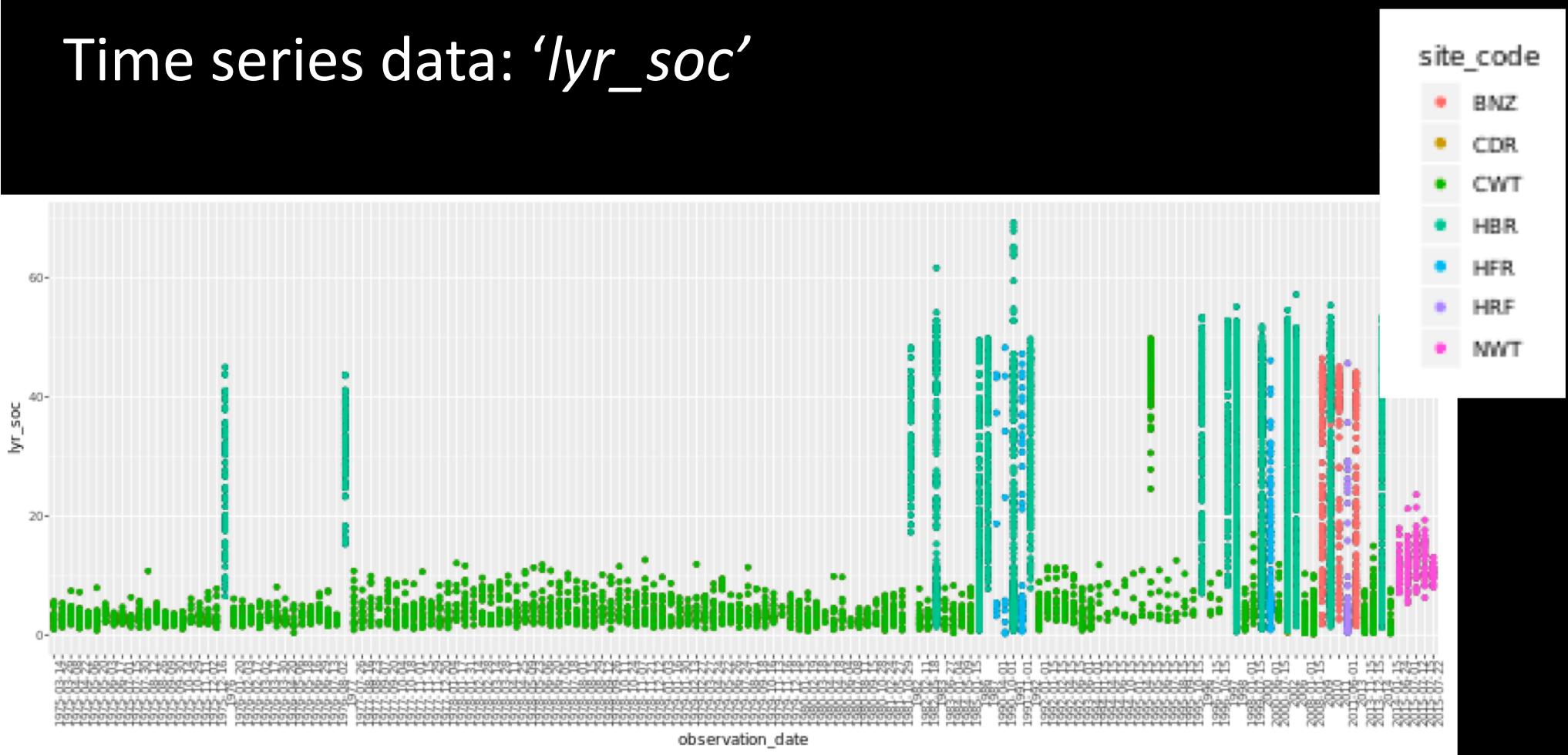
1. White paper (Kate & Sharon)
2. Finalize Database (Time Series?)
3. Publish raw data (doi where needed, Stevan & Will)
4. Publish database (Will)
5. Publish science papers! (various)

# Time series data:



## LTER: BNZ, CDR, CWT, HBR, HFR, KNZ, LUQ, NWT, SEV

## Time series data: '*lyr\_soc*'



LTER: BNZ, CDR, CWT, HBR, HFR, KNZ, LUQ, NWT, SEV

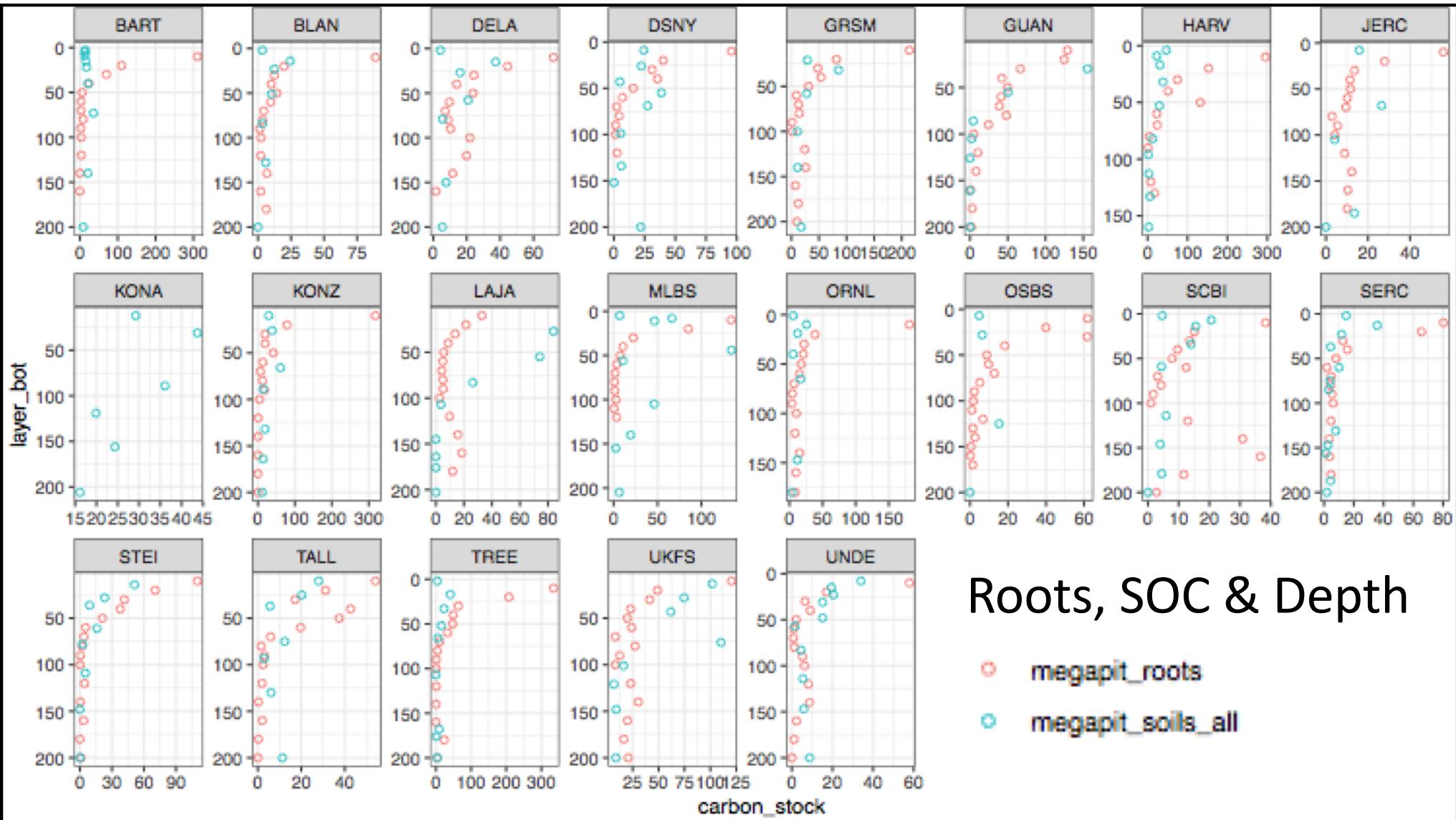
# Projects

<u>Link</u>	<u>Project</u>	<u>Lead</u>	<u>Status</u>
<a href="#">1</a>	White paper	K. Lajtha & S. Billings	Revising text
<a href="#">2</a>	Database paper	W. Wieder	Finalizing database, 2019 AGU poster
<a href="#">3</a>	Bulk Density & SOC	S. Grandy	S. Weintraub is exploring PTFs w/ NEON data
<a href="#">4</a>	Soil stoichiometric coupling & variability	E. Kyker-Snowman	In progress
<a href="#">5</a>	Big gradients and SOC controls	K. Georgiou	Machine learning with WoSIS database, 2019 AGU poster
<a href="#">6</a>	Linking root and soil carbon across sites	A. Malhotra	In progress
<a href="#">7</a>	Detecting a signal from the noise: Power analysis and SOC variability	D. Pierson	In progress, how do we capture heterogeneity with nested design?

## I want to get involved!

1. Contact project leads
2. Start your own project & get others to join in
3. Contribute analysis & ideas
4. Write!



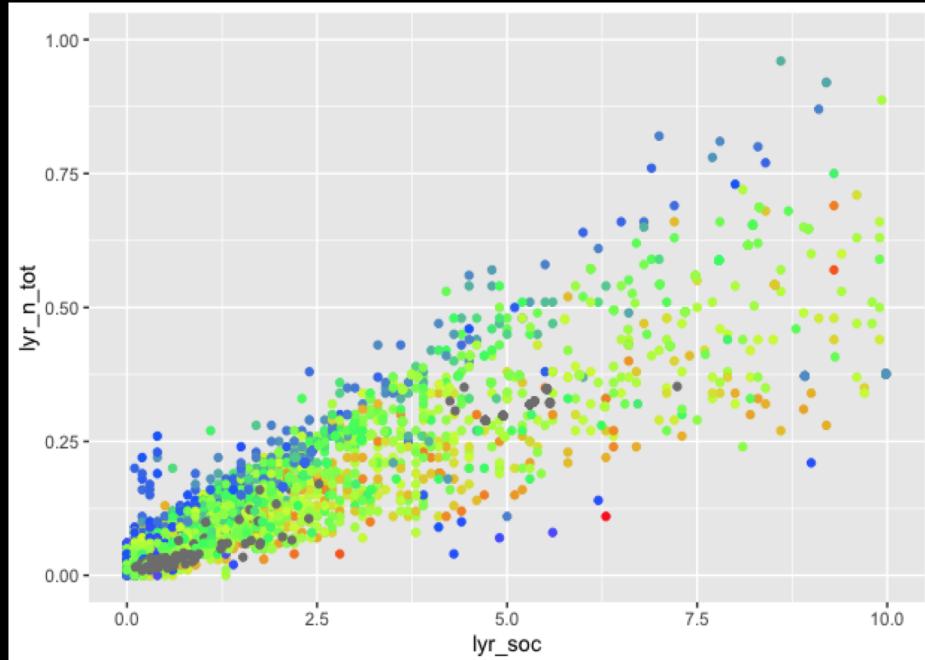


Roots, SOC & Depth

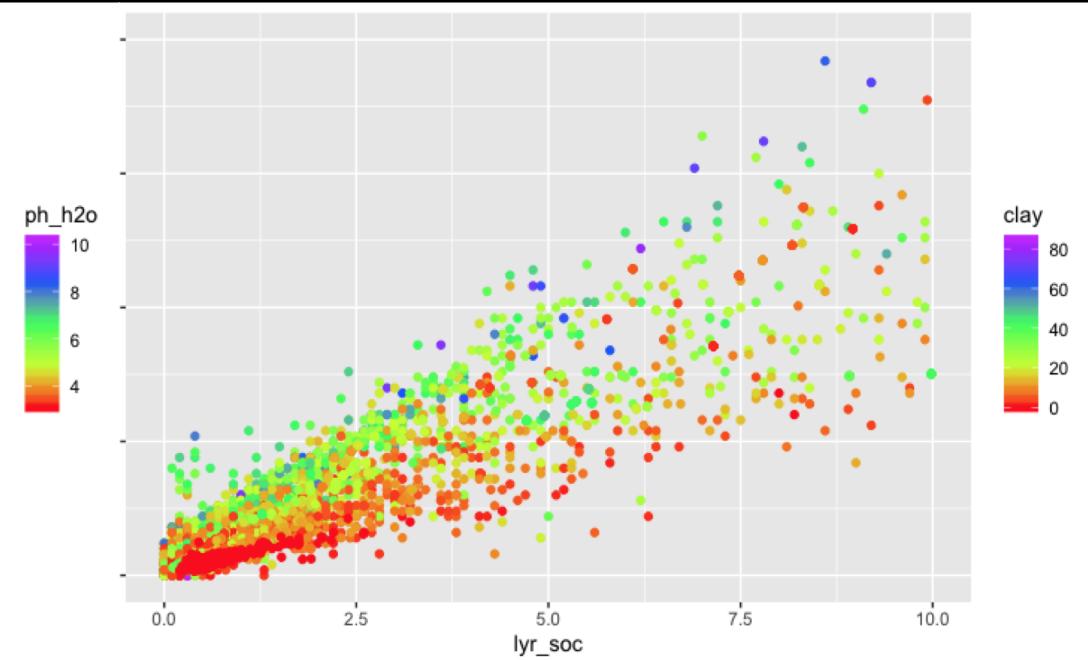
- megapit\_roots
- megapit\_soils\_all

# Soil controls over SOM stoichiometry

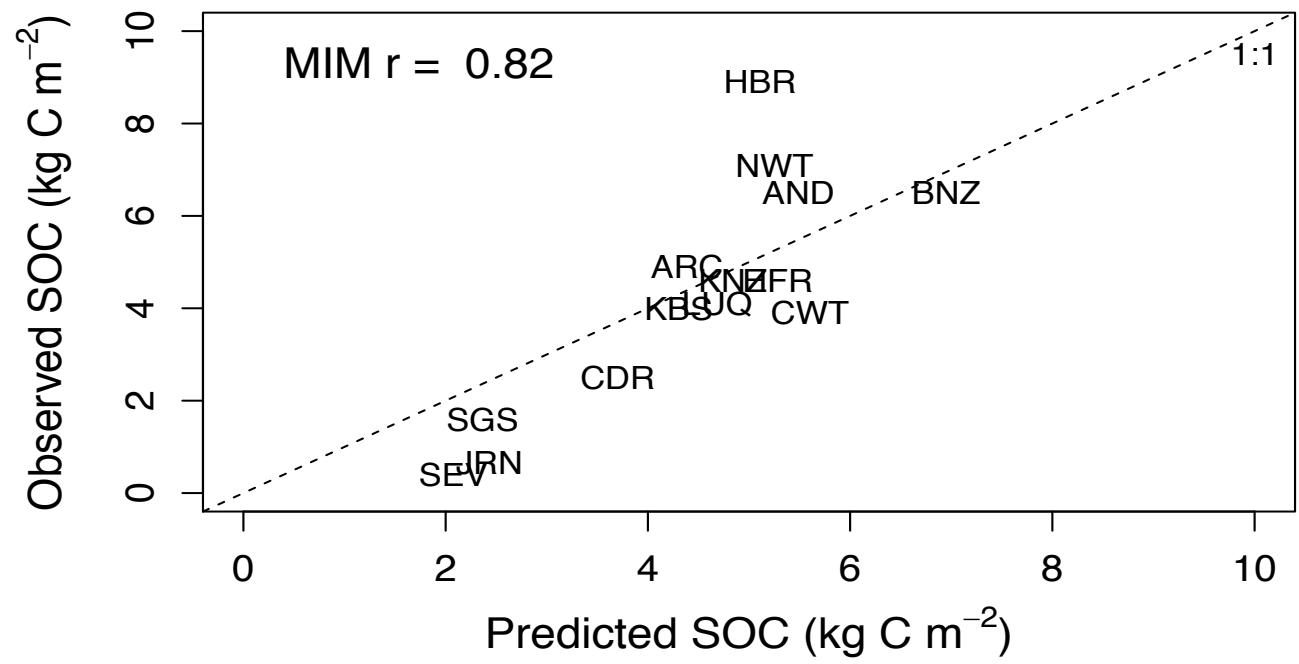
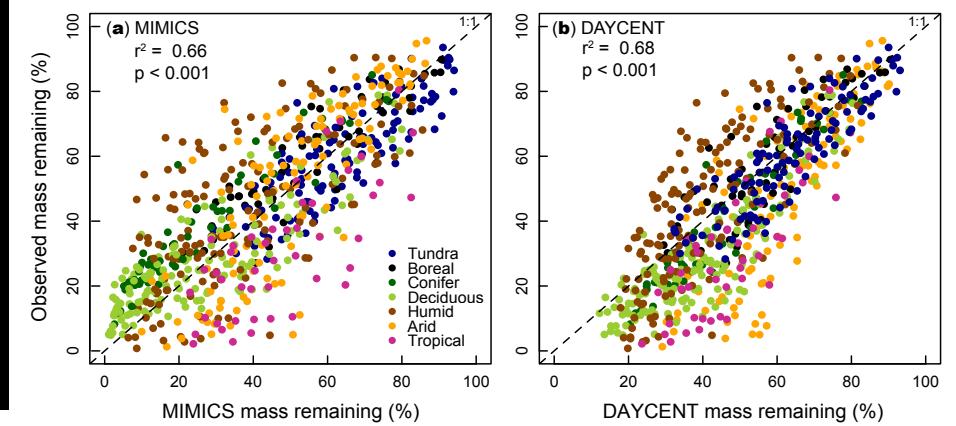
pH



Clay



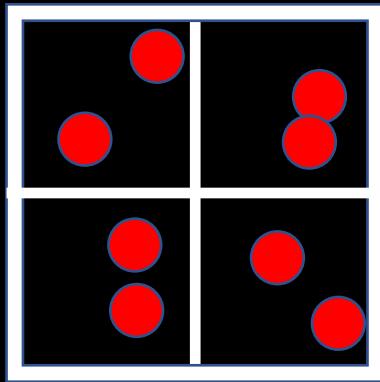
# Modeling + Within & among site variation



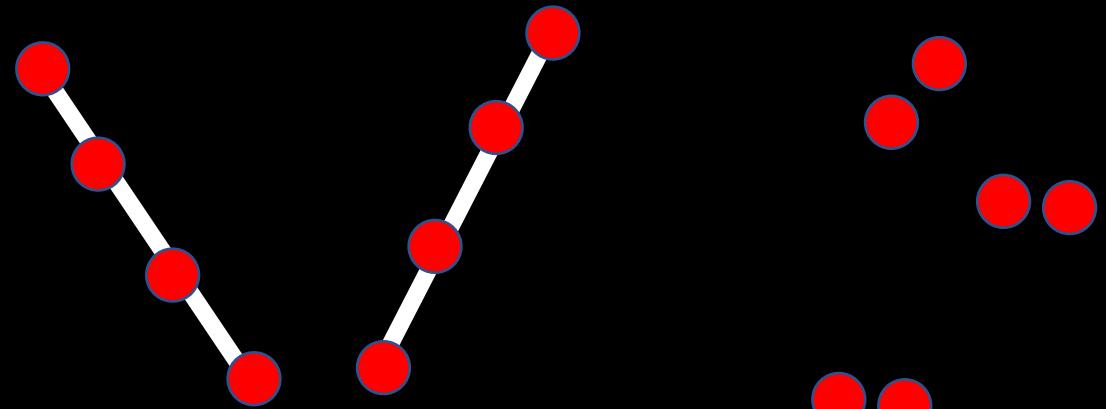
Wieder et al. GMD 2015

Wieder unpublished

# Remaining challenges



- + climate, vegetation, productivity, litter chemistry, roots
- + soil chemical & physical properties
- + depth
- + manipulations
- + time series



**Data Analysis Challenges**

- Alignment
- Averaging

# Contributions Welcome!



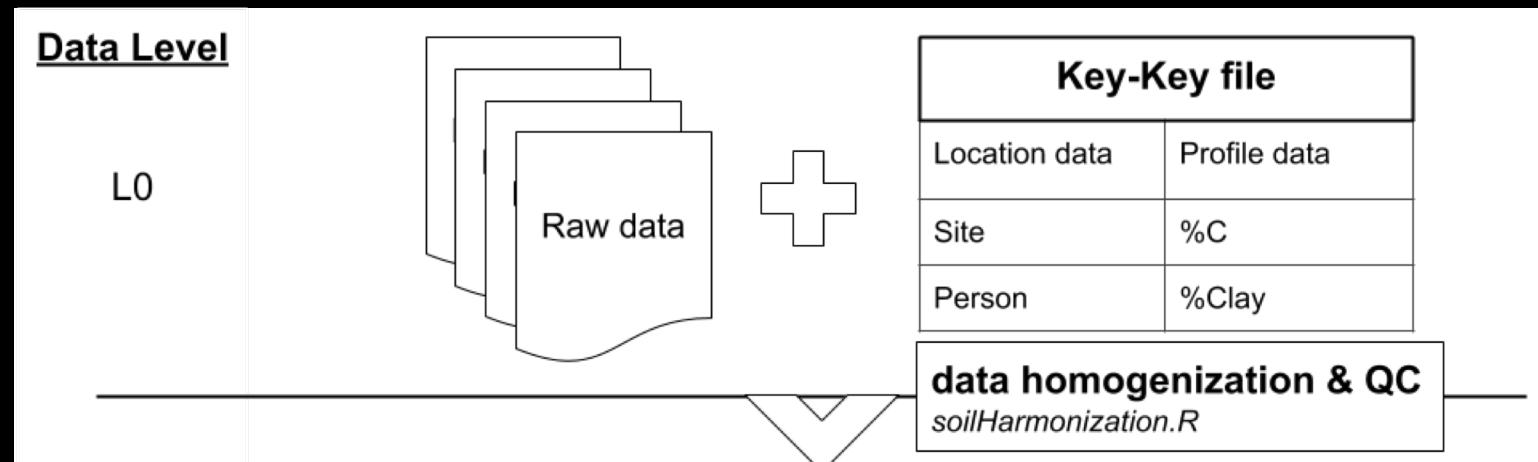
[Website](#)  
[Shiny app](#)  
[GitHub](#)

<https://lter.github.io/som-website/index.html>  
<https://cosima.nceas.ucsb.edu/lter-som/>  
<https://github.com/lter/lterwg-som>

# Data Synthesis Challenges 2.0

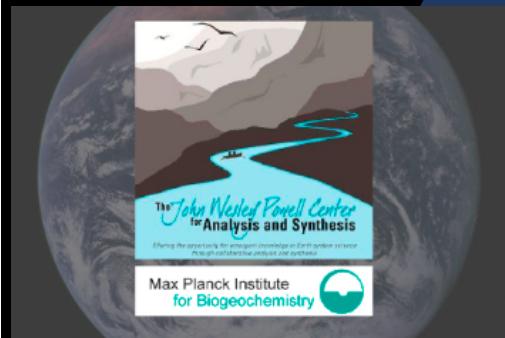
## Mapping ‘bottle neck’

- Controlled vocabulary, dictionary
- Mapping from EML metadata into harmonized data product
- Maintaining parent-child relationships





Sampling breadth



Fractions

+  $^{14}\text{C}$

Manipulations  
+ Time

Coordination  
among  
networks?

