Utilizing Interdisciplinary Strategies for Next Generation Ecosystem Experiments Tropics Data Organization

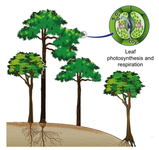
Emily Robles, Deb Agarwal, Danielle Christianson, Boris Faybishenko, Robinson Negron Juarez,
Gilberto Pastorello, Charuleka Varadharajan
Lawrence Berkeley National Laboratory

OUR PROJECT

The goal of NGEE-Tropics is to develop a predictive understanding of how tropical forest carbon balance and climate system feedbacks will react to changing environmental drivers over the 21st Century. Data collected by NGEE-Tropics researchers offers insight into how tropical forests in Central and South America respond to..

Drought

Anthropogenic Disturbance



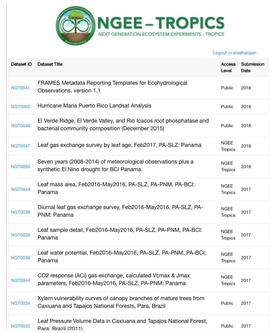
CO_2 and Temperature

Hydrology

Natural Disturbance

Nutrient Constraints

A variety of data types are collected including sapflow, ecohydrological, and meteorological measurements. The data team collaborates with researchers to curate their data packages before approval and publication on the NGEE-Tropics archive.

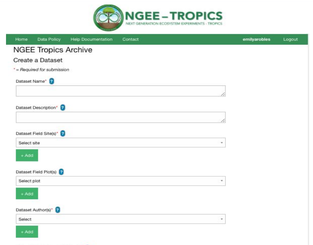


GOALS

1. Design an easy to use data archive with a streamlined submission process
2. Define standards for reporting file and package level metadata
3. Promote researcher engagement in data curation
4. Improve the quality, longevity, and reproducibility of NGEE-Tropics data

THE SUBMISSION AND APPROVAL PROCESS

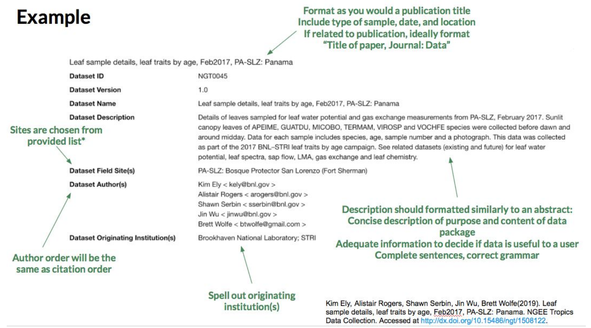
Through the data archive, users can internally curate and publish data with a digital object identifier (DOI). Package level metadata is easily collected and organized using our data submission form, which is visible to archive users before downloading any data.



1:1 meetings and project wide presentations are used to train scientists on the elements of quality data packages and include topics such as..

- Package organization
- Submission steps and demonstrations
- Data package and file examples

Example



PUBLICATION WORKFLOW

Researcher collects data

↓

Data contributor creates a data package through the archive portal and completes the submission form

↓

Data team performs thorough checks, ensuring quality expectations are met

↓

Feedback is provided directly to the data contributor

↓

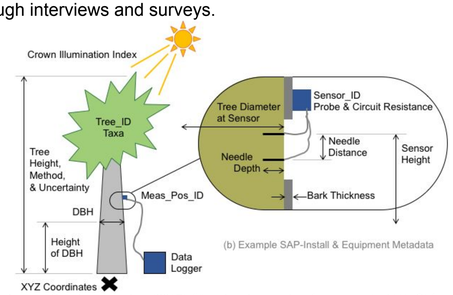
Once any necessary changes have been made, the data package is approved

A COMMUNITY CENTERED APPROACH

Interdisciplinary group work and community outreach were utilized to meet our main objectives.

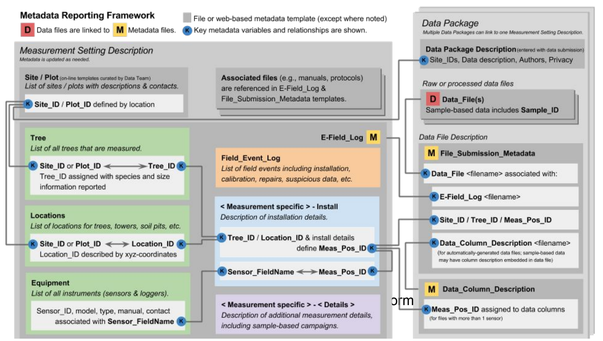
FILE LEVEL METADATA

To record file level metadata, the NGEE Tropics Archive and metadata reporting templates (FRAMES) were designed using user-experience research methods to incorporate user feedback through interviews and surveys.



Danielle Christianson, et al. (2017). "A metadata reporting framework (FRAMES) for synthesis of ecohydrological observations."Ecological Informatics Volume 42, November 2017, Pages 148-158.

The resulting three Excel and online templates describe the measurement setting, data collection, and data file organization. This standardization enables cross-site comparison for different sensor types in various formats, QA/QC, and processing levels.



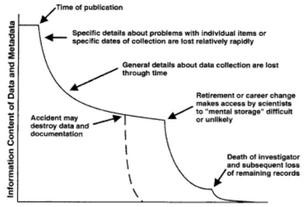
PACKAGE LEVEL METADATA

Package level metadata for each dataset is reviewed using a series of quality checks. These expectations align with community agreed standards, including those implemented by data repositories and scientific journals, and are tailored specifically for NGEE-Tropics.

A focus on community input ensured that these standards fit within the existing workflows of researchers.

ONGOING OBJECTIVES

1. Create data packages with sufficient metadata for reuse by researchers to answer multiple scientific questions
2. Maximize the longevity of NGEE-Tropics data to increase its impact
3. Increase awareness and prioritization of data package quality through educational opportunities for research teams



Michener, William K., et al. "Nongeospatial Metadata for the Ecological Sciences." *Ecological Applications*, vol. 7, no. 1, 1997, pp. 330–342.

The scientific community benefits from your data, and **"No data set is perfect and self explanatory"** without complete metadata to accompany it.

Don't let your data die with you!

IMPACT

The NGEE-Tropics archive currently has..

107 total data packages, **45** of which are publicly available, and **172** unique users.
Data packages on the archive have been downloaded **1316** times.

A focus on generating high quality metadata as part of creating the NGEE-Tropics data legacy will benefit the tropical research community for decades into the future.

ACKNOWLEDGMENTS

Danielle Christianson, et al. (2017). "A metadata reporting framework (FRAMES) for synthesis of ecohydrological observations."Ecological Informatics Volume 42, November 2017, Pages 148-158. Accessed at http://dx.doi.org/10.15486/ngt/1419956.

Kim Ely, Alistair Rogers, Shawn Serbin, Jin Wu, Brett Wolfe(2019). Leaf sample details, leaf traits by age, Feb2017, PA-SLZ: Panama. NGEE Tropics Data Collection. Accessed at http://dx.doi.org/10.15486/ngt/1508122.

Michener, William K., et al. "Nongeospatial Metadata for the Ecological Sciences." *Ecological Applications*, vol. 7, no. 1, 1997, pp. 330–342. *JSTOR*, www.jstor.org/stable/2269427.

This work was supported by the U.S. Department of Energy, Office of Biological and Environmental Research, as a part of the NGEE-Tropics