

Poster# 77

NGEE Tropics Data Collection and Management

Danielle Christianson¹, Charuleka Varadharajan¹, Val Hendrix¹, Megha Sandesh¹, Bradley Christoffersen², Matteo Detto^{3,4}, Boris Faybishenko¹, Kolby Jardine¹, Robinson Negron-Juarez¹, Bruno Gimenez⁵, Gilberto Pastorello¹, Thomas Powell¹, Jeffrey Warren⁶, Brett Wolfe³, Jeffrey Chambers^{1,7}, Lara Kueppers^{1,7}, Nathan McDowell², and Deb Agarwal¹

¹ Lawrence Berkeley National Laboratory

² Los Alamos National Laboratory

³ Smithsonian Tropical Research Institute

⁴ Princeton University

⁵ National Institute of Amazonian Research (INPA)

⁶ Oak Ridge National Laboratory

⁷ University of California, Berkeley

Contact: Danielle Christianson [dschristianson@lbl.gov]

The NGEE Tropics Data Team works closely with the project's scientists to 1) efficiently generate and synthesize ecological, hydrological, and meteorological datasets from tropical forests in Central and South America for scientific analysis and model parameterization / benchmarking, 2) host all project data, in a community accessible archive and release those data to the public after 18 months in compliance with the NGEE Tropics Data Policy, 3) standardize data and metadata collection, and 4) create priority data products such as meteorological model drivers, perform Quality Assurance/Quality Control (QA/QC) of field based observations, and synthesize sapflux measurements collected across 9 field sites collected during the 2015-2016 El Niño.

Project data are archived using the NGEE Tropics Archive tool, which allows users to upload and search data packages. Currently, NGEE Tropics team members and collaborators can create, save, edit, and submit draft data packages of any content and provide essential metadata in a user-friendly form. Once curated, the data package's metadata are visible to all users of the Archive tool. Three levels of access are enabled for the data files in the package – Public, NGEE-Tropics and Private. The system enables users to obtain a NGEE Tropics custom DOI for their data. Currently there are a total of 21 records in the NGEE Tropics Archive, with 4 being publicly shared (with DOIs).

The metadata reporting framework, implemented in conjunction with the NGEE Tropics Archive enables cross-site and cross-method comparison, data interpretability, and QA/QC. The metadata reporting framework leverages several existing metadata protocols, such as AmeriFlux and Smithsonian Tropical Research Institute. It has been implemented for sensor-based observations, such as sapflux, leaf surface temperature, and soil water content. The framework consists of templates that define a multi-scale measurement position hierarchy, descriptions of measurement settings, and details about data collection and data file organization. The framework also enables data providers to define data-access permission settings, provenance, and referencing to enable appropriate data usage, citation, and attribution. Six core NGEE Tropics field sites in Brazil, Panama, and Puerto Rico, as well as collaborators, have used the metadata templates to submit data packages to the NGEE Tropics Archive.

The NGEE-Tropics Archive and metadata reporting templates enable long-term preservation of project data, enforce compliance with the NGEE Tropics data policy, and also provide the organization and standardization needed to transform diverse and complex ecohydrological data into scientific understanding.