Smithsonian Environmental Research Center Global Change Research Wetland Data Management Plan

I. Policies for Data Sharing and Public Access

Our data sharing and public access policy is consistent with Creative Commons Attribution 4.0 International (CC-BY 4.0), a public license that makes the data freely available provided the source of the data are attributed properly, and that the user does not claim an endorsement by the Smithsonian Institution for their specific application of the data. The policy was adopted because the data are collected by both US Federal and non-Federal employees, making the application of Federal copyright law to the data unclear. CC-BY 4.0 is used by government agencies and non-profit organizations (https://creativecommons.org/licenses/by/4.0/). These polices are consistent with the SI Public Access Plan (V1.2).

II. Data Collected, Formats and Standards

Raw data is archived after it has been subjected to standardized QA/QC processes. Derived data are calculated from raw data, which we archive along with the code (SAS or R) used to execute the calculations. Derived data are reduced by replicated unit (plot or chamber) and reported as single values that represent the unit. We do not archive treatment-level summaries of the data as they are easily calculated by the user. The data are stored in Excel-format files, each of which also contain a Metadata worksheet listing variable names, variable units, missing data codes, and accessory information that places the data into context. We adopted the metadata standards used by Harvard Forest (http://harvardforest.fas.harvard.edu/data-archive/policies-guidelines). Raw data collected on paper are scanned and archived on a Smithsonian Institution network (see below) in Archive PDF (pdf/a) format. These scanned documents are not posted for public download but are available from the managing PI upon request. The data are archived and publically available two years (or less) after the samples are collected. This period allows 12 months for analysis and QA/QC, and 12 months for publication of novel results. All files are posted on a Smithsonian-maintained site dedicated to the project (http://serc.si.edu/GCREW). Archived variables are listed in Table 1.

III. Data Storage, Preservation, and Dissemination

Standard procedures for processing samples, entering raw data, performing QA/QC analyses, reducing data, and tracking corrections are documented in files associated with the respective datasets on the Global Change Research Wetland website (http://serc.si.edu/GCREW). All files posted on the website are assigned a name that ends in the posting date (dd-mm-yyyy). Once posted, the original raw and derived data files (used to generate the posted files) are locked by the managing PI (see below) as read-only. Any subsequent changes to the file are made by the managing PI and logged into a "Change Log" worksheet found in the same Excel file that contains the data and the metadata. Any changes to data or metadata made after the file was locked triggers a new version of the dataset identified by the new date. Retired versions of the file are not archived on the public website, but are available upon request.

All data files are saved to a Smithsonian Institution network drive; we do not depend on local drives other than for temporary transfer from sensor, device or instrument memory to the network. Changes to files stored on the network drive are backed up by the Institution every evening, and a complete backup occurs every weekend. Backup tapes are stored offsite and the backup data are kept for one year.

Primary responsibility for managing the data rests with Dr. Patrick Megonigal, a PI on the LTREB grant and a full time employee of the Smithsonian Institution since 2000. Responsibility for the data will be passed to a Smithsonian Institution colleague when required, as was done upon the retirement of Dr. Bert Drake who began the longest of the datasets in 1986. We intend for the data to be available in perpetuity.