

PI: Shawn Landry, Ph.D.	Department: Geosciences
Institution: University of South Florida	Agency: National Science Foundation

Data Management Plan

I. Types of Data

The research team is committed to compliance with NSF policies on the preservation, dissemination, and sharing of research data. The data inputs and outputs are detailed in the table below. Relevant software include R, SAS, and SPSS for statistical analysis, ArcGIS for spatial analysis, QT Modeler for LiDAR data processing, and i-Tree ECO for field plot data management. In general, the products of the research will include conference presentations, peer-reviewed publications, and presentations to urban forest managers.

Data Inputs	Data Outputs
Tree attributes collected at sampled field plots in Tampa, Gainesville and Orange County.	Statistical data analysis, results, and reports. Field plot data included in i-Tree ECO database submitted to the iTreertools.org data repository.
Pre- and post-storm estimates of tree canopy cover in buffers surrounding sampled field plots.	Statistical data analysis, results, and reports.
Pre-storm very high resolution tree canopy cover data for Tampa, with height above ground of tree canopy and building elements (i.e., normalized digital surface model). Available Census and other ancillary spatial data.	Statistical data analysis and results. Spatial data (other than Census) will be made available as data layers in an Open Data repository of the USF Water Institute (directed by co-PI Landry) at http://waterinstitute.usf.edu/data-and-maps .

II. Data and Metadata Standards and Dissemination Methods

Metadata will be documented using the Ecological Metadata Language standard (EML) using the open-source software package Morpho (<https://knb.ecoinformatics.org/#tools/morpho>). Metadata will include the basic descriptors of the whole data set (e.g., title, originator); the research activities (site locations and descriptions, experimental design and methods, names of variables, catalog of samples, etc.); data set status (latest update and verification, proprietary rights and accessibility, etc.); and data structure (file sizes and formats). Researchers will document metadata for each new data set developed in the study, and publish the metadata on the Knowledge Network for Biocomplexity repository. For example, Shawn Landry. 2015. Tree Canopy Cover Data, Tampa, FL USA, 2006. Knowledge Network for Biocomplexity. shawnlandry.3.2. (<https://knb.ecoinformatics.org/#view/shawnlandry.3.2>). Tree canopy and other related spatial (i.e., GIS) data layers will also be incorporated into the USF Water Institute (directed by co-PI Landry) Open Data repository at <http://waterinstitute.usf.edu/data-and-maps>.

III. Policies for Access and Sharing and Provisions for Appropriate Protection/Privacy

Data Sharing Policy

Data to be shared include input and output data, and spatial analysis, results, and maps and KNB repository of metadata. The data will be shared at the end of the project and the restrictions will require approval from the project team. Journal, conference, and/or thesis and doctoral dissertation publications (if generated) will be shared in PDF format upon publication with no restrictions.

Privacy. There are no privacy concerns with the data collected as part of this study. No data related to human subjects will be collected, other than the publicly available sociodemographic datasets provided by the Census American Community Survey.

Data Citation Policy. Data should be cited using the standard APA citation style when used by others.

IV. Policies and Provisions for Re-use, Re-distribution

The majority of data will be made publicly available, either as soon as available in the case of existing datasets or after publication in order to temporarily protect intellectual property. All pre-storm i-Tree ECO data has already been submitted to the i-Tree website and according to the data sharing agreement within the i-Tree ECO software the data are already publicly available (<http://www.itreetools.org/>). Metadata of all new datasets will be made publicly available through the KNB. The purpose of open sharing of data is for the re-use and re-distribution and production of derivatives (with appropriate citation of our publications and databases) by the scientific community and other NGO and government agencies.

V. Resources and Facilities for Data Storage, Preservation and Archiving

For long-term storage, researchers have institutional file storage systems with replication/backup and redundancy. Systems are on scheduled backup cycles and include back-up generator in case of power outages.