2-2-2017

CAREER: Robust aquatic habitat representation for water resources decision-making

Sarah Null

Utah State University, sarah.null@usu.edu

DATA MANAGEMENT PLAN

Data management is a core component of this project. Many NSF and Utah State University (USU) programs focus on big data and on data management. This project will take advantage of existing data management infrastructure to support the full data life cycle. Goals of data management for this project include: a) protect data from corruption or loss, b) transform raw data into cleaned and error-checked datasets, c) provide and maintain metadata to accompany datasets, and d) share data and models with the larger scientific community using the Consortium of Universities for the Advancement of Hydrologic Science (CUAHSI)'s HydroShare platform (see Data Publication and Sharing section below).

The research proposed requires collection, acquisition and/or management of data from a broad range of disciplines. Spatial environmental data includes physical river characteristics, spatially-distributed streamflow, water quality, fisheries presence and abundance, bioenergetics or other biotic characteristics, barrier infrastructure, and habitat networks. Additional acquired data includes economic loss functions for water demands and pricing. Data derived from this project includes measured physical habitat and water quality, habitat imagery (thermal infrared, LiDAR, and visible imagery), comparisons of modeled and measured habitat representation for target species, model formulation, model code, model results, sensitivity analyses quantifying robust decision-making, tradeoffs between economic benefit and environmental harm of instream barriers, educational museum products, and evaluations on increasing public science literacy through museum exhibit education and outreach activities.

Metadata: Metadata is important with modeling studies so that input data sources and uncertainty regarding input data is transparent to future model users and developers. HydroShare and other data discovery systems have developed rigorous metadata standards, which will be adopted. For example, each published dataset will have unique metadata specific to the data type and relationship to other data sources.

Data Storage and Backup: Dr. Null will be responsible for retaining all measured data, model input data, model code, model results and output, and public science literacy evaluations. All data will be compiled on computers in Dr. Null's laboratory. Data will be backed up to external hard drives and to cloud storage at least once per week. Final data products will be stored using CUAHSI's HydroShare.

Data Publication and Sharing: Data will be disseminated using the following methods:

1) Data and final models will be published using CUAHSI's HydroShare platform and shared with the wider water resources and hydrologic community. HydroShare is a collaborative environment for sharing hydrologic data and models to give water researchers and managers the technological infrastructure they need to address critical issues related to water quantity, quality, accessibility, and management. There are no privacy, copyright, or confidentiality issues associated with the project data (water demand and price data will be aggregated for large regions). Data will also be shared with fellow researchers, as requested. Interested data consumers and researchers will be required to sign a data use contract that requires them to acknowledge the PI and NSF in any resulting publications or presentations using shared data.

2) The PI will also maintain a water resources systems modeling webpage (on Dr. Null's faculty website) with pertinent systems modeling course syllabi, news and press-releases for museum exhibits, photographs of data collection, news and press-releases for newspaper articles or publication releases, cleaned and error-checked models, graphical user interfaces (GUIs) that have been developed from project research, and thought-provoking syntheses of barrier removal systems modeling research, training, and public science literacy. Video recordings of faculty and student talks will also be posted to the website when available.

3) PI and student trainees will present data, results, and major findings at national and international meetings such as the Weber River Watershed Symposium, American Geophysical Union (AGU) Fall Meeting, or International Society for River Science.

4) The PI and students will publish results in peer-reviewed journals in the fields of ecological engineering, water resources management, river restoration, and systems analysis, enabling a broad distribution of findings and scientific advances across sub-disciplines.

5) Papers will be stored as pdfs on the Digital Commons site of Utah State University's Merrill-Cazier Library (http://digitalcommons.usu.edu/). Digital Commons is a repository for publications, data, video, sound, executable and other file types, managed by the Berkeley Electronic Press. Content in Digital Commons is optimized for fast and accurate indexing by Google and Google Scholar. Assigned URLs remain unchanged indefinitely.

Student Data Management Training: The PI will recommend that graduate students funded on this project enroll in Hydroinformatics, a 3 unit, graduate-level course offered through the Civil and Environmental Engineering Department at USU. The course teaches data management software, metadata and semantics, data storage file formats and standards, and web-based data distribution and access using web services. Further, the PI will train undergraduate and graduate students to publish and share data and models using HydroShare. This will teach students data management skills that they will need in technical ecological engineering and systems modeling fields. It will allow them to interact with other researchers in the CUAHSI and HydroShare communities.