

Data Management and Access Plan

Data management activities for this project will take advantage of the ongoing efforts of the LTER and/or ULTRA projects ongoing at each research sites. Data collected at each site will be incorporated into the information management systems at these sites as well as into a centralized, integrated data system created for this project. The multi-disciplinary nature of this system will represent a novel data management activity that should be of interest to emerging coupled socio-ecological research activities in networks around the U.S., i.e. LTER, ULTRA, NEON, WATERS.

Types of data and samples to be managed in the project:

- **Biophysical data.** Soil carbon and nitrogen pools and processes, plant community composition, aboveground biomass and net primary production, sediment denitrification potential, soil climate (temperature, moisture). Soil samples from each site will be shipped to the Cary Institute (Peter Groffman) and analyzed for total carbon and nitrogen content, potential net N mineralization and nitrification and soil inorganic N pools. Subsamples will be shipped to UC Irvine (Diane Pataki) for isotope analysis. Sediment samples will be shipped to Florida International University (FIU, James Heffernan) for denitrification assays. Soil climate data collected at each site will be sent to Arizona State University (ASU, Sharon Hall) for processing and collation. Plant data collected at each site will be sent to the University of Minnesota (UMN, Jeanine Cavendar-Bares and Sarah Hobbie) for processing and collation.
- **Surveys.** There will be two surveys; 1) a household telephone survey that includes questions about land management practices, knowledge and sources of information for land management, watershed knowledge, willingness to participate in environmental activities, environmental perceptions, and perceptions of neighborhood environmental quality of life and 2) an in-depth, in-person sample of households that will target their lawn management practices, socioeconomic and demographic characteristics, and environmental attitudes. In particular, the survey will identify fertilizer and outdoor water use practices. Colin Polsky (Clark University) will take the lead on compiling data from the telephone surveys and Laura Ogden (FIU) and Rinku Roy Chowdhury (Indiana University) will co-direct the design and data compilation from the in-person surveys that will be done in each city.
- **Parcel structure remote sensing.** For each study site, we will develop an object-based image analysis (OBIA) system that extracts land cover information using the best available remotely sensed data and ancillary GIS data layers. The land cover classes will consist of: tree canopy, grass/shrub, bare soil, water, buildings, roads/railroads, and other paved surfaces. This work will be conducted at the University of Vermont Spatial Analysis Laboratory (UVM SAL) under the supervision of Jarlath O'Neill Dunne (UVM SAL) and Rinku Roy Chowdhury (Indiana University).
- **Scaled products.** Project data will be used to produce; 1) Regional (MSA)-scale estimates of the effects of urban land use change on soil, vegetation and whole-ecosystem pools of carbon and nitrogen (total and reactive), plant diversity (native and exotic, functional and species) and landscape denitrification potential; and 2) Preliminary continental scale estimates of the above effects based for the vast majority of the US occupied land area (excluding Alaska and Hawaii). This work will be led by Colin Polsky along with Morgan Grove (U.S. Forest Service) and Peter Groffman.

Details of sample collection, transport and analysis of each of the data streams listed above are presented in the "project description" and "project management" sections of this proposal. Soil, sediment and

vegetation samples will be stored at the institutions where they are processed at least until data analyses are completed.

Standards for data format and metadata:

All biophysical data will be documented using Ecological Metadata Language (EML), which is standard for the LTER network. The metadata will be available via the LTER metadata catalog. Data collected at individual LTER sites will be entered into the information management systems at each site as well as into an integrated project system (described below).

All geospatial datasets will be stored in formats recognized by the Geospatial Data Abstraction Library (GDAL) and documented according to the standard specified by the Federal Geographic Data Committee (FGDC). Datasets relevant to individual LTER sites will be entered into the information management systems at each site as well as into an integrated project system (described below).

We will combine parcel-level data from (a) vegetation and homeowner surveys with (b) other relevant measures of the social landscape (e.g., Census data, PRIZM lifestyle characteristics, zoning/land-use restrictions), and (c) the biophysical measurements to produce a combined multi-scale, multi-site dataset. This data synthesis will draw on our substantial experiences producing similar combined georeferenced ecological-social-locational databases in the Baltimore, Boston and Twin Cities metropolitan areas. The leadership for this complex geo-data management task will be shared by the UVM-SAL and the Clark University HERO program (Boston).

Policies for access:

Data collected at individual LTER sites will be entered into the information management systems at each site and will be subject to standard LTER access policies, e.g., public access in one year for most datasets. PRIZM lifestyle data represent a special circumstance where data access is constrained by a license agreement and will be kept internal to project participants. Survey data will be stripped of all personally identifiable information to protect the privacy of the interview subjects. Links to the integrated database established for this project will be added to all LTER and/or ULTRA sites involved.