**iUTAH Data Collection Plan**

**Effort Name:**

High frequency monitoring of environmental variables at iUTAH GAMUT sites

**Collaborators:**

IUTAH GAMUT Working Group:

Zach Aanderud

Michelle Baker

Dave Bowling

Jobie Carlisle

Chris Cox

Joe Crawford

Jim Ehleringer

Dave Eiriksson

Jeffery Horsburgh

Amber S Jones

Scott Jones

**Brief Summary (including data collection methods, timing and location of data collection, and the parties responsible for data collection):**

Suites of environmental sensors will be deployed to measure climate, hydrologic, and water quality variables at stream sites and terrestrial sites within the three GAMUT watersheds (Logan River, Red Butte Creek, Provo River). There will be 8-12 sites in each watershed depending on access and spatial variability. To the extent possible, site instrumentation and program deployment will be standardized. Data will be generated continuously and with high frequency (~15 minutes). Data collection will be ongoing for the duration of iUTAH. The Watershed Technicians will be responsible for the data collection.

Raw sensor data will be quality controlled and post-processed by the Technicians within six months of data collection. Stage-discharge relationships will be developed at each aquatic site, and discharge will be derived at the same schedule as quality control performance.

**Individual(s) Responsible for iUTAH Metadata Completion:**

Watershed technicians: Dave Eiriksson, Chris Cox, Joe Crawford

**Datasets Expected to be Generated:**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dataset** | **Data Type\*** | **Method of Creation** | **Resulting Data Format** | **Data Storage** | **Final Data Product** | **Timeframe\*\*** | **Access During Collection** | **Access After Completion** | **Anonymization for IRB** |
| Raw Sensor Data | Type A | Sensor results logged by datalogger | Datalogger files, Observations Data Model (ODM) databases | IUTAH Servers | ODM Databases with web tools for access | Raw sensor data will be available immediately upon data collection. | All data will be publicly accessible. | NA | NA |
| Quality Controlled Sensor Data | Type A | Technicians use ODM Tools Python to post process sensor data | Data series contained in ODM databases, Python scripts of edits | iUTAH Servers | ODM databases with web tools for access | Quality control and post processing of the raw sensor data will be completed within six months of initial collection. Release of the previous six months will occur quarterly: Jan 1, April 1, July 1, October 1. | Technicians and data managers will have access to quality-controlled data prior to release. | After each release, the data will be publically accessible. | NA |
| Derived Data Product: Discharge | Type A | Technicians use ODM Tools Python to derive discharge based on sensor data | Data series contained in ODM databases, spreadsheets with stage-discharge relationships | iUTAH Servers | ODM databases with web tools for access | Initial stage-discharge relationships will be established by technicians by 9/1/2014.  Discharge will be derived within six months of stage measurement. Release of the previous six months will occur quarterly: Jan 1, April 1, July 1, October 1. | Technicians and data managers will have access to relationships and derived data prior to release. | After each release, the data will be publically accessible. | NA |

**\*Data Typology:**

**Type A** - Primary iUTAH datasets or research products. These include raw and QA/QC calibrated sensor data from iUTAH facilities, baseline sampling datasets across iUTAH facilities and sites, and general datasets collected by iUTAH for the community of iUTAH participants.

**Type B** - Datasets or other research products for which monetary or material support was provided by iUTAH, but that are created by a specific investigator, student, or iUTAH group to support a particular research question or goal.

**Type C** - Type A and Type B datasets or products that include personally identifiable information or information about human subjects/participants and are subject to IRB restrictions.

**Type D** – Datasets or other research products procured by iUTAH or iUTAH participants supported by iUTAH that are subject to licensing, copyright, or use restrictions/agreements from the data source that may prohibit general distribution of the data.

**\*\*Data Typology Timeframe**:

**Type A** - Automated data streams from iUTAH facilities will be streamed directly into live databases and will be made available online in near real time. Quality controlled and derived data products from iUTAH facilities will be published within six months of data collection. All other primary datasets will be published as they become available (e.g., as soon as results are created).

**Type B** - Finalized data will be submitted to the iUTAH MDF within one year of the completion of data creation activities. Students collecting data must submit their finalized data as a condition of their thesis/dissertation defense. For long running data creation activities (i.e., efforts that last longer than one year), the following will be required:

1. The initial metadata description will be reviewed and updated at least once per year.
2. Data will be submitted at least every 6 months. Data will not be released until the dataset is finalized by the data creator.
3. Finalized data will be submitted within one year of collection or by the end of the project, whichever comes first.

**Type C** - Type C datasets will be subject to time requirements described for Type A and Type B datasets. However, they may require the additional step of anonymization or aggregation with methods described in the Data Collection Plan.

**Type D** - Type D datasets will be published as soon as possible (within one month) and to the greatest extent allowable by the licensing, copyright, and/or data use agreements under which they were created/procured.  Some Type D datasets may be permanently restricted.