# Working Group Report – LTER All Scientists Meeting 2015

## Session Information

Title: Automatic Sensor Data Processing and Q/C with the GCE Data Toolbox

Organizer: Wade Sheldon (GCE LTER), Co-organizer: Brian Herndon (CWT LTER)

Date: Sep 02, 2015 08:00-09:45

## Overview

The working group session was attended by 24 people (including organizers), representing 12 different LTER sites or organizations (see Attendees). Approximately half of the participants were completely unfamiliar with the GCE Data Toolbox software, and the remainder included a mix of new and experienced users.

Wade Sheldon opened the session by presenting an introductory PowerPoint slideshow (1). He began with the creation and evolution of the GCE Data Toolbox software, which is a comprehensive open source software framework for metadata and data processing for Mathwork's MATLAB® software platform. An overview of the Toolbox's design and functionality was presented as well as descriptions of the supporting documentation, Trac Wiki web site (https://gce-svn.marsci.uga.edu/trac/GCE_Toolbox), and other existing resources for learning to use the Toolbox and getting support.

This segued into Brian Herndon giving a user "testimony" of his experience implementing this software in a harvesting and streaming data workflow at CWT LTER. Trials and errors as an information and data manager were discussed and how this middleware/software has been a valuable tool particularly for environmental sensor generated data.

Wade Sheldon then conducted a brief demonstration of the graphical user interface of the software. An environmental sensor-generated data file was imported into the Toolbox and various data visualization, management and Q/C functions were demonstrated. Importing datasets from other agencies (USGS, NOAA) and middleware (DataTurbine) were covered, including the options available for data, metadata and Q/C flag merging.

Wade Sheldon then presented a more detailed PowerPoint slideshow describing the GCE Data Toolbox Q/C Framework (2). Best management practices for designing Q/C criteria (rules) for data qualifying, managing flagged values, and data archiving and reporting were discussed. Ways to input and track metadata were also described and demonstrated.

Questions were fielded as they arose during the working group, and a final question and answer session finished out the session. Questions from the audience included:

1. Support for ISO and ESRI metadata standards?
2. How does this integrate with website design?
3. How to deal with harvesting large files?
4. Is there a versioning process or provenance tracking feature?
5. Is it R and Java compatible?
6. How easy is it to learn and what training materials are available?
7. How much control do you have over filtering flags?

8. Does the toolbox load all or some of the data, or does it have to be a clean parse?
9. Is this software compatible or interface with Aquarius?

**Presentations**

1. Automating Sensor Data Processing and Q/C with the GCE Data Toolbox - Wade Sheldon, Brian Herndon, Sep 02, 2015

2. The GCE Data ToolboxQ/C Framework - Wade Sheldon, Sep 02, 2015

**Outcomes**

Due to the large number of potential and new/inexperienced users attending the workshop, discussion focused on inter-operability with other software used by LTER sites (e.g. R, DEIMS, Aquarius) and user training and support rather than advanced Q/C uses cases as originally planned. Consequently, follow-up activities are focused on user engagement, future software development use cases, and future training activities, e.g.

- A total of 11 participants requested to be added to the GCE Data Toolbox email listserv (gcetoolbox-l@uga.edu) for ongoing announcements and discussion, in addition to the 8 participants that were already subscribed. This task will be completed by Sep 11, 2015.
- Interoperability with the R statistical software package and DEIMS (Drupal Environmental Information Management System) were high priorities of attendees, and will be prioritized in future toolbox development. EML-based crosswalks are already available or supported by all these tools, so EML approaches will be explored in Fall 2015, starting with DEIMS.
- A full training workshop, similar to the 2013 LTER ARRA-supported training, was also a high priority. We will explore opportunities for a 2-3 day virtual or in-person training workshop to be held in early 2016.
- A Wiki page will be added to the GCE Data Toolbox Trac website for capturing Q/C use cases and examples; this page will be created in Sept 2015 using material and examples presented in this workshop (2). Interested participants will be given access to add and edit these pages.

**Attendees**

Aaron Ellison (HFR)
Bengt Logebladh (SITES, Sweden)
Brian Herdon (CWT) – co-organizer
Don Henshaw (AND)
Ed Johnson (CWT)
Elise Osenga (Aspen Global Change Institute, USA)
Emery Boose (HFR)
Fox Peterson (AND)
Hap Garrett (PIE)
Henry Brandes (NWT)
Hope Humphries (NWT)
Inigo San Gil (MCM)

Inke Forbrich (PIE)
Jeff Patriarche (MCM)
Jennifer Morse (NWT)
John Anderson (JRN)
Ken Ramsey (JRN)
Kristin Vanderbilt (SEV)
M. Gastil-Buhl (MCR)
Mary Martin (HBR)
Ola Langvall (SITES, Sweden)
Suzanne Remillard (AND)
Wade Sheldon (GCE) – organizer
Yang Xia (KNZ)