

ODM Tools Python: Open Source Software for Managing Environmental Sensor Data

Amber Spackman Jones, Jeffery S. Horsburgh, Stephanie L. Reeder, Jacob Meline
Utah Water Research Laboratory, Utah State University



UtahState
University



CUAHSI
HIS

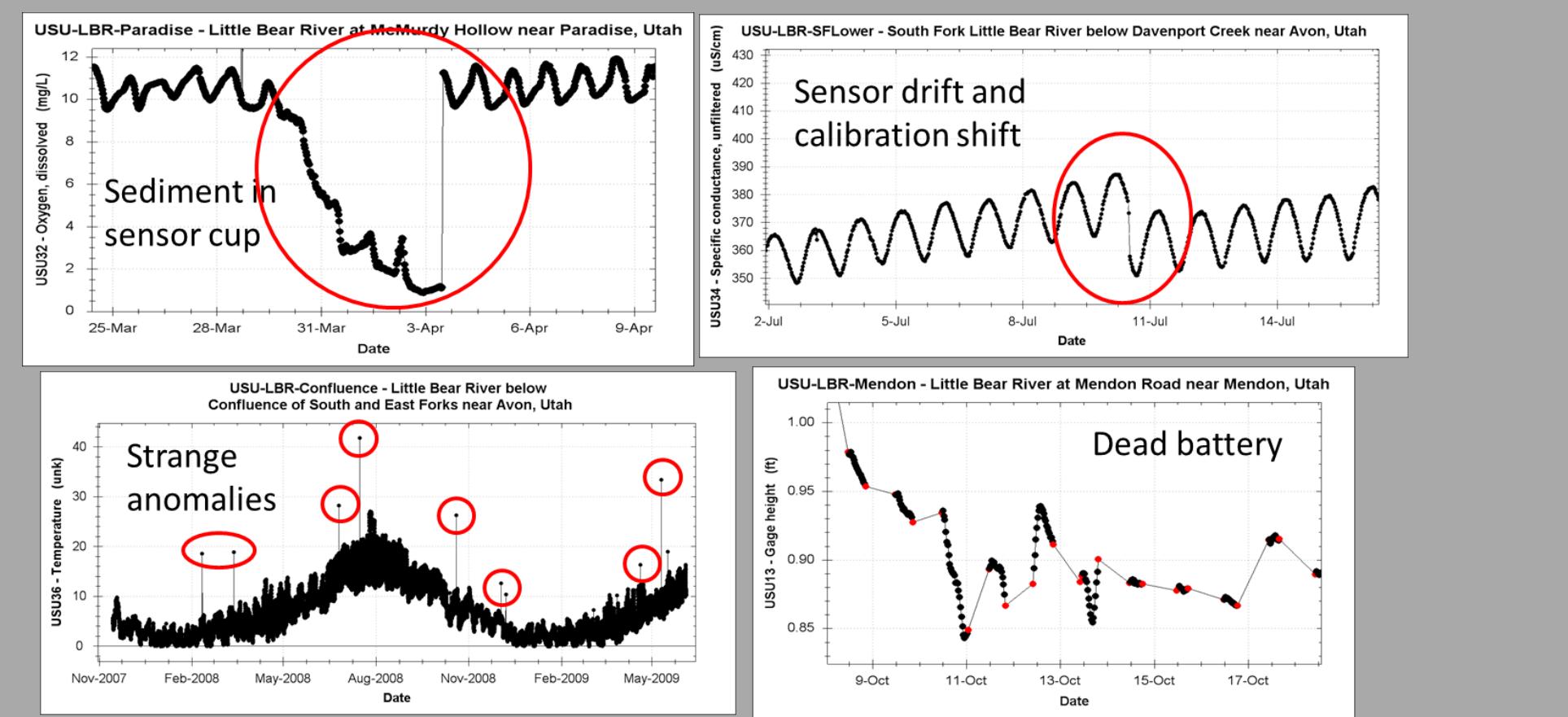


This project is funded by National Science Foundation grant EPS-1208732. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

MOTIVATION: Environmental data are being collected at high frequencies, for extended durations, and with spatial distributions requiring infrastructure for data management.

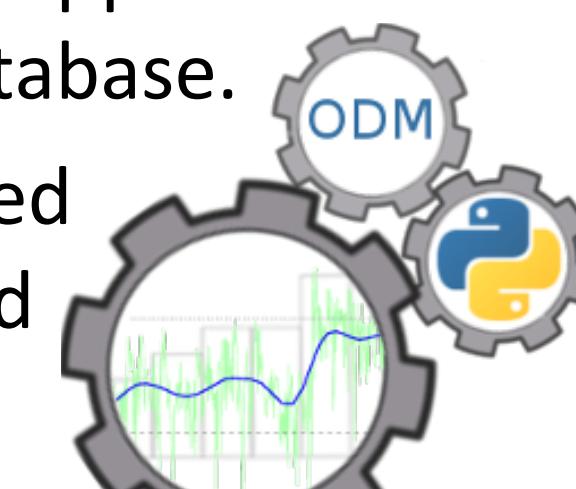


These data often need quality control and post processing performed to remove anomalous values, account for drift, and arrive at an 'approved' dataset. Standard tools for performing data QAQC enhance the ability to track data provenance and repeat edits to facilitate data reuse.



UNDERLYING FRAMEWORK:

- ODM:** a data model for organizing, storing, and describing point observations data originally developed as part of the CUAHSI HIS.
- ODM Tools Python** is an open source software application for managing data stored in an ODM database.
- The original **ODM Tools** software was developed as part of the CUAHSI HIS. We have re-designed and deployed a new version in Python.



IMPROVED FUNCTIONALITY:

- Multiple Platform Support:** Windows, Mac, and Linux
- Multiple RDBMS Support:** Can connect to ODM databases deployed with SQLServer, MySQL, or PostgreSQL.
- Query and Export:** Simple filters are implemented for data series selection. Data and associated metadata can be exported from the database.
- Dockable Windows:** Various components can be displayed, hidden, and docked for user flexibility.



OPEN SOURCE REPOSITORY: ODM Tools Python releases and code are available on GitHub
<https://github.com/UCHIC/ODMToolsPython>

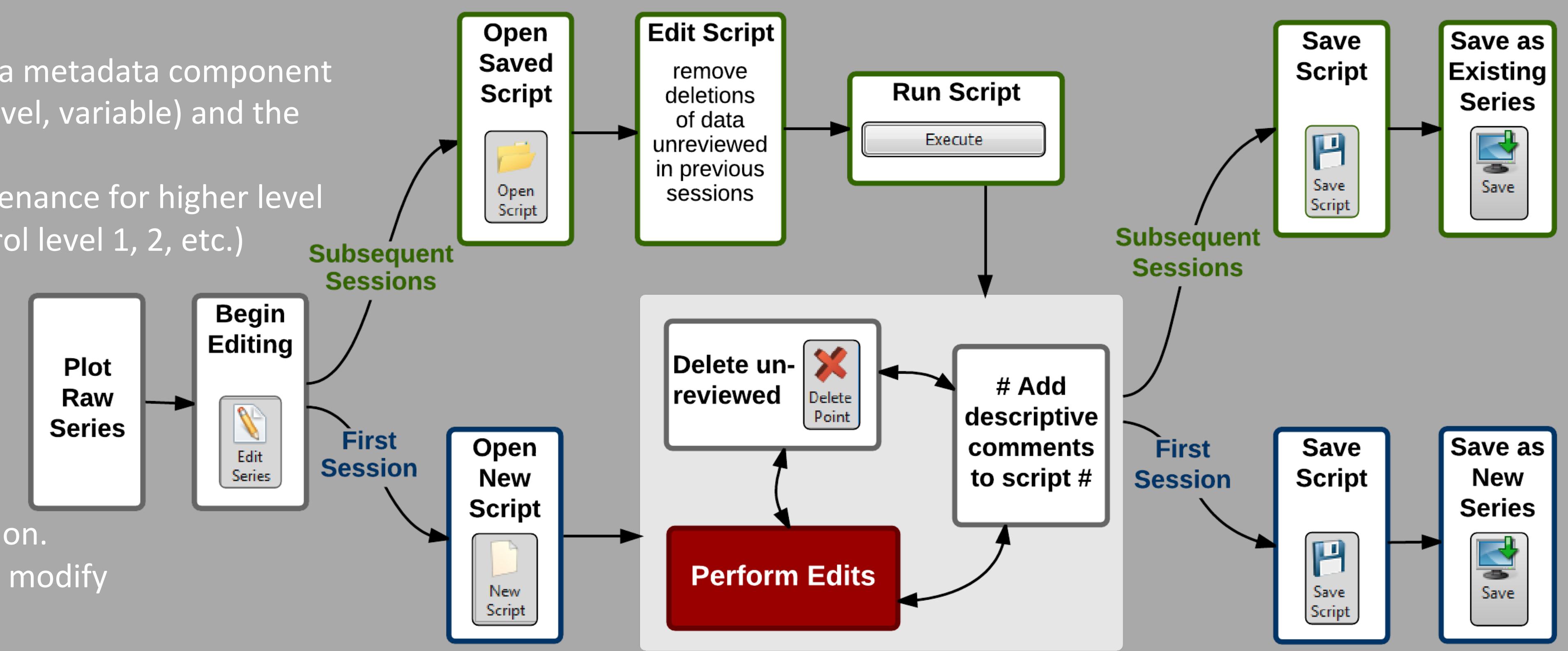
VISUALIZATION:

- Facilitates inspection of data to identify and assess patterns and problems.
- Multiple series can be plotted simultaneously to facilitate comparisons with other variables (internal consistency) and other sites (external consistency).
- Edited data immediately available for inspection after modification.
- Various plot types implemented.



WORKFLOW:

- New series can be derived with a metadata component modified (e.g., quality control level, variable) and the remaining metadata retained.
- Generating and preserving provenance for higher level data products (i.e., quality control level 1, 2, etc.) is streamlined.
- Raw data are always preserved—edits are performed on a copy of the database stored in memory.
- The editing script can be saved for future reference and execution.
- Users can compare, review, and modify the editing steps.



EDITING AND AUTOMATED SCRIPTING:

- All edits performed on a data series are automatically recorded using a Python script editor and console integrated with the ODM Tools Python interface.
- Tools are built-in for common editing tasks, and the Python console and script editor can be used to perform advanced editing tasks.
- Each action is translated into executable code that is captured in the script.
- Scripting ensures that editing steps are traceable and reproducible. Annotations to the edits can also be made as comments in the script.

