

Cyberinfrastructure Tools for Managing GAMUT Data and Infrastructure, Part 3

6/5/2014

Amber Spackman Jones
Stephanie Reeder
Jeffery S. Horsburgh

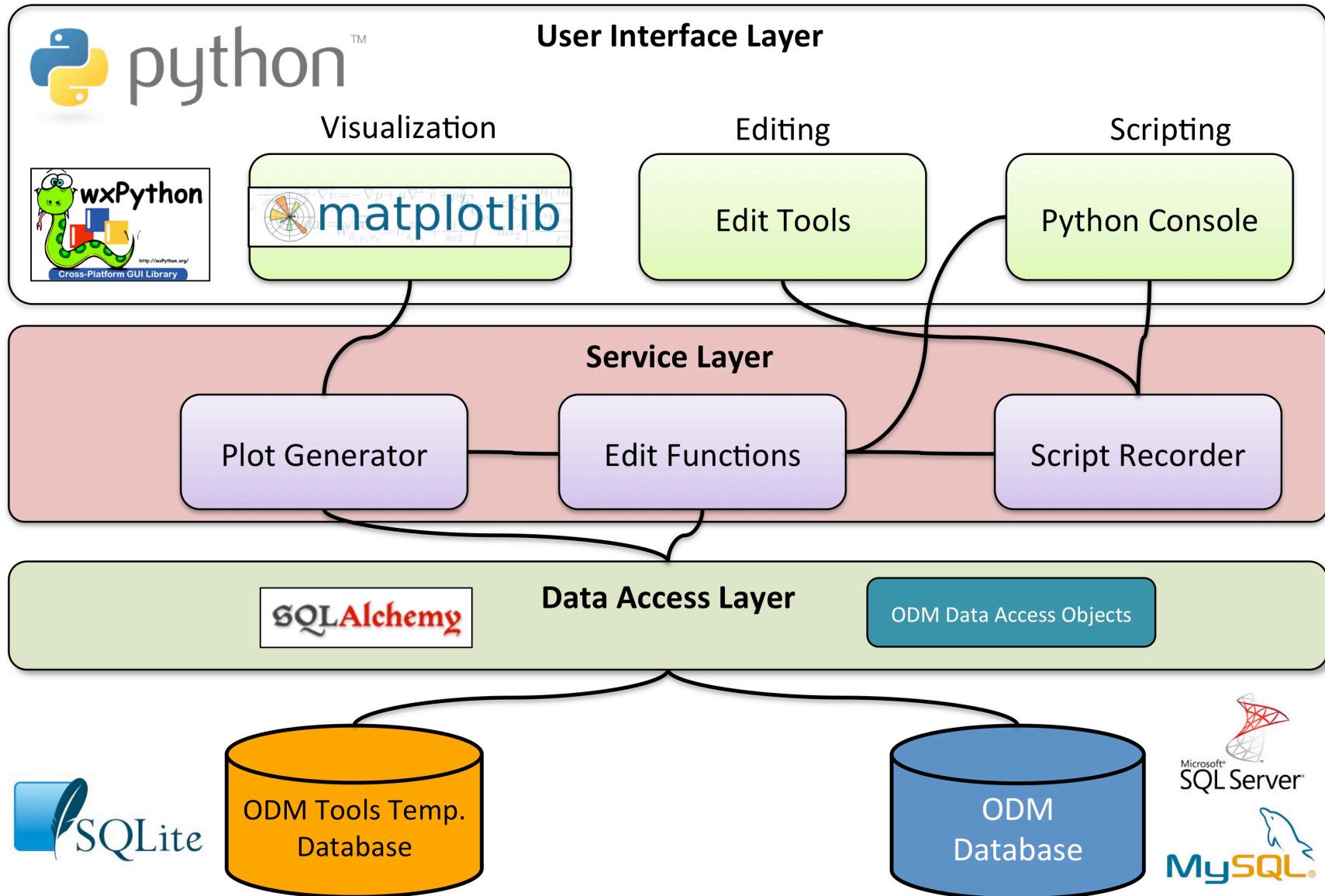
Housekeeping, Misc

- Synoptic Samples Entry: status update
- New Web Data Viewer released!
<http://data.iutahepscor.org/tsa>
- Pictures needed for each site for individual site pages:
http://data.iutahepscor.org/mdf/river_info/iUTAH_Provo_OD/PR_BJ_AA/
- Automated Alerts: what additional alerts are needed?
- QAQC Document Update: timeframe, what to include?

ODM Tools Python: Design Goals

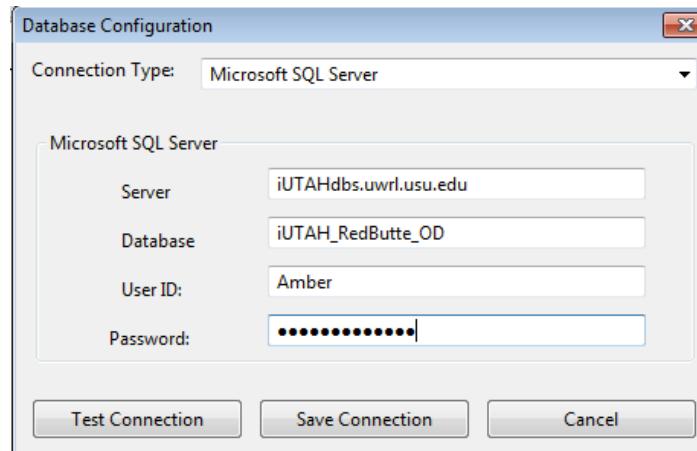
- Multi-platform support (Windows, Linux, Mac)
- Multi-database support (Microsoft SQL Server and MySQL)
- Implement a scripting interface to save the provenance of data edits in QC process
- Modernize the Graphical User Interface (GUI)

Architecture



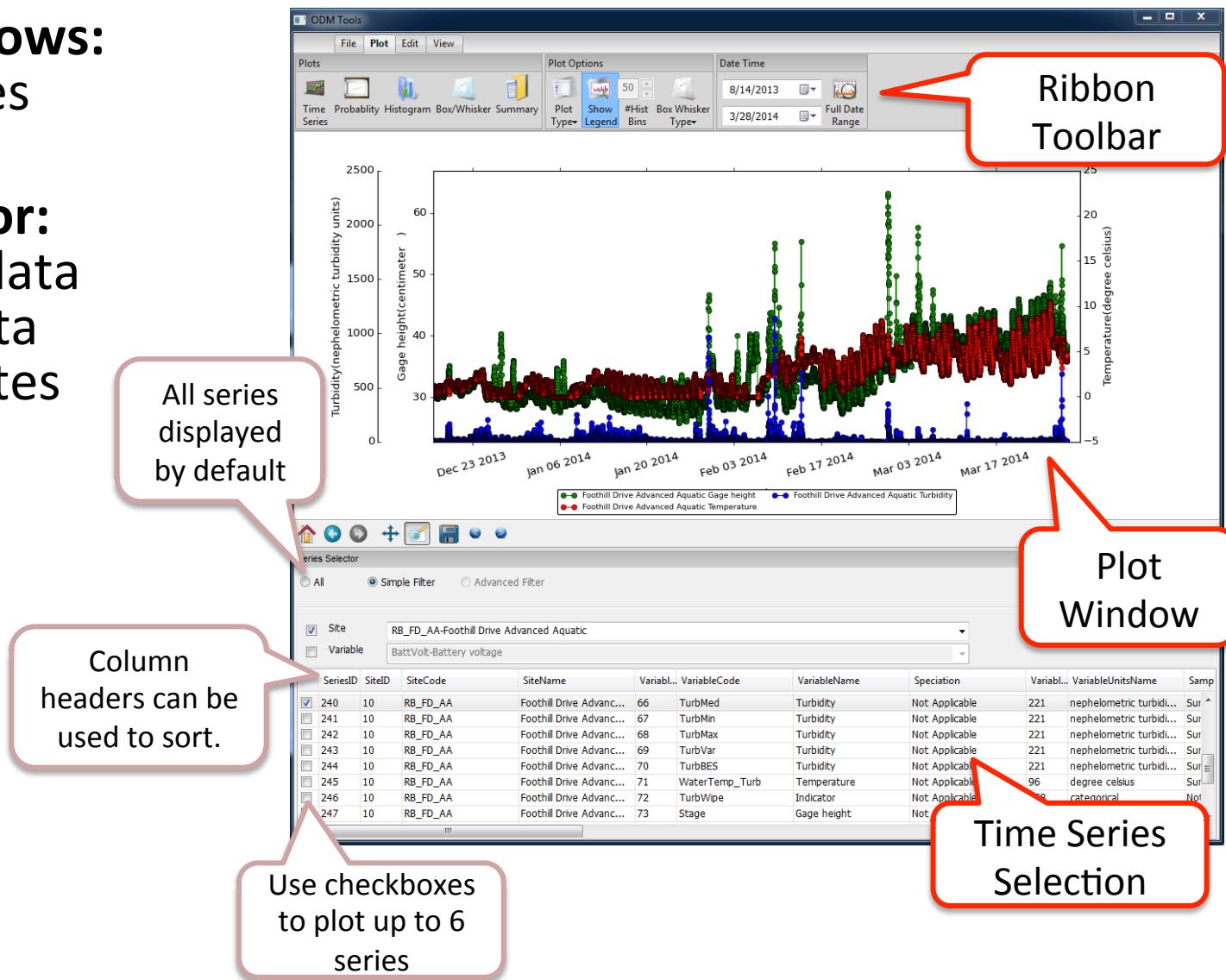
Getting Started

- Documentation:
<https://github.com/UCHIC/ODMToolsPython/wiki/ODMTools-Python-Documentation>
- Download and install application
<https://github.com/UCHIC/ODMToolsPython/releases>
- Connect to database
 - Connection information saved so that connection will persist when the program is re-opened.
 - New connection can be made using ‘File’ tab.
 - **Note:** User credentials saved to file and visible in program- take care!



Plotting: Series Selection

- **Default windows:** Plot and Series Selector.
- **Series Selector:** Table listing data series and data series attributes
- **Data Series =** Site + Variable + QCLevel + Method + Source



Plotting: Simple Filter

- The Simple Filter restricts the list of series for display and selection
- Can filter by Site, Variable, or Both
- **Note:** A series can be selected for plotting and then the filter applied to hide that series. Filter must be re-applied in order to hide series.
- **Note:** Advanced Filter not yet implemented.

Series Selector

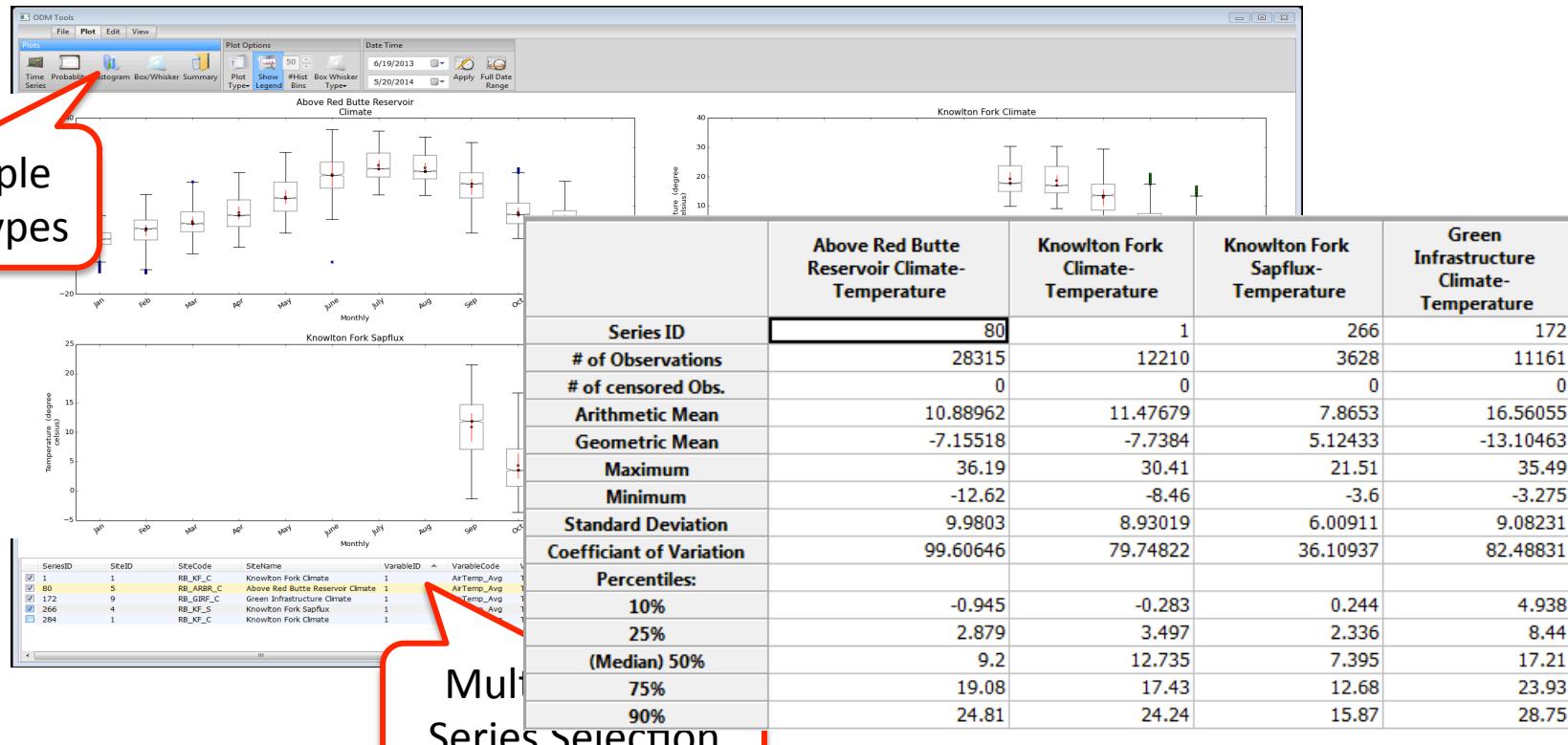
All Simple Filter Advanced Filter

Site RB_ARBR_AA-Above Red Butte Reservoir Advanced Aquatic
 Variable AirTemp_Avg-Temperature

SeriesID	SiteID	SiteCode	SiteName	VariableID	VariableCode	VariableName	Speciation	VariableUnitsID	VariableUnits...	SampleMedium	ValueType	TimeSupport	TimeUnitsID	TimeUnitsName	DataType	General
1	1	RB_KF_C	Knowlton Fork Climate	1	AirTemp_Avg	Temperature	Not Applicable	96	degree celsius	Air	Field Observation	15.0	102	minute	Average	Climate
80	5	RB_ARBR_C	Above Red Butte Reservoir Climate	1	AirTemp_Avg	Temperature	Not Applicable	96	degree celsius	Air	Field Observation	15.0	102	minute	Average	Climate
172	9	RB_GIRF_C	Green Infrastructure Climate	1	AirTemp_Avg	Temperature	Not Applicable	96	degree celsius	Air	Field Observation	15.0	102	minute	Average	Climate
266	4	RB_KF_S	Knowlton Fork Sapflux	1	AirTemp_Avg	Temperature	Not Applicable	96	degree celsius	Air	Field Observation	15.0	102	minute	Average	Climate
284	1	RB_KF_C	Knowlton Fork Climate	1	AirTemp_Avg	Temperature	Not Applicable	96	degree celsius	Air	Field Observation	15.0	102	minute	Average	Climate

Plotting: Multiple Types and Series

- Time Series
- Probability
- Histogram
- Box/Whisker
- Summary Statistics
- Multiple y-axes for different variables
- Colors automatically assigned
- Subplots
- Plotting options

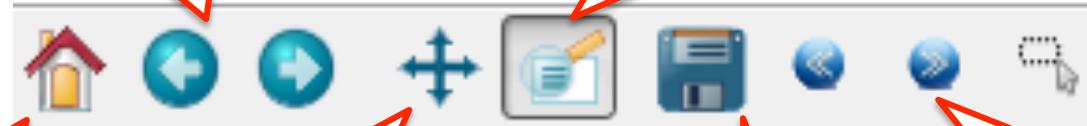


Plotting: Options

- **Time Series and Probability:** can show Lines, Points, or Both
- **Time Series and Probability:** Legend displays Site, Variable, and QCLevel
- **Legend** can be dragged and dropped
- **Histogram:** number of bins/bars can range from 0-99
- **Box/Whisker:** can group by month, season, or overall
- **Margins** can be adjusted for Histogram and Box/Whisker
- Can subset data series by restricting **Date Range**



Plotting: Dynamic Panning/Zooming



Home Button: Returns to the original, default plot extents.

Pan Button: Clicking and dragging on the plot will pan over the plot.

Zoom Button: Left click to draw a box for zooming. Right click to draw a box to zoom out.

Pan Arrows: Use to pan forward and backward in time.

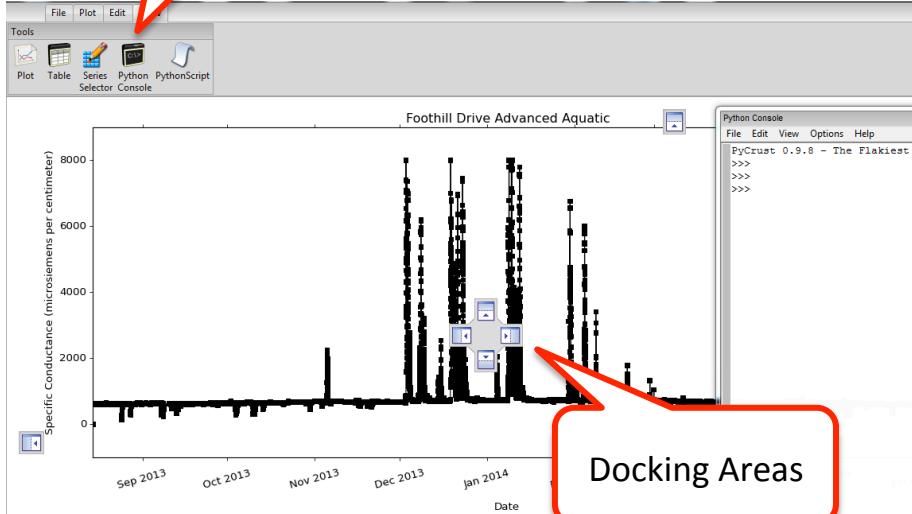
Save Button: save the current plot as an image.

Zoom Arrows: Revert to a previous or a subsequent view

Dockable Windows

Show/Hide
Windows

Python Console:
Used for custom
edits.



```
1 #from odservices import EditService
2 #from odservices import SeriesService
3 #edit_service = EditService(series_id=232, connection_string='mssql+pyodbc://Amber:xxxxxxxx@(local)/iUTAH')
4 #series_service = SeriesService(connection_string='mssql+pyodbc://Amber:b00ger@(local)/iUTAH_RedButte_OD')
5 ## To run commands from the python console uncomment and run the following command ##
6 #edit_service = Tools
7
```

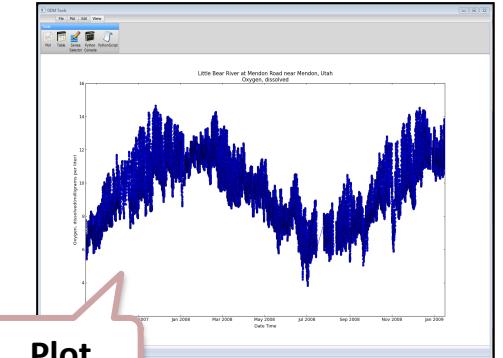
Tabbed
Docking

Python Script:
Shows steps in the
editing process

```
>>>
>>>
>>>
```

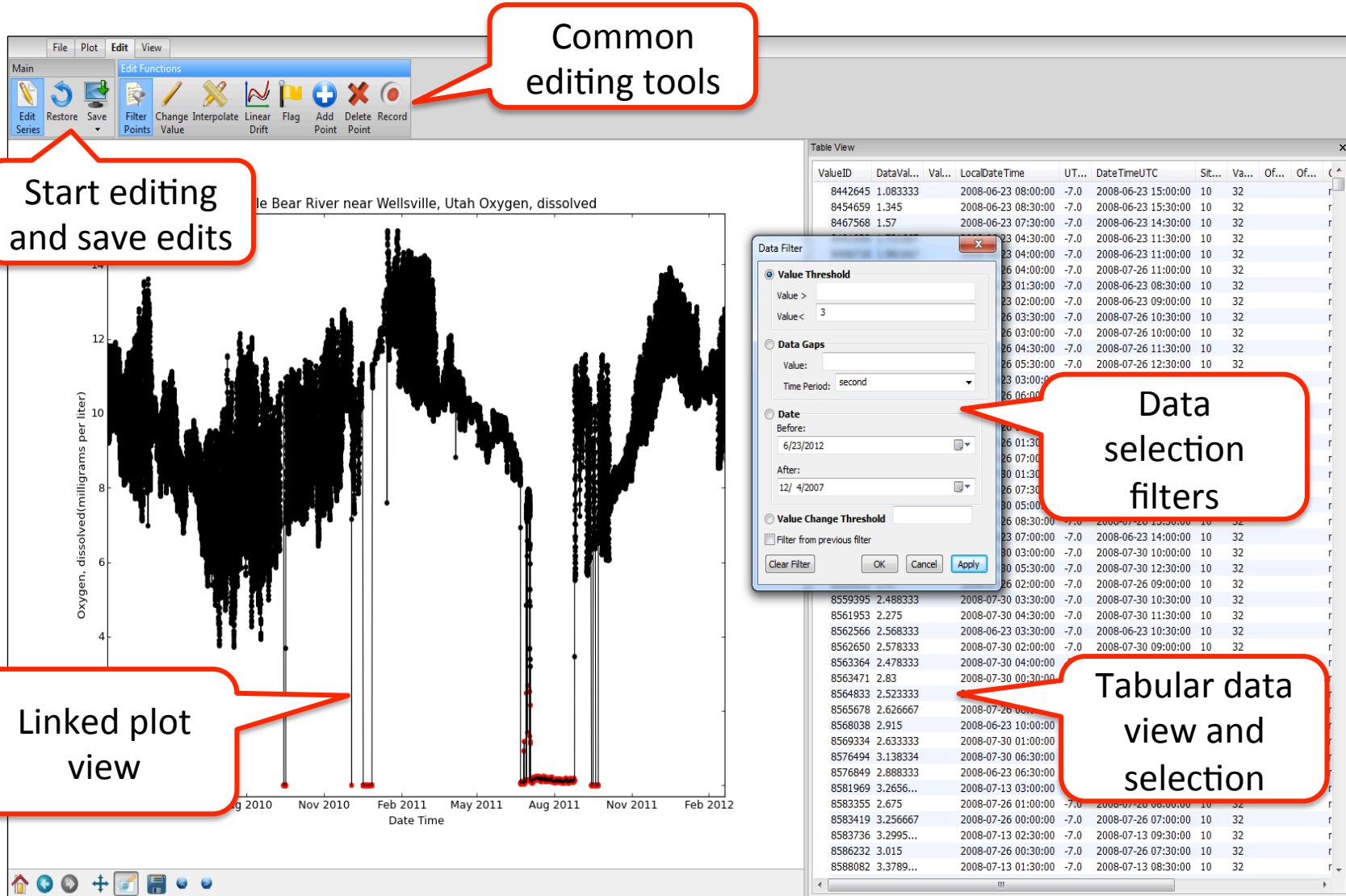
LocalDateTime
2013-08-14 12:15:00
2013-08-14 12:30:00
2013-08-14 12:45:00
2013-08-14 13:00:00
2013-08-14 13:15:00
2013-08-14 13:30:00
2013-08-14 13:45:00
2013-08-14 14:00:00
2013-08-14 14:15:00
2013-08-14 14:30:00
2013-08-14 14:45:00
2013-08-14 15:00:00
2013-08-14 15:15:00
2013-08-14 15:30:00
2013-08-14 15:45:00
2013-08-14 16:00:00
2013-08-14 16:15:00
2013-08-14 16:30:00
2013-08-14 16:45:00
2013-08-14 17:00:00
2013-08-14 17:15:00
2013-08-14 17:30:00
2013-08-14 17:45:00
2013-08-14 18:00:00
2013-08-14 18:15:00
2013-08-14 18:30:00
2013-08-14 18:45:00
2013-08-14 19:00:00
2013-08-14 19:15:00
2013-08-14 19:30:00
2013-08-14 19:45:00
2013-08-14 20:00:00
2013-08-14 20:15:00
2013-08-14 20:30:00
2013-08-14 20:45:00
2013-08-14 21:00
2013-08-14 21:15
2013-08-14 21:30
2013-08-14 21:45
2013-08-14 22:00
2013-08-14 22:15
2013-08-14 22:30
2013-08-14 22:45
2013-08-14 23:00
2013-08-14 23:15
2013-08-14 23:30
2013-08-14 23:45
2013-08-14 23:59
2013-08-15 00:00

```
PyCrust 0.9.8 - The Flakiest Python Shell
>>>
>>>
>>>
```

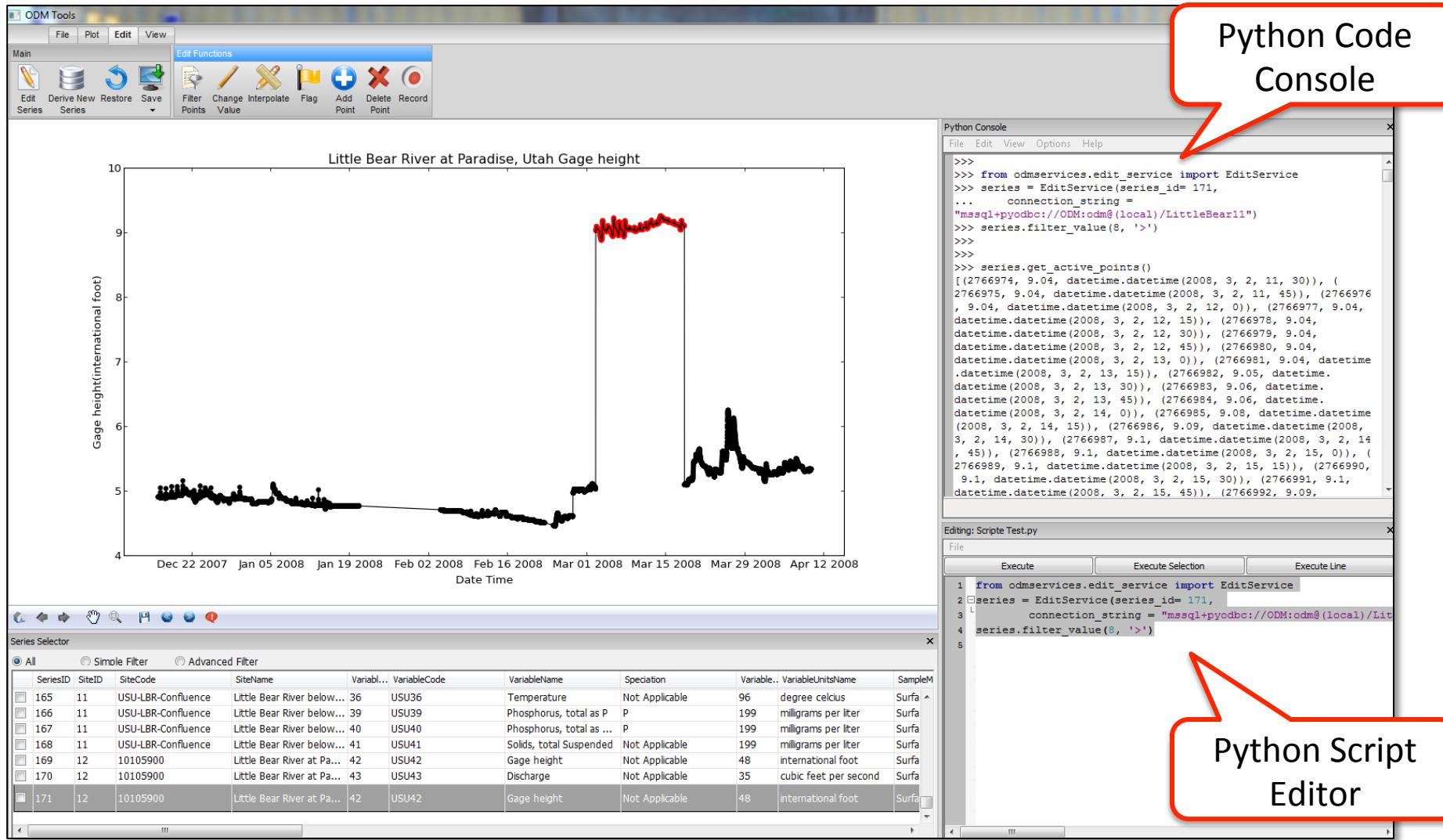


Value	Unit	Variable
1413905	611.3	
1413906	610.8	
1413907	610.8	
1413908	610.8	
1413909	610.8	
1413910	610.8	
1413911	610.8	
1413912	610.8	
1413913	610.8	
1413914	610.8	
1413915	610.8	
1413916	610.8	
1413917	610.8	
1413918	610.8	
1413919	610.8	
1413920	610.8	
1413921	610.8	
1413922	610.8	
1413923	610.8	
1413924	610.8	
1413925	610.8	
1413926	610.8	
1413927	610.8	
1413928	610.8	
1413929	610.8	
1413930	610.8	
1413931	610.8	
1413932	610.8	
1413933	610.8	
1413934	610.8	
1413935	610.8	
1413936	610.8	
1413937	610.8	
1413938	610.8	
1413939	610.8	
1413940	610.8	
1413941	610.8	
1413942	610.8	
1413943	610.8	
1413944	610.8	
1413945	610.8	
1413946	610.8	
1413947	610.8	
1413948	610.8	
1413949	610.8	
1413950	610.8	
1413951	610.8	
1413952	610.8	
1413953	610.8	
1413954	610.8	
1413955	610.8	
1413956	610.8	
1413957	610.8	
1413958	610.8	
1413959	610.8	
1413960	610.8	
1413961	610.8	
1413962	610.8	
1413963	610.8	
1413964	610.8	
1413965	610.8	
1413966	610.8	
1413967	610.8	
1413968	610.8	
1413969	610.8	
1413970	610.8	
1413971	610.8	
1413972	610.8	
1413973	610.8	
1413974	610.8	
1413975	610.8	
1413976	610.8	
1413977	610.8	
1413978	610.8	
1413979	610.8	
1413980	610.8	
1413981	610.8	
1413982	610.8	
1413983	610.8	
1413984	610.8	
1413985	610.8	
1413986	610.8	
1413987	610.8	
1413988	610.8	
1413989	610.8	
1413990	610.8	
1413991	610.8	
1413992	610.8	
1413993	610.8	
1413994	610.8	
1413995	610.8	
1413996	610.8	
1413997	610.8	
1413998	610.8	
1413999	610.8	
1413900	610.8	
1413901	610.8	
1413902	610.8	
1413903	610.8	
1413904	610.8	
1413905	610.8	
1413906	610.8	
1413907	610.8	
1413908	610.8	
1413909	610.8	
1413910	610.8	
1413911	610.8	
1413912	610.8	
1413913	610.8	
1413914	610.8	
1413915	610.8	
1413916	610.8	
1413917	610.8	
1413918	610.8	
1413919	610.8	
1413920	610.8	
1413921	610.8	
1413922	610.8	
1413923	610.8	
1413924	610.8	
1413925	610.8	
1413926	610.8	
1413927	610.8	
1413928	610.8	
1413929	610.8	
1413930	610.8	
1413931	610.8	
1413932	610.8	
1413933	610.8	
1413934	610.8	
1413935	610.8	
1413936	610.8	
1413937	610.8	
1413938	610.8	
1413939	610.8	
1413940	610.8	
1413941	610.8	
1413942	610.8	
1413943	610.8	
1413944	610.8	
1413945	610.8	
1413946	610.8	
1413947	610.8	
1413948	610.8	
1413949	610.8	
1413950	610.8	
1413951	610.8	
1413952	610.8	
1413953	610.8	
1413954	610.8	
1413955	610.8	
1413956	610.8	
1413957	610.8	
1413958	610.8	
1413959	610.8	
1413960	610.8	
1413961	610.8	
1413962	610.8	
1413963	610.8	
1413964	610.8	
1413965	610.8	
1413966	610.8	
1413967	610.8	
1413968	610.8	
1413969	610.8	
1413970	610.8	
1413971	610.8	
1413972	610.8	
1413973	610.8	
1413974	610.8	
1413975	610.8	
1413976	610.8	
1413977	610.8	
1413978	610.8	
1413979	610.8	
1413980	610.8	
1413981	610.8	
1413982	610.8	
1413983	610.8	
1413984	610.8	
1413985	610.8	
1413986	610.8	
1413987	610.8	
1413988	610.8	
1413989	610.8	
1413990	610.8	
1413991	610.8	
1413992	610.8	
1413993	610.8	
1413994	610.8	
1413995	610.8	
1413996	610.8	
1413997	610.8	
1413998	610.8	
1413999	610.8	
1414000	610.8	
1414001	610.8	
1414002	610.8	
1414003	610.8	
1414004	610.8	
1414005	610.8	
1414006	610.8	
1414007	610.8	
1414008	610.8	
1414009	610.8	
1414010	610.8	
1414011	610.8	
1414012	610.8	
1414013	610.8	
1414014	610.8	
1414015	610.8	
1414016	610.8	
1414017	610.8	
1414018	610.8	
1414019	610.8	
1414020	610.8	
1414021	610.8	
1414022	610.8	
1414023	610.8	
1414024	610.8	
1414025	610.8	
1414026	610.8	
1414027	610.8	
1414028	610.8	
1414029	610.8	
1414030	610.8	
1414031	610.8	
1414032	610.8	
1414033	610.8	
1414034	610.8	
1414035	610.8	
1414036	610.8	
1414037	610.8	
1414038	610.8	
1414039	610.8	
1414040	610.8	
1414041	610.8	
1414042	610.8	
1414043	610.8	
1414044	610.8	
1414045	610.8	
1414046	610.8	
1414047	610.8	
1414048	610.8	
1414049	610.8	
1414050	610.8	
1414051	610.8	
1414052	610.8	
1414053	610.8	
1414054	610.8	
1414055	610.8	
1414056	610.8	
1414057	610.8	
1414058	610.8	
1414059	610.8	
1414060	610.8	
1414061	610.8	
1414062	610.8	
1414063	610.8	
1414064	610.8	
1414065	610.8	
1414066	610.8	
1414067	610.8	
1414068	610.8	
1414069	610.8	
1414070	610.8	
1414071	610.8	
1414072	610.8	
1414073	610.8	
1414074	610.8	
1414075	610.8	
1414076	610.8	
1414077	610.8	
1414078	610.8	
1414079	610.8	
1414080	610.8	
1414081	610.8	
1414082	610.8	
1414083	610.8	
1414084	610.8	
1414085	610.8	
1414086	610.8	
1414087	610.8	
1414088	610.8	
1414089	610.8	
1414090	610.8	
1414091	610.8	
1414092	610.8	
1414093	610.8	
1414094	610.8	
1414095	610.8	
1414096	610.8	
1414097	610.8	
1414098	610.8	
1414099	610.8	
1414100	610.8	
1414101	610.8	
1414102	610.8	
1414103	610.8	
1414104	610.8	
1414105	610.8	
1414106	610.8	
1414107	610.8	
1414108	610.8	
1414109	610.8	
1414110	610.8	
1414111	610.8	
1414112	610.8	
1414113	610.8	
1414114	610.8	
1414115	610.8	
1414116	610.8	
1414117	610.8	
1414118	610.8	
1414119	610.8	
1414120	610.8	
1414121	610.8	
1414122	610.8	
1414123	610.8	
1414124	610.8	

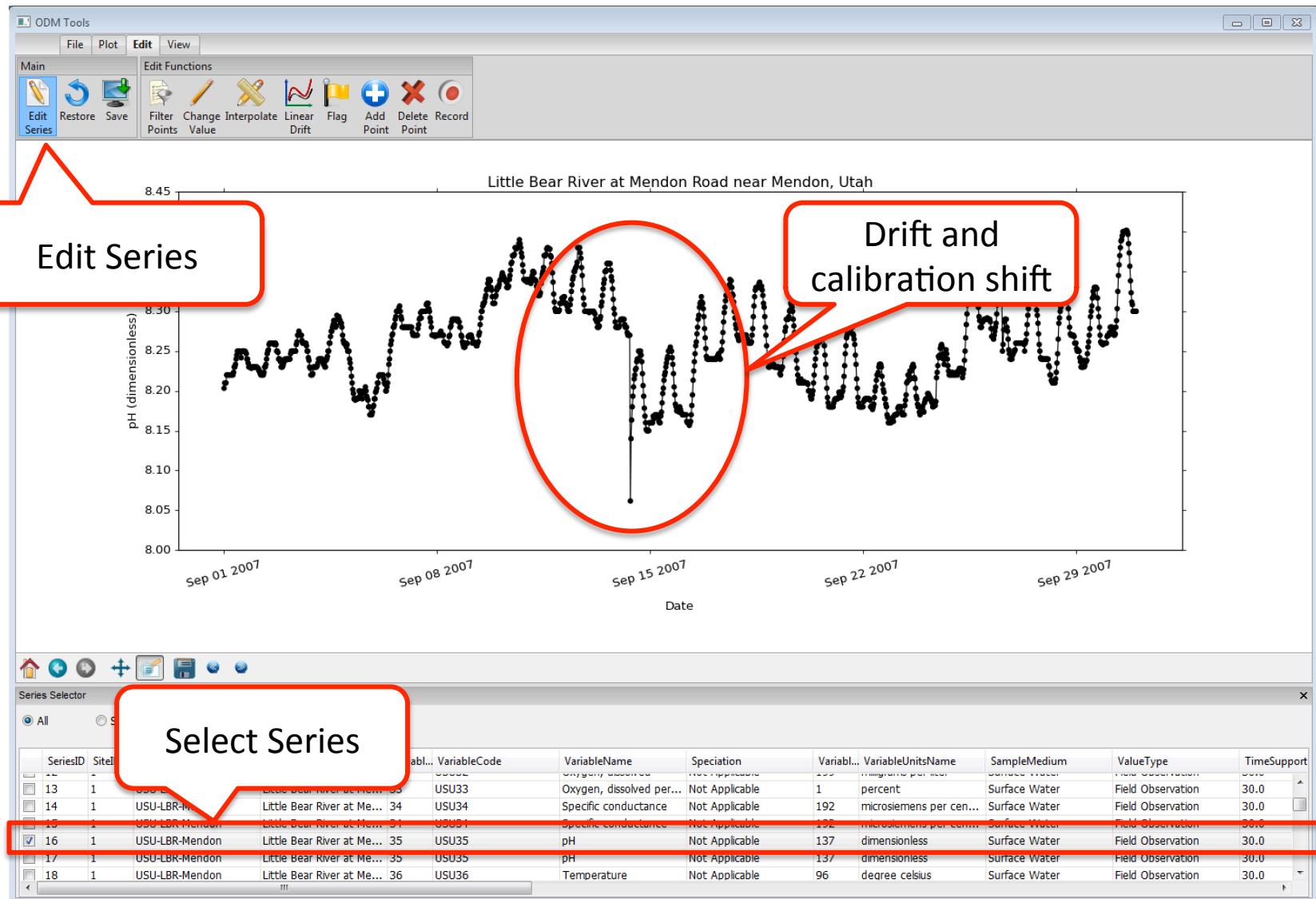
Editing for Quality Control



Data Editing for Quality Control

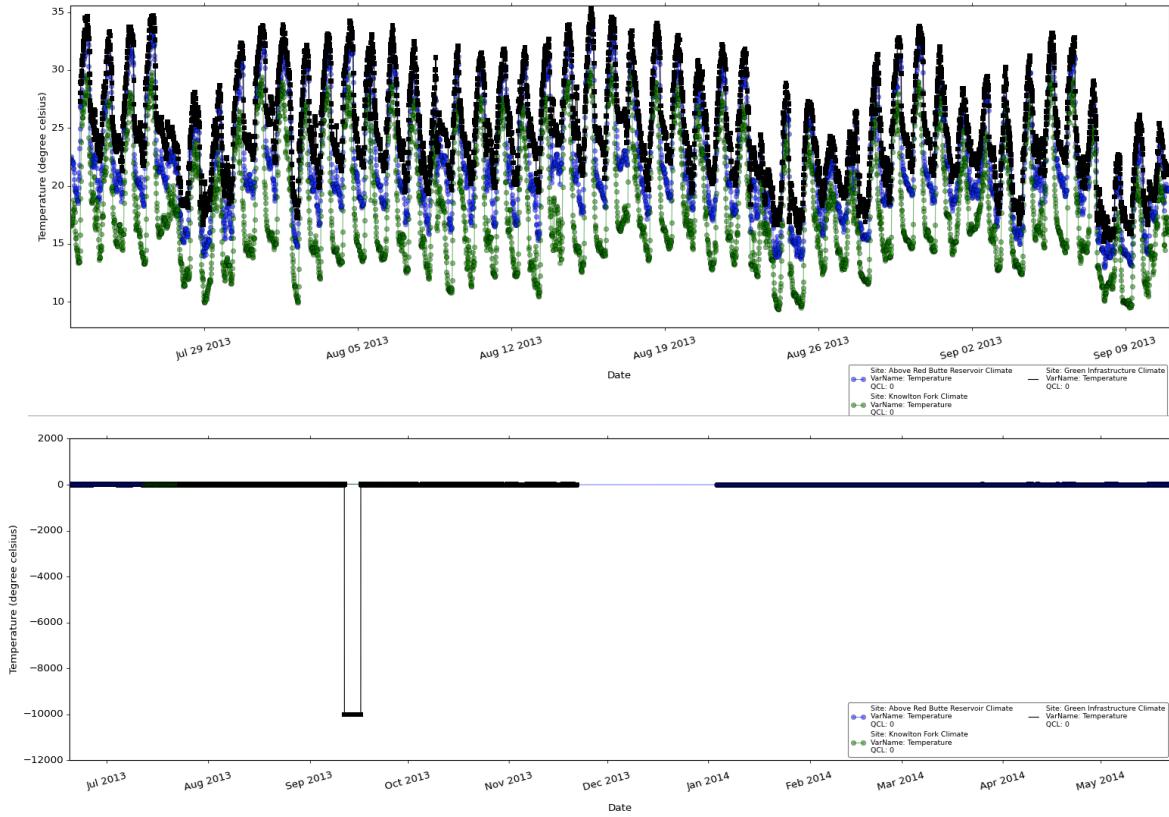


Step 1: Select a Time Series for Editing

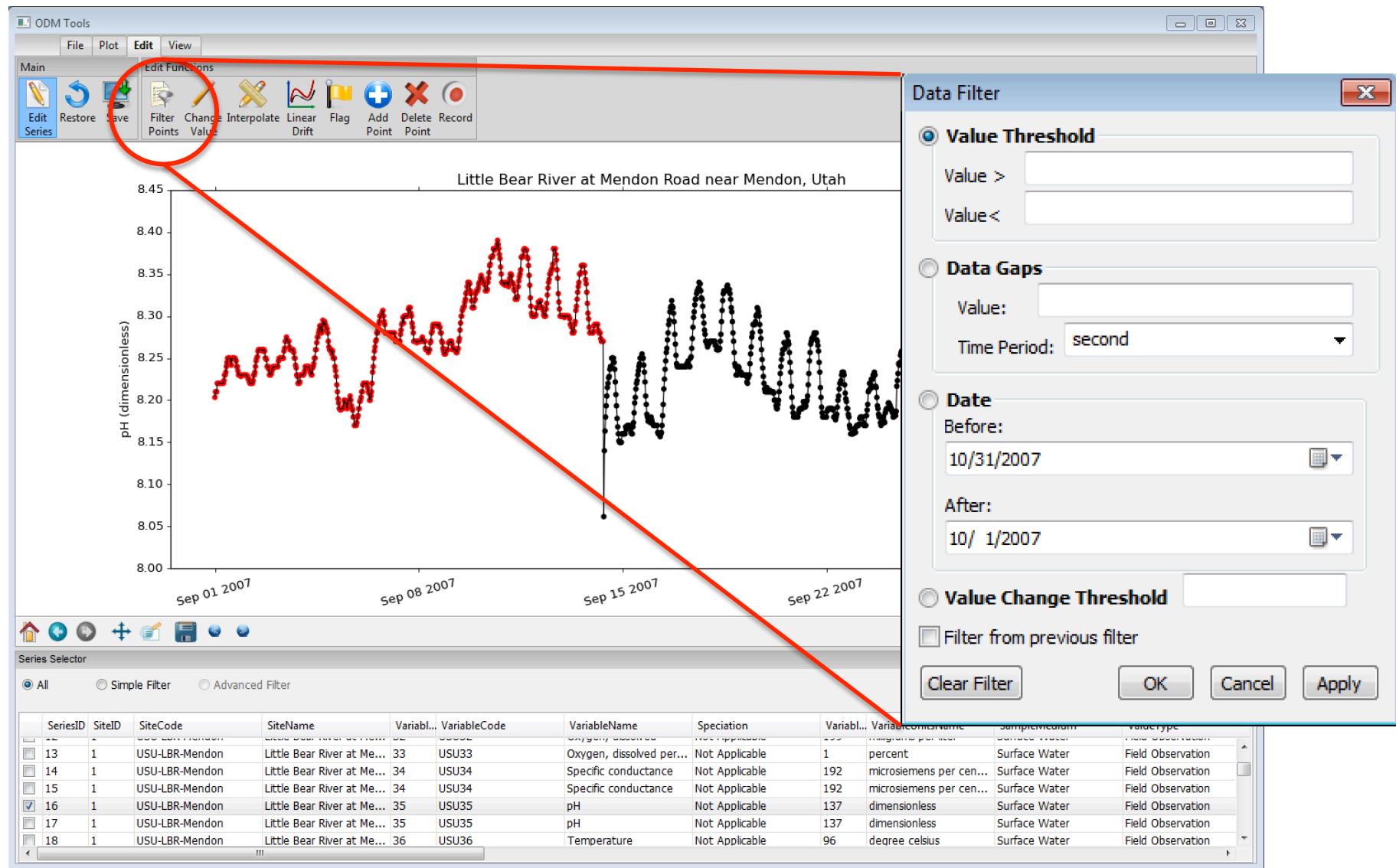


Editing: Selecting Series

- Series Selector still used to show/hide series in the plot
- Multiple data series can be viewed while editing- symbology changes
- ‘No Data’ (-9999) values are now plotted- zoom level changes.
- When a data series is being edited, points are loaded to the Table window.

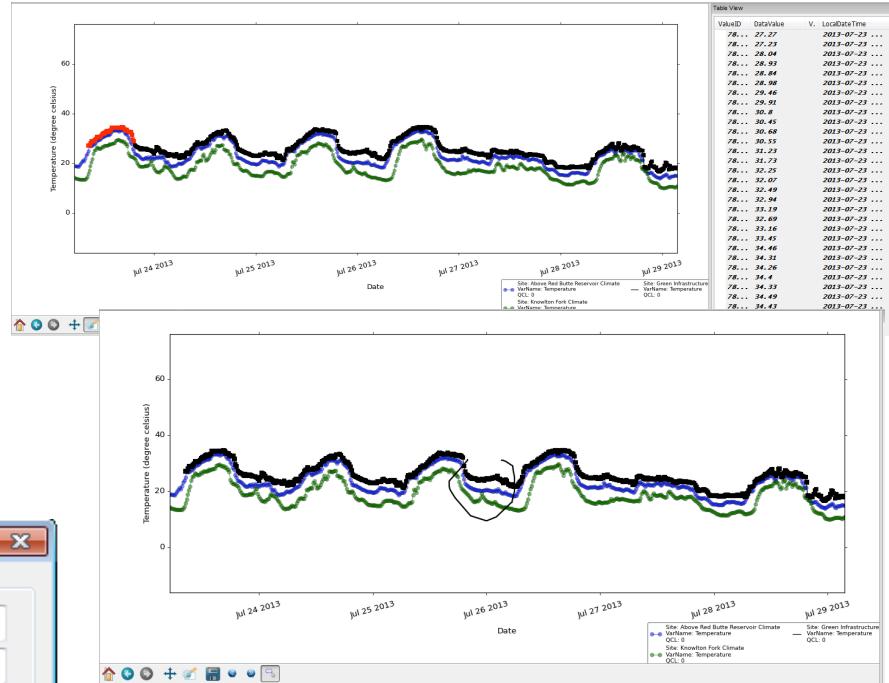


Step 2: Select Data to Edit

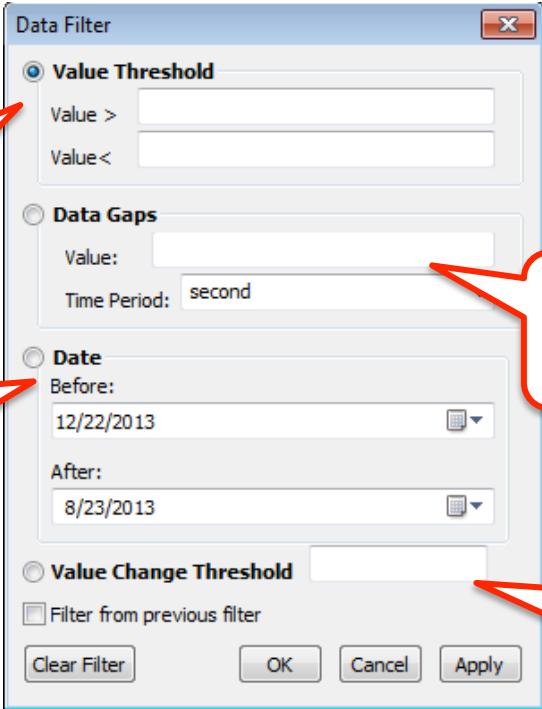


Editing: Selecting Points

- **Table Window:** use ctrl+click, shift+click, click and drag
- **Plot Window:** use Lasso to draw shape around points
- **Filters:** select based on criteria
- **De-select:** select other points, use 'Clear Filter' or 'Reset Selection'.



Value Threshold:
selects data points $>$, $<$, or $><$ specified criteria.

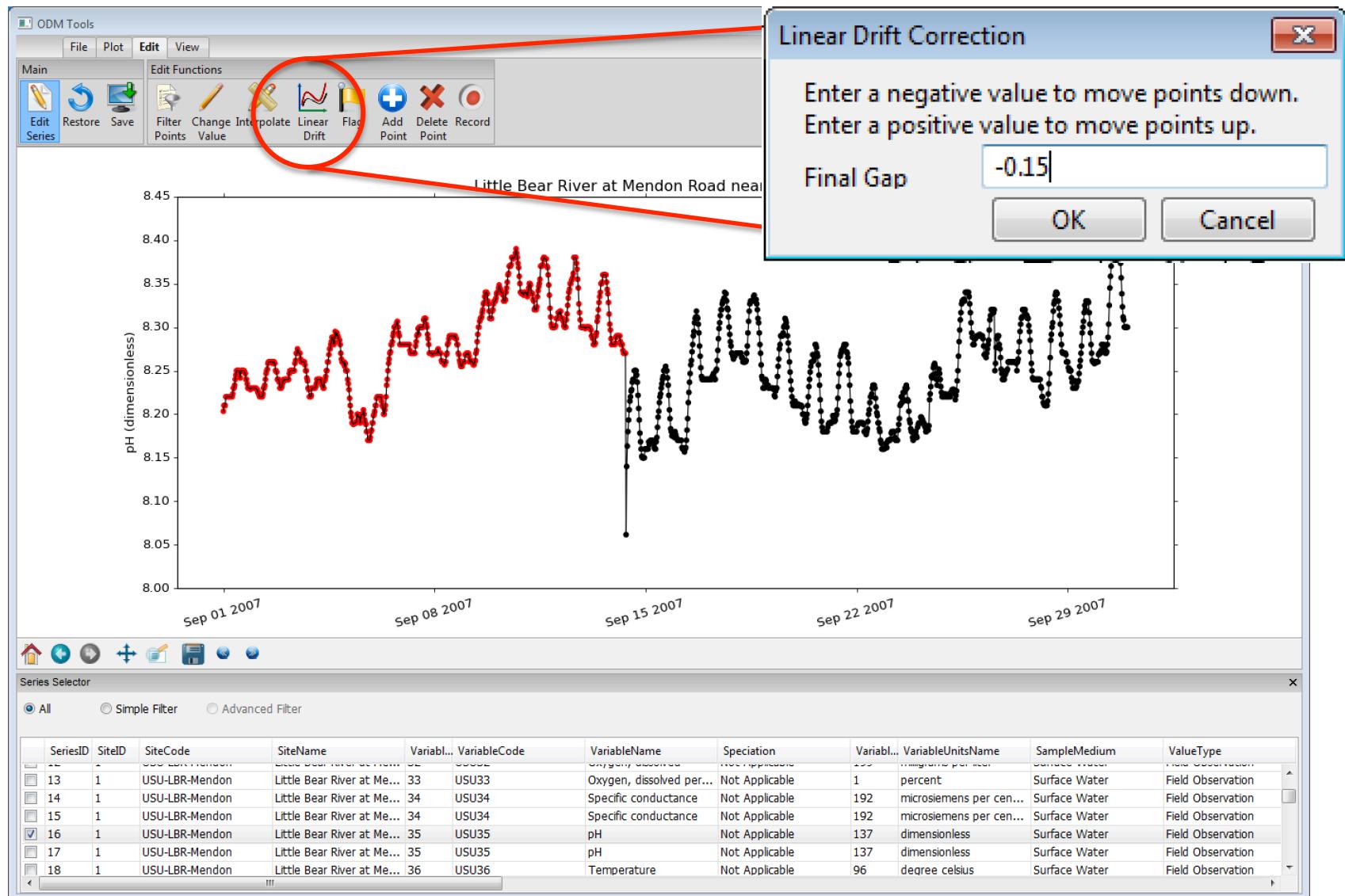


Data Gaps: selects beginning and end points of a specified gap in continuous data.

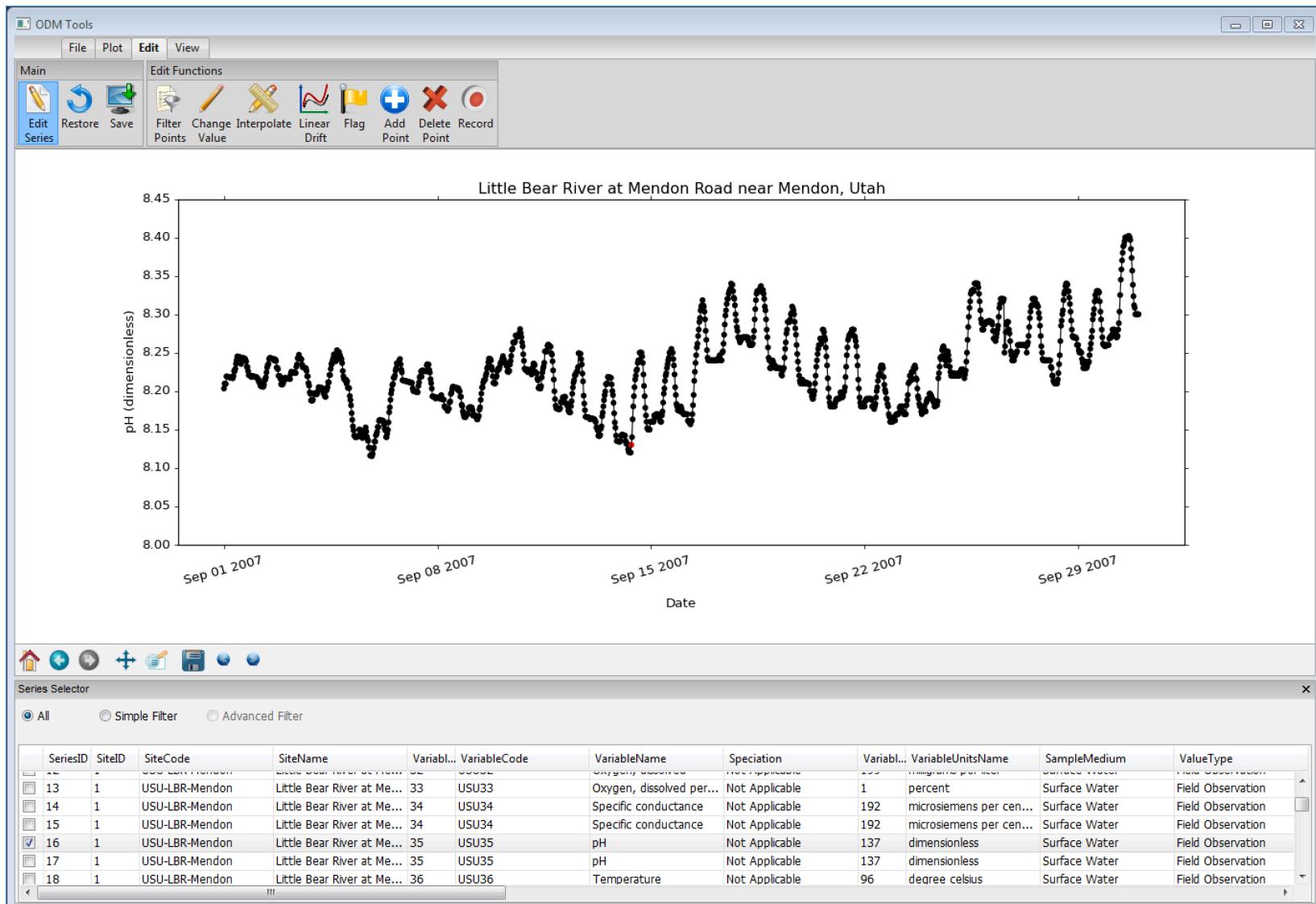
Date Range: selects before, after, between, or outside specified dates (time not yet implemented).

Value Change Threshold:
selects subsequent pairs of points with a value difference $>$ specified criterion.

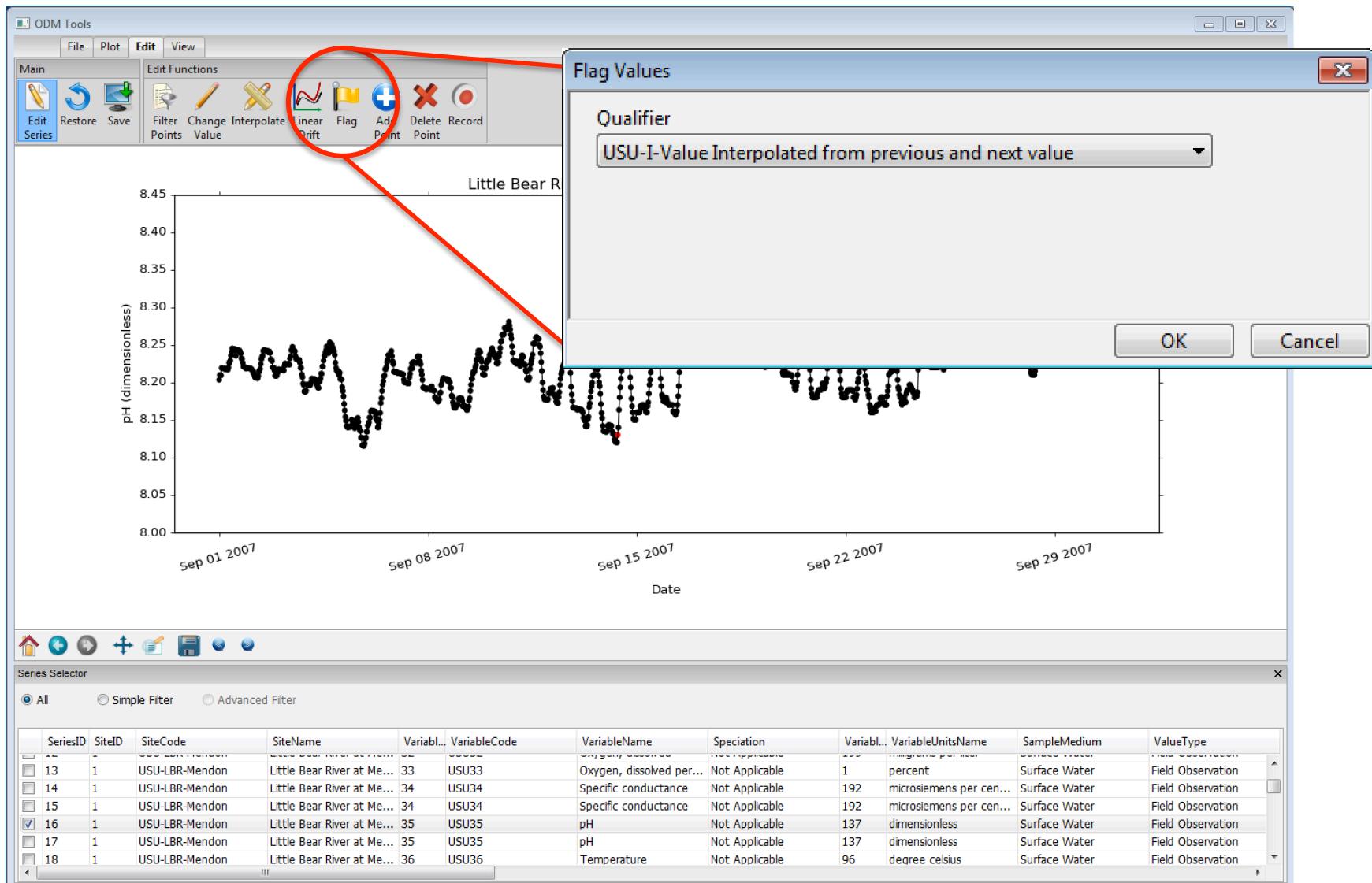
Step 3: Linear Drift Correction



Step 4: Interpolate



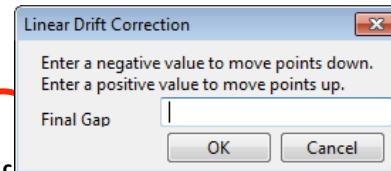
Step 5: Flag



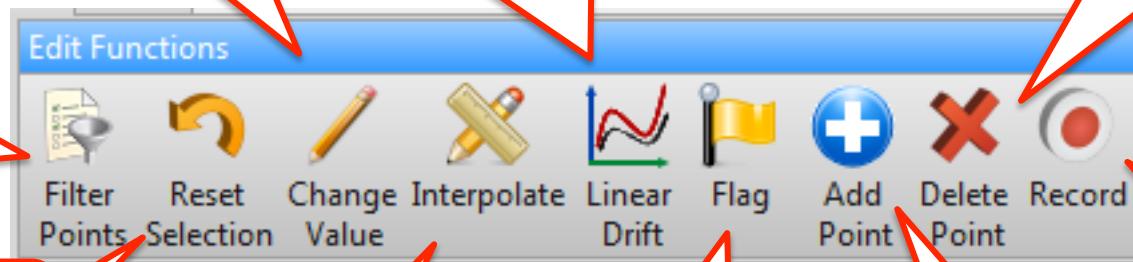
Editing Functions



Change Value: Add, subtract, multiply or set equal to a constant value.



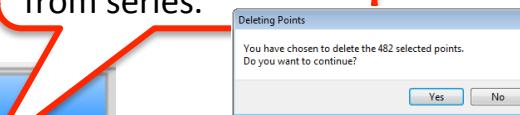
Linear Drift: Uses reference values at beginning and end of selected data and a gap value to account for calibration shifts and sensor drift.



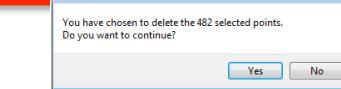
Filter Points: Use filters to select points

Reset Selection: Remove selection from all points.

Interpolate: Linearly interpolates points based on previous and subsequent values.

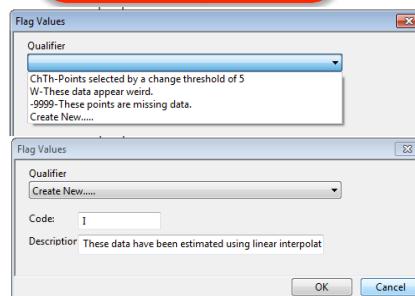


Delete: Removes selected data values from series.

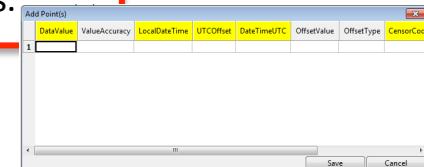


Record: Initiates recording of edits to Python script.

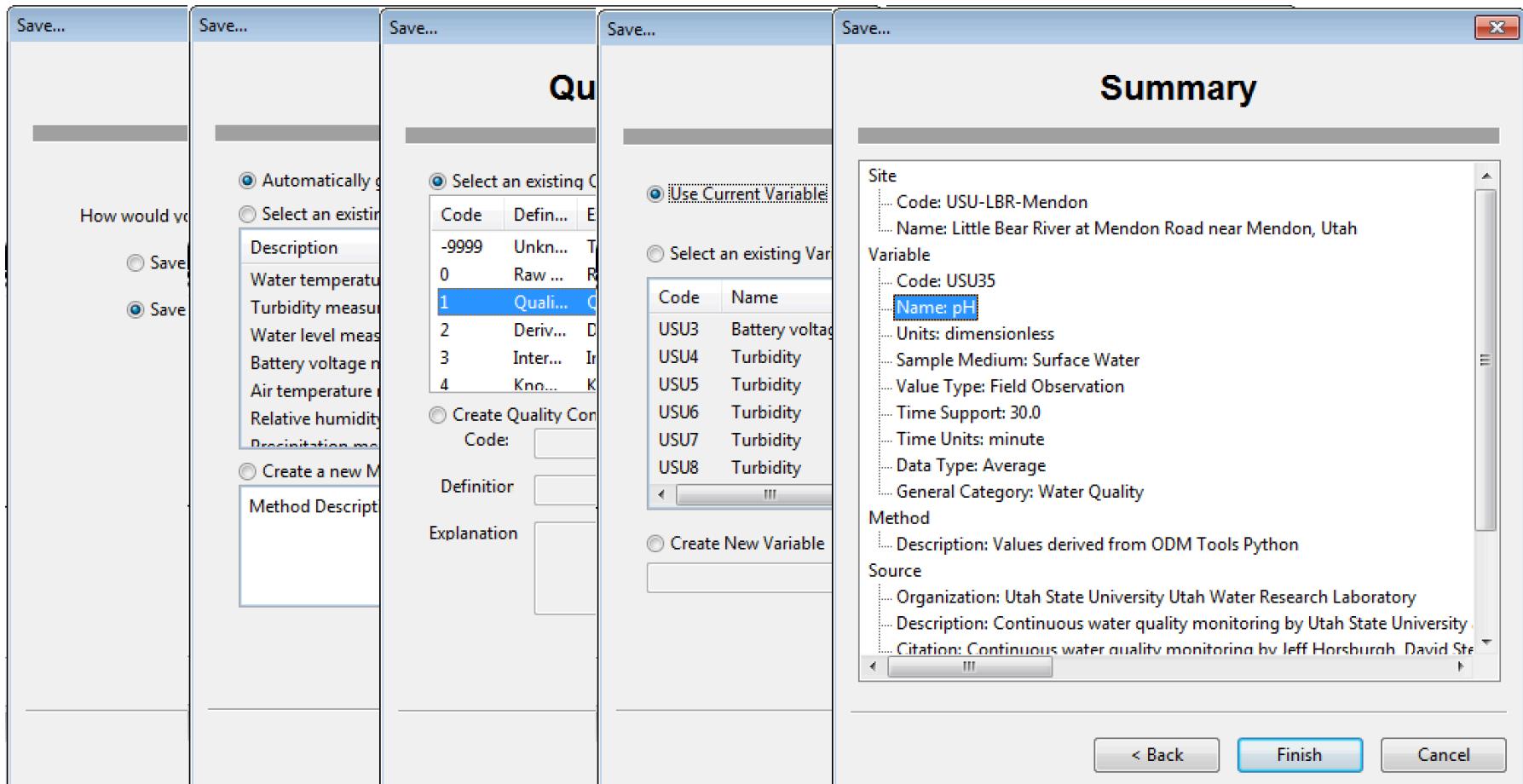
Flag: Applies a data qualifying comment to selected points.



Add Point: Inserts individual data values.

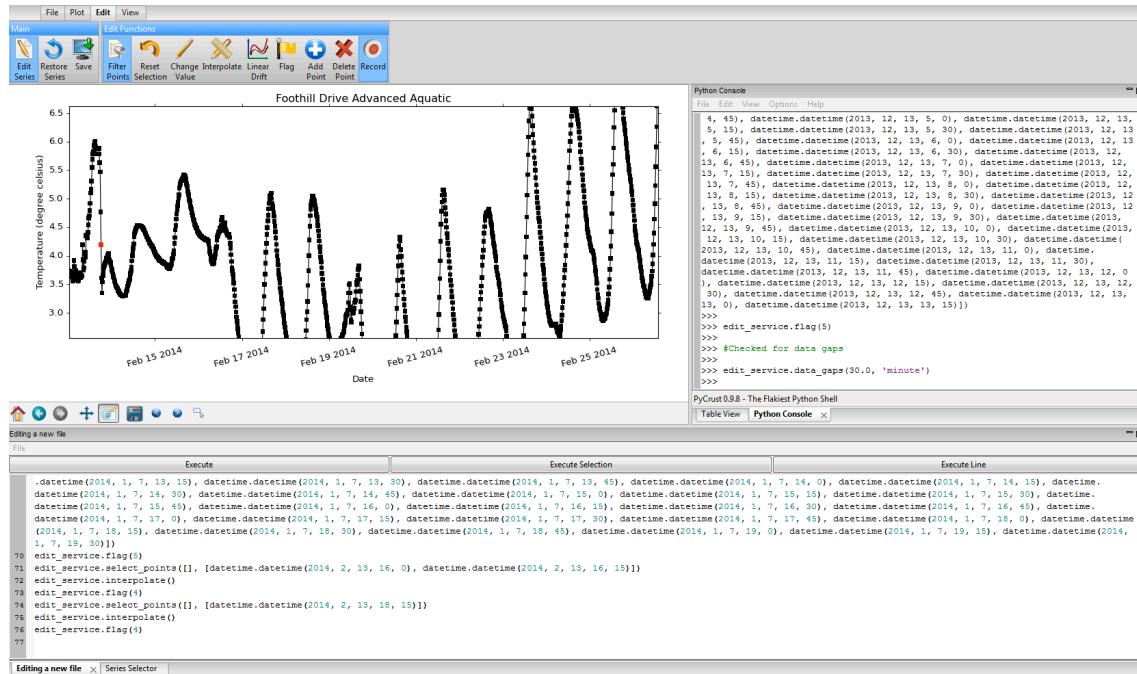


Step 6: Save Modified Data Series



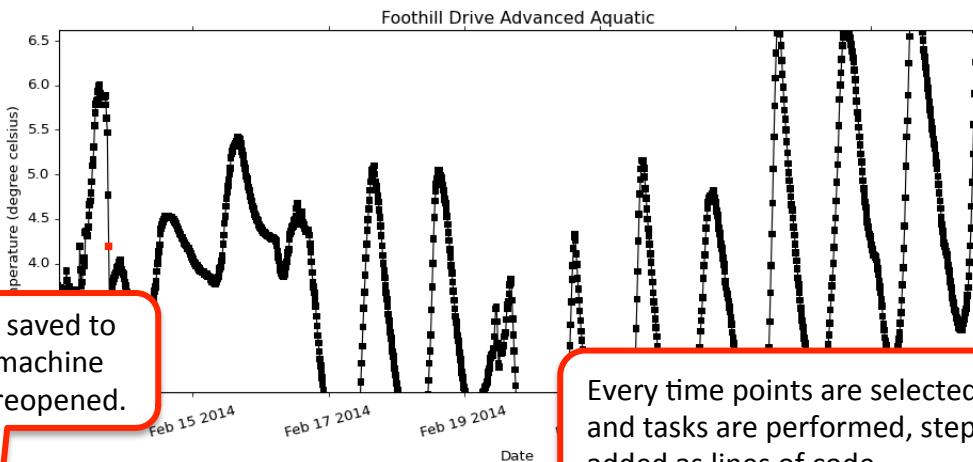
Scripting Edits

- All actions recorded as lines of code in the Python script editor
 - Script can then be saved for later execution and review.
 - Fosters traceability, transparency, reproducibility
 - Can add comments to the script to annotate the rationale for edits
 - Editing can be performed in multiple sessions with the script incrementally saved and re-run for each new session
 - Script could be generated by one user and then reviewed and revised by another



Scripting Edits

Click on the 'Record' button in the Ribbon. View the Script Editor window.



Script saved to local machine then reopened.

Every time points are selected and tasks are performed, steps added as lines of code.

A screenshot showing the Python Console and the Script Editor window. The Python Console window shows a series of command-line entries in black text on a white background. The Script Editor window shows a script with numbered lines of code. Lines 31 through 43 are visible, containing commands like 'edit_service.flag(5)', '#Checked for data gaps', 'edit_service.data_gaps(30.0, 'minute')', and 'edit_service.select_points([])'. A red arrow points from the text 'Can edit script. Comments added using leading #' to the first few lines of the script editor.

Can open previous file in script editor and run (sent to Python console).

Currently, cannot append file (every time 'Record' is used, a new script is started), so should copy/paste saved script to new file for subsequent sessions

Workflow

