ODM Tools
Version 1.2

An application for interfacing with the CUAHSI Hydrologic Information System Observations Data Model

October 2007

Prepared by:

Jeffery S. Horsburgh
Environmental Management Research Group
Utah Water Research Laboratory
Utah State University

Distribution

The ODM Tools application and all associated source code and documentation are available at the following URL: http://water.usu.edu/cuahsi/ODM/

The ODM Tools software, source code, and documentation are provide free of charge under the Berkely Software Distribution (BSD) license. Please see the following license information:

Disclaimers

Although much effort has been expended in the development and testing of the ODM Tools application, errors and inadequacies may still occur. Users must make the final evaluation as to the usefulness of ODM Tools for his or her application.

The ODM Tools application and this software manual are based upon work supported by the National Science Foundation under Grant No. 03-26064. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Acknowledgements

The team of engineers and software developers that developed the ODM Tools application includes:

Jeffery S. Horsburgh, Research Engineer, Environmental Management Research Group, Utah Water Research Laboratory, Utah State University, Logan, UT.
David K. Stevens, Professor of Civil and Environmental Engineering, Utah State University, Logan, UT.
Michelle Hospodarsky, Programmer, Environmental Management Research Group, Utah Water Research Laboratory, Utah State University, Logan, UT
Justin Berger, Programmer, Environmental Management Research Group, Utah Water Research Laboratory, Utah State University, Logan, UT

Funding for the ODM Tools application was provided by the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) under NSF Grant 03-26064. In addition, much input and feedback has been received from Rick Hooper, President and Executive Director of CUAHSI, and the CUAHSI Hydrologic Information System development team. Their contribution is acknowledged here.

Parts of the ODM Tools application were modeled after programs developed by United States Geological Survey (USGS) personnel for the Panola Mountain Experimental Watershed. Brent Aulenbach of USGS is acknowledged for his contribution to this work.

Technical Support

There is no formal ongoing support for this freely distributed open source software. However, we are interested in feedback. If you find errors, have suggestions, or are interested in any later versions, please contact:

Jeffery S. Horsburgh
Utah State University
8200 Old Main Hill
Logan, UT 84322-8200
jeff.horsburgh@usu.edu

Table of Contents

1.0 Introduction and Software Description

The CUAHSI Hydrologic Information System (HIS) Project is developing information technology infrastructure to support hydrologic science. One of the components of the HIS is a point Observations Data Model (ODM) (Tarboton et al., 2007), which is a relational database schema that was designed for storing time series data. The purpose of the ODM is to provide a framework for optimizing data storage and retrieval for integrated analysis of information collected by multiple investigators. It is anticipated that the CUAHSI HIS ODM will be implemented by a number of local work groups throughout the country and that these work groups will use the ODM as a mechanism for publication of individual investigator data and for registering these data with the National HIS.

Under this premise, the ODM Tools application was created to allow administrators of local instances of the ODM to visualize, manage, manipulate, edit, and export data that have been imported to their local instance of the ODM. The development of the ODM Tools application has several advantages. First, ODM Tools protects the security and consistency of a work group HIS ODM database because it provides users with a set of automated tools for performing many of the most common database transactions. Second, ODM Tools allows users to export data from their ODM instance with an accompanying metadata file. This allows users to work with local copies of data series exported from their ODM database while preserving the provenance of the data via the metadata file. ODM Tools also provides a mechanism by which users can interact with the ODM database without having to learn the complexities of its relational structure. Finally, for more advanced users, the source code of the ODM Tools application provides an example of how applications can be built on top of the CUAHIS HIS ODM.

1.1 General Functionality

The main objective of the ODM Tools application is to provide managers and users of work group instances of the ODM with a set of value added tools that they can use to better manage their data. These tools are organized into three general areas: 1) query and export; 2) visualize; and 3) edit. The Query and export functionality allows users to find the data that they are interested in and export it to a simple format that can be used with a variety of analysis software. The Visualize functionality allows users to quickly plot and summarize data using a variety of plot types and descriptive statistics. The Edit capability of ODM Tools was designed to provide users with a simple set of tools that they can use to edit existing data series and to create new data series from existing data series.

1.2 Platform and Minimum System Requirements

ODM Tools was designed to run on Microsoft Windows XP or Windows 2003 Server based computers. It is recommended that machines running the ODM Tools software have at least 100 MB of free disk space and 1 gigabyte of RAM. In addition, computers running the ODM Tools application must have the Microsoft .Net Framework Version 2.0 installed prior to installing ODM Tools. Instructions for obtaining the .Net Framework Version 2.0 from Microsoft are included in the Installation Instructions section below.

ODM Tools is a client application. It must be connected to an instance of the CUAHSI HIS ODM Version 1.0 that has been implemented in Microsoft SQL Server 2005. The SQL Server database can be located on the same machine as the ODM Tools application, or ODM Tools has the capability to connect to a remote ODM database provided that the database server name and ODM database name are known and the user has been given access to and authentication information for that server and database.

2.0 Installation Information

2.1 Installation Prerequisites

Prior to running the ODM Tools installation, you must first install the Microsoft .Net Framework Version 2.0. The .Net Framework Version 2.0 is free, and is required to run software applications developed in Microsoft's Visual Studio .Net 2005. Instructions for downloading and installing the .Net Framework Version 2.0 can be obtained from the Microsoft website via the following URL:

http://www.microsoft.com/downloads/details.aspx?FamilyID=0856eacb-4362-4b0d-8edd-aab15c5e04f5&displaylang=en

Once the .Net Framework Version 2.0 has been installed, you can continue with the ODM Tools installation.

NOTE: ODM Tools requires that you have an ODM database implemented in Microsoft SQL Server 2005. If you do not already have an instance of Microsoft SQL Server running, you can download and install Microsoft SQL Server 2005 Express from Microsoft for free. It is recommended that you download and install both SQL Server 2005 Express and SQL Server Management Studio Express. You can get these products and instructions for installing them at the following Microsoft URL:

http://www.microsoft.com/sql/editions/express/default.mspx

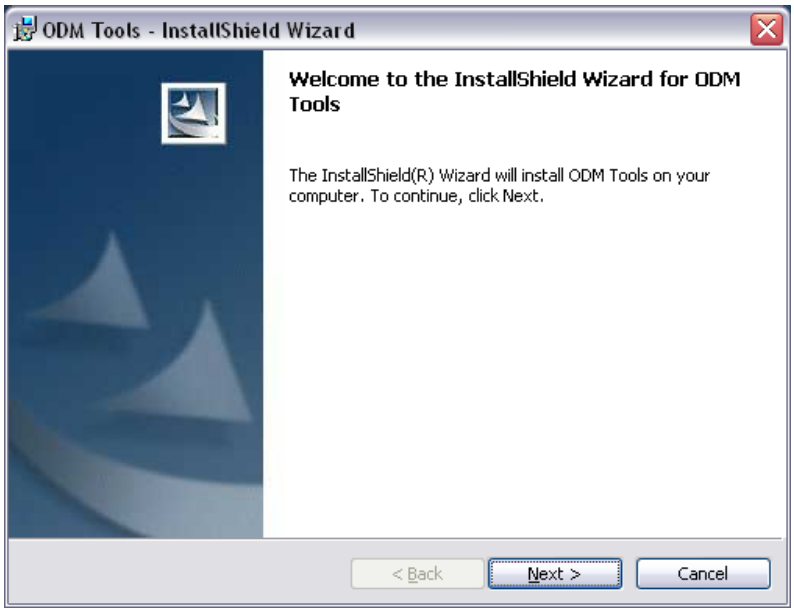
Directions for attaching the test databases distributed with the ODM Tools software installation to your instance of Microsoft SQL Server 2005 are included in Appendix A of this document.

2.2 Installing the ODM Tools Application

Install the ODM Tools using the following steps:

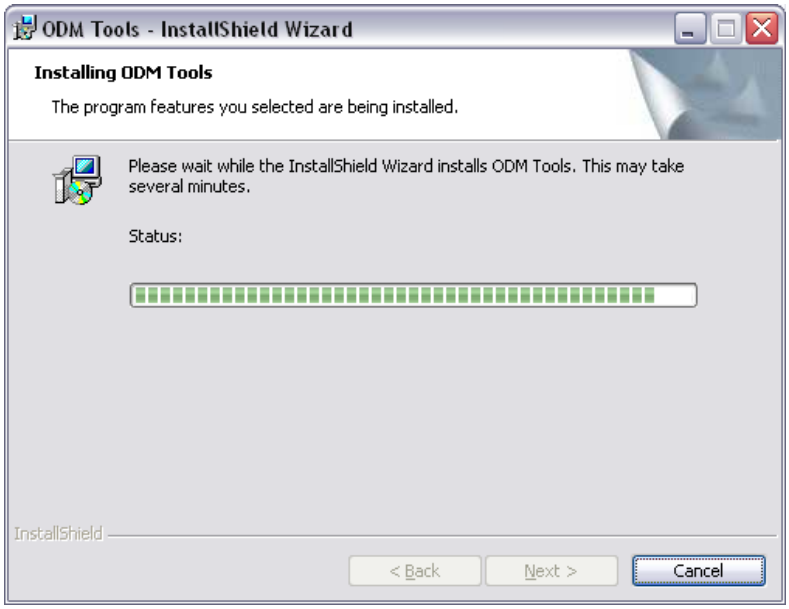
1. First, ensure that you have installed the Microsoft .Net Framework Version 2.0. See the previous section if you have not done so.

2. Double click on the setup.exe installation file. This will begin the installation of the ODM Tools application. After a few moments, the following window will appear:

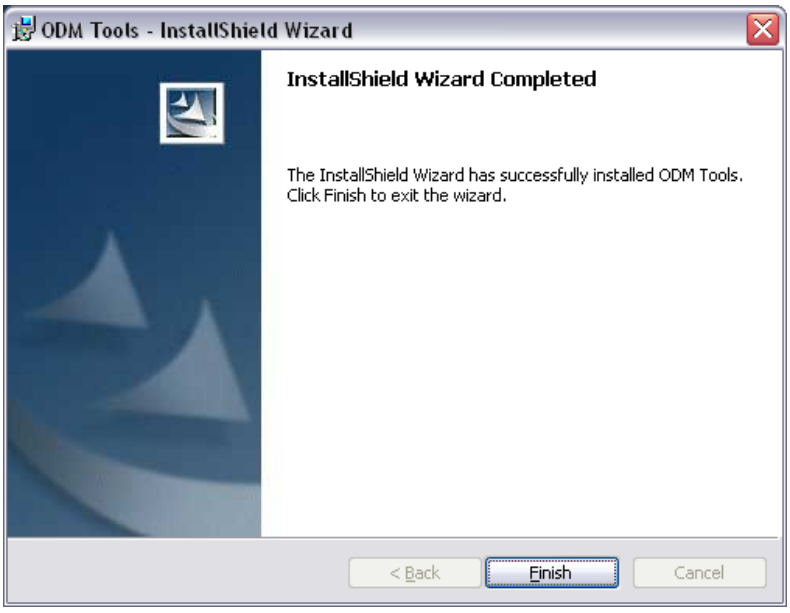


3. Click the "Next" button to continue with the ODM Tools installation. You will see the following window.

4. Read the license and then click on the radio button next to "I accept the terms in the license agreement" to accept the license. Click the "Next" button. The following window will appear for a few moments.



5. When the installation is complete, you will see the following window. Click the "Finish" button to complete the ODM Tools installation.

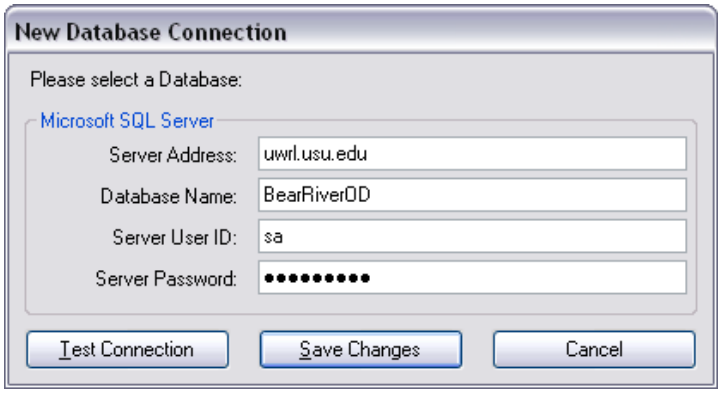


3.0 Connecting the ODM Tools Application to an Instance of ODM

NOTE: The steps in this section assume that you already have an ODM database set up and running within an instance of Microsoft SQL Server 2005. If you do not have SQL Server 2005 or a working ODM database, please consult the information in the Installation Prerequisites Section above and in the appendices of this document for information on how to remedy this situation.

Before you can begin using the functionality of the ODM Tools application, you must connect the application to an ODM database implemented in Microsoft SQL Server 2005. The first time you open the ODM Tools you will be prompted for the information required to connect to the database. Complete the following steps to connect the ODM Tools application to your ODM Database.

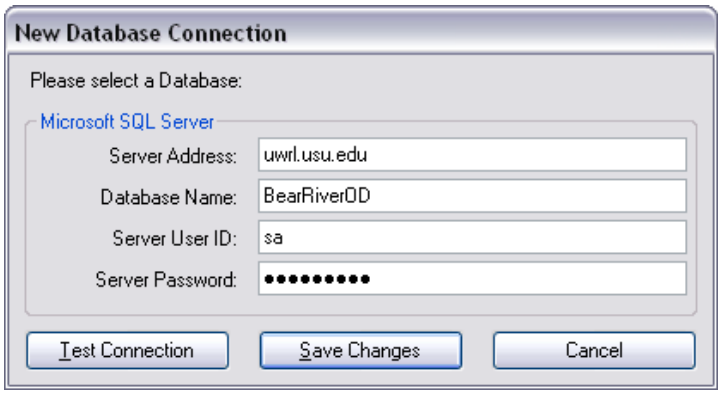
1. Start the ODM Tools application by clicking on Start --- All Programs --- CUAHSI HIS --- ODM Tools. You can also start ODM Tools by double clicking on the shortcut on your desktop. The following window will appear.



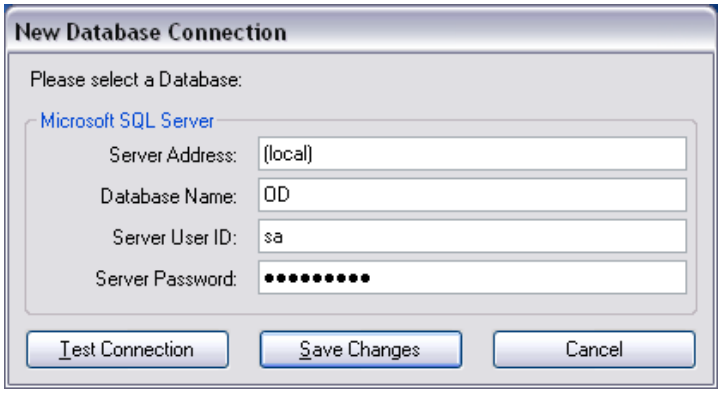
NOTE: ODM Tools can connect to an ODM database implemented within any version of Microsoft SQL Server 2005 (i.e., Express, Standard, Enterprise) using SQL Server authentication. Both local and remote SQL Server databases can be accessed using the ODM Tools application. ODM Tools assumes that you have already been given a SQL Server authentication login with a username and password that has access to the database that you want to connect to.

NOTE: You must know the server address to connect with a database implemented on that server. This is the name of the computer on which your instance of SQL Server is installed (in the above example it is a remote server called uwrl.usu.edu). If you have installed the ODM Tools application on the same computer as your SQL Server instance, the Server Address is "(local)" (if you are using SQL Server 2005 Standard or Enterprise installed as the default instance) or "(local)\SQLEXPRESS" (if you are using SQL Server 2005 Express). See the following examples. If you do not know the server address, consult with the administrator of your SQL Server database.

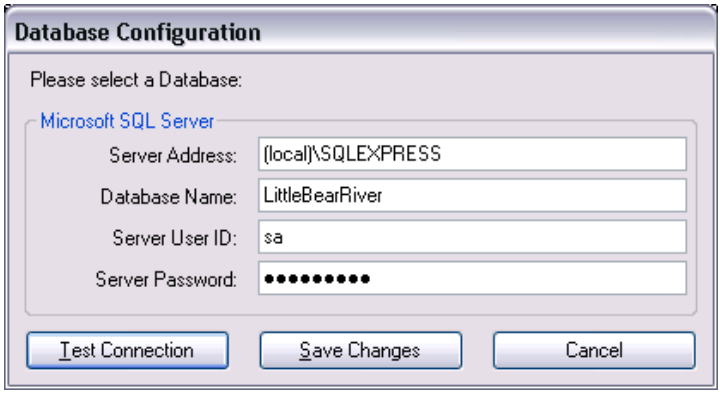
Example database connection information for an ODM database on a remote server using SQL Server 2005 Standard or Enterprise installed as the Default instance:



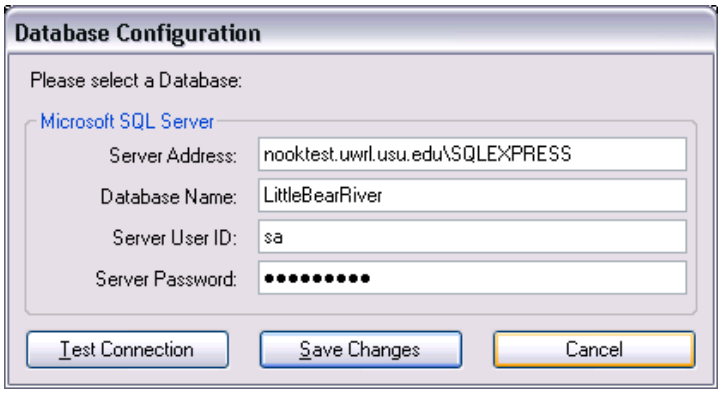
Example database connection information for an ODM database on the local machine using SQL Server 2005 Standard or Enterprise installed as the Default instance:



Example database connection information for an ODM database on the local machine using SQL Server 2005 Express:



Example database connection information for an ODM database on a remote machine using SQL Server 2005 Express:



NOTE: By default, SQL Server 2005 Express does not allow remote connections. You must enable this feature before you can connect ODM Tools to a remote SQL Server 2005 Express database. See the instructions at the following URL for enabling remote connections using SQL Server 2005 Express: http://support.microsoft.com/kb/914277.

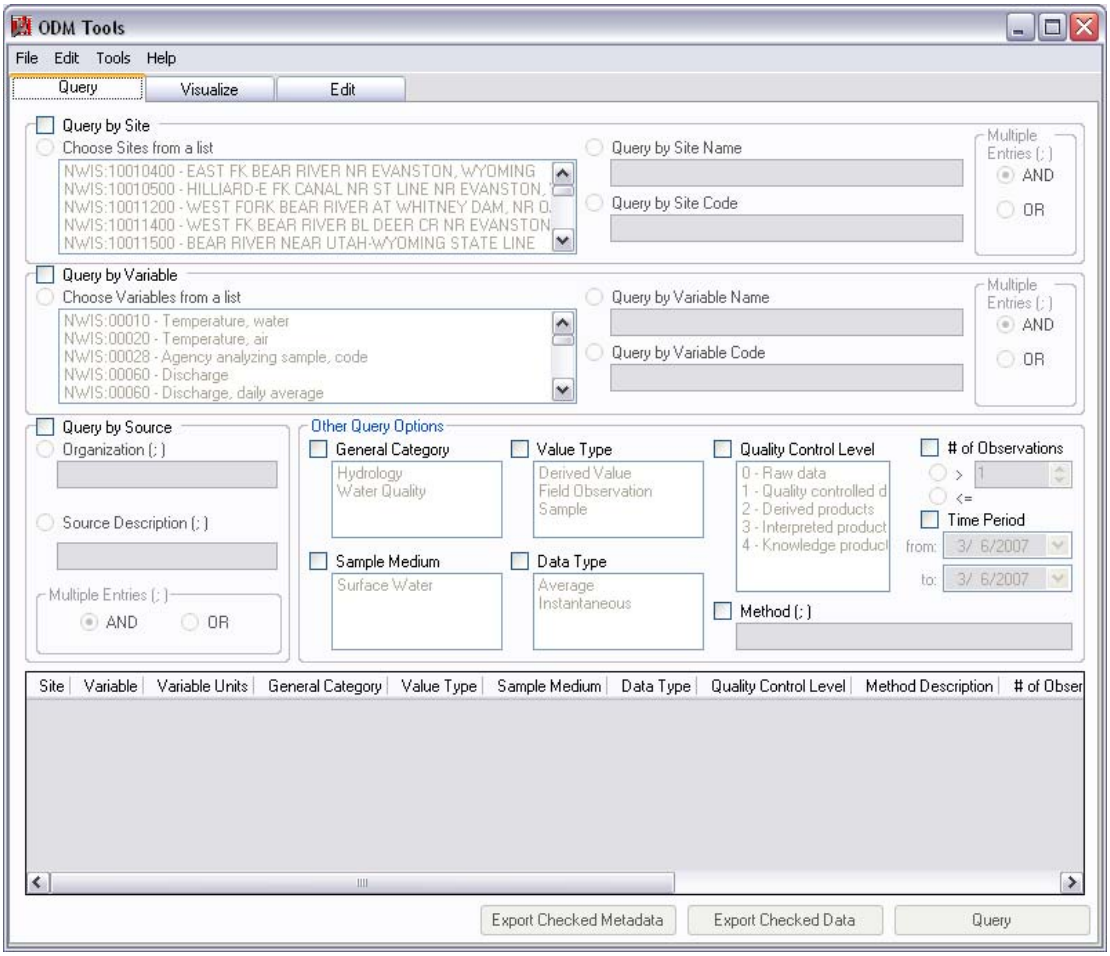
2. Type the information for your ODM database into the text boxes provided, including the server address, the database name, and your SQL Server user name and password. Click the "Test Connection" button to test the connection to your ODM Database. If your connection is successful, the following window will appear.

3. Click "OK" to return to the New Database Connection screen. Click the "Save Changes" button to save your database connection information. Once you have completed this step, your database connection information will be saved in a configuration file and the ODM Tools application will open with a connection to your ODM database. The next time you open the ODM Tools application it will use the database connection that you just set up.

If you wish to change the database connection after you have completed steps 1-3 above (for instance if you want to point ODM Tools at a different ODM database) you can access the database connection information by clicking on the "Edit" pull down menu in the ODM Tools application and selecting "Database Connection." This will open the same dialog as before, and you can change your database connection. Click the "Save Changes" button to save any changes to the database connection and return to the ODM Tools application.

4.0 Querying and Exporting Data Series

The CUAHSI HIS ODM has within it the concept of a "data series." Each data series in the ODM represents a unique Site, Variable, Method, QualityControlLevel, and Source combination, and the SeriesCatalog table in the ODM provides a listing of all of the distinct series of data values stored in the ODM. ODM Tools provides the ability to query an instance of the ODM for specific data series based on information contained in one or more fields in the SeriesCatalog table. Once specific data series are identified, users can then export them to a delimited text file in the CUAHSI HIS MyDB format. The following figure shows the Query tab of the ODM Tools application, and the following sections describe the data series query and export functionality of ODM Tools.



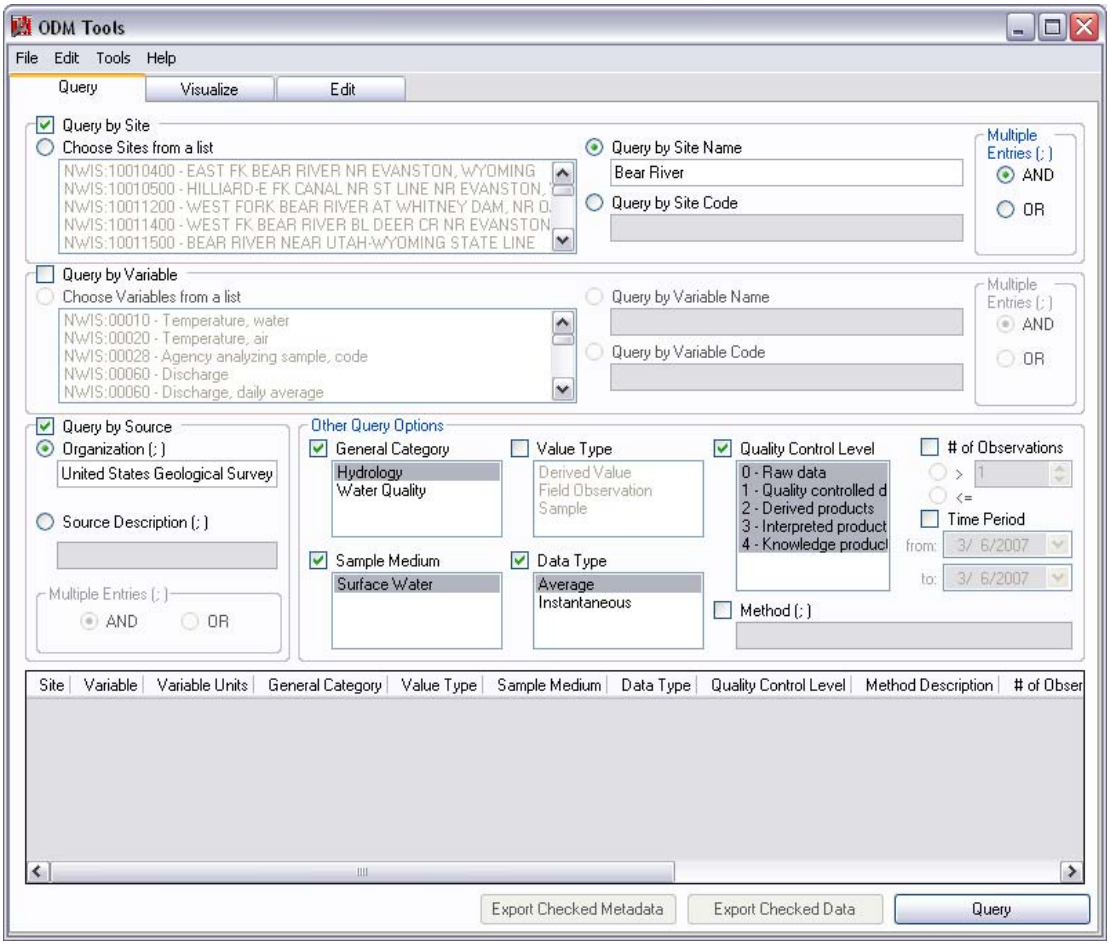
4.1 Selecting Data Series

The Query tab allows users to query the ODM database to find specific data series and then either export these data series or send them to the Visualize or Edit tabs. Users can search for data series by site (either by selecting from a list or searching for a text string in the site code or name), by variable (either by selecting from a list or searching for a text string in the variable code or name), by Source, by General Category, by Sample Medium, by Value Type, by Data Type, by Quality Control Level, by Method, by number of observations, and by Time Period.

The following example illustrates how data series can be identified using the query options on the Query Tab.

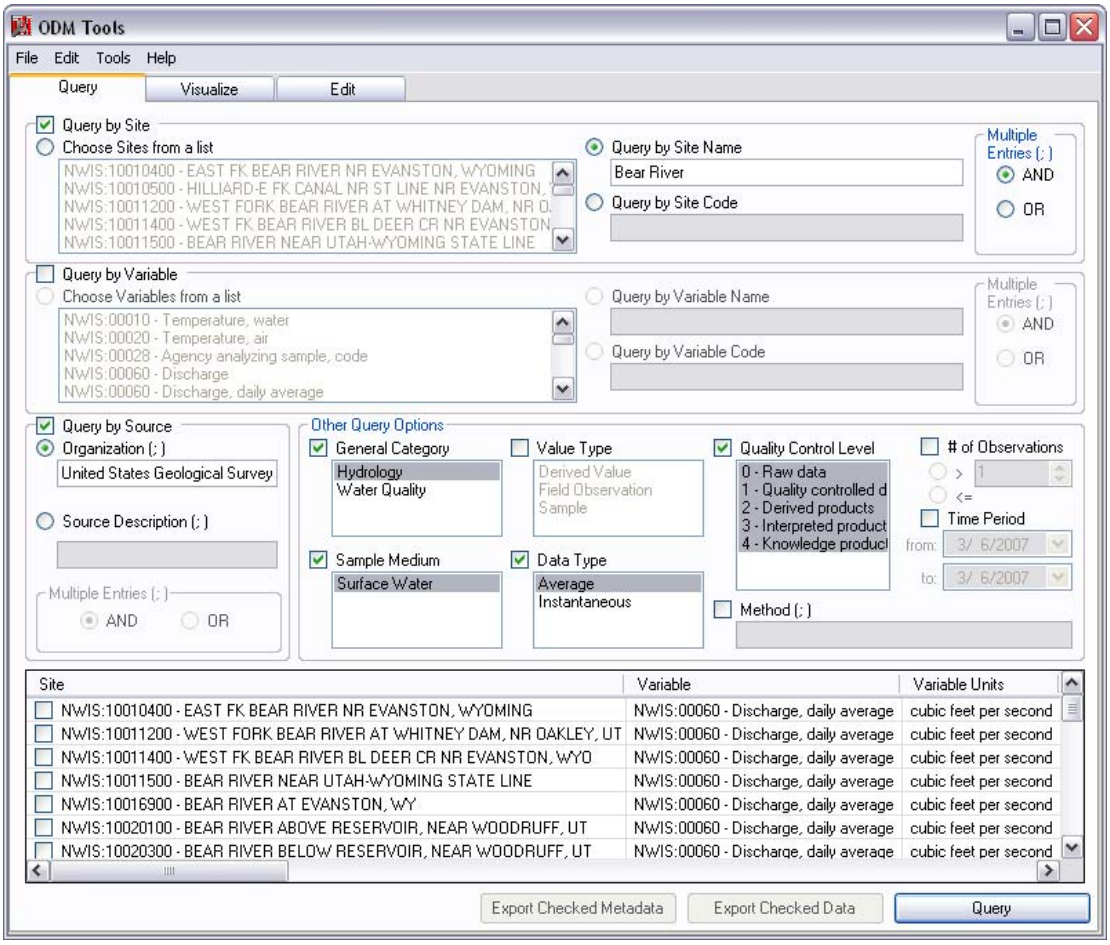
1. Make sure that the Query tab is active by clicking on its tab at the top of the ODM Tools application.

2. Enter criteria for one or many of the data series attributes listed on the Query tab. First check the box next to the attributes for which you want to add criteria (i.e., Query by Site, or Query by Variable). Next, either select criteria from the lists given or type your search text into the text boxes. Your screen might look something like the following:



NOTE: Multiple search strings can be used in the Query by Site Name, Query by Site Code, Query by Variable Name, and Query by Variable Code search options. To do so, enter the search strings into the text box separated by a semicolon and then select "AND" or "OR" from the Multiple Entries options. For example, to search for Sites with Site Name containing the text "Bear River" or "Logan River", enter the string "Bear River; Logan River" in the Query by Site Name text box and then select "OR" as the Multiple Entries option.

3. Click on the "Query" button at the bottom of the form. You will notice that the data grid at the bottom of the form is populated with any data series that are returned by your query. See the following figure for an example:



Each of the records in the table at the bottom of the form represents one data series. You can view all of the attributes of each data series by using the scroll bars to the right and bottom of the tabular list of data series. You can further limit the number of data series returned by your query by adding additional or more specific criteria in the query options on the form and then re-running the query by clicking on the "Query" button.

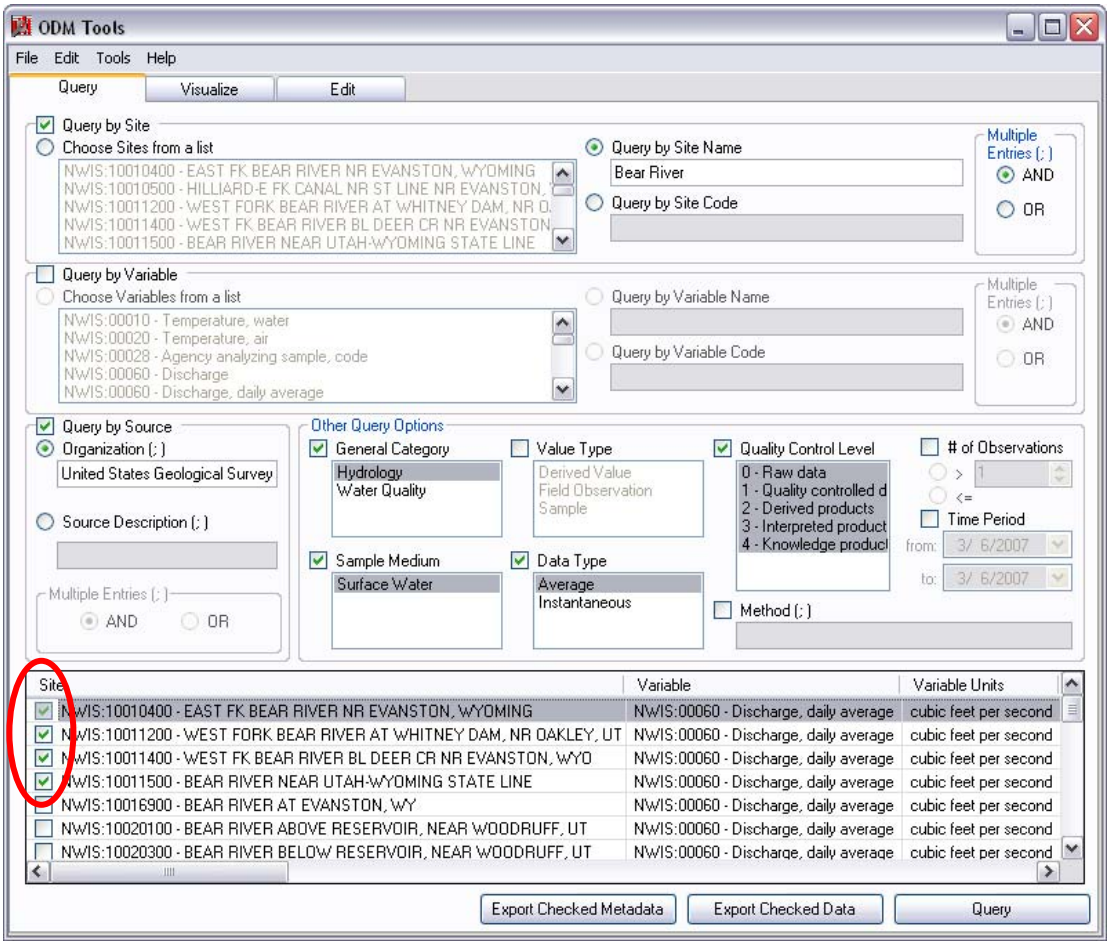
4.2 Exporting Data Series to MyDB

ODM Tools uses the MyDB table format for data export. MyDB is a simplified version of the ODM that consists of a single table with many of the most important fields from the ODM. Appendix B of this document describes the MyDB table format in detail. MyDB was designed by the CUAHSI HIS Team and in the future will serve as the underlying data source for many of the tools developed as part of the HIS Analyst Toolkit under development by the HIS Team. MyDB is a delimited text file that can easily be loaded into many data visualization and analysis software programs such as Microsoft Excel.

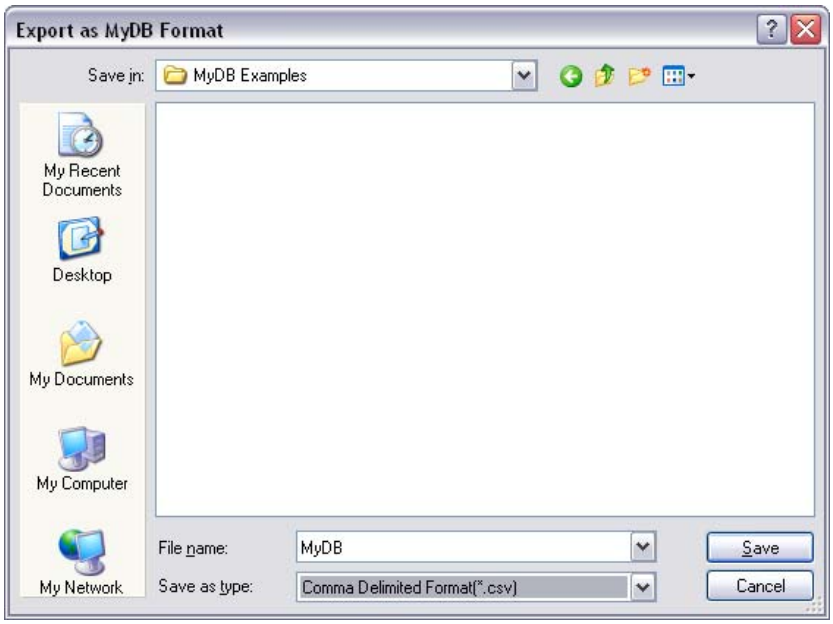
Once you have identified a set of data series using the query options, you can select them for export by clicking on the check box in the left most column of the data grid that contains the list. One or more data series can be selected for export to a single MyDB file by clicking on their selection check boxes. To export data series from ODM, use the following steps:

1. Using the query functionality, select a set of data series using the steps described in Section 4.1 above.

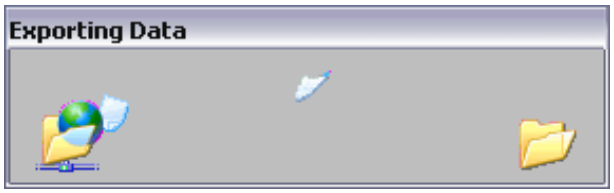
2. Click the check boxes next to one or more data series to select them for export to MyDB. Your screen may look something like the following:



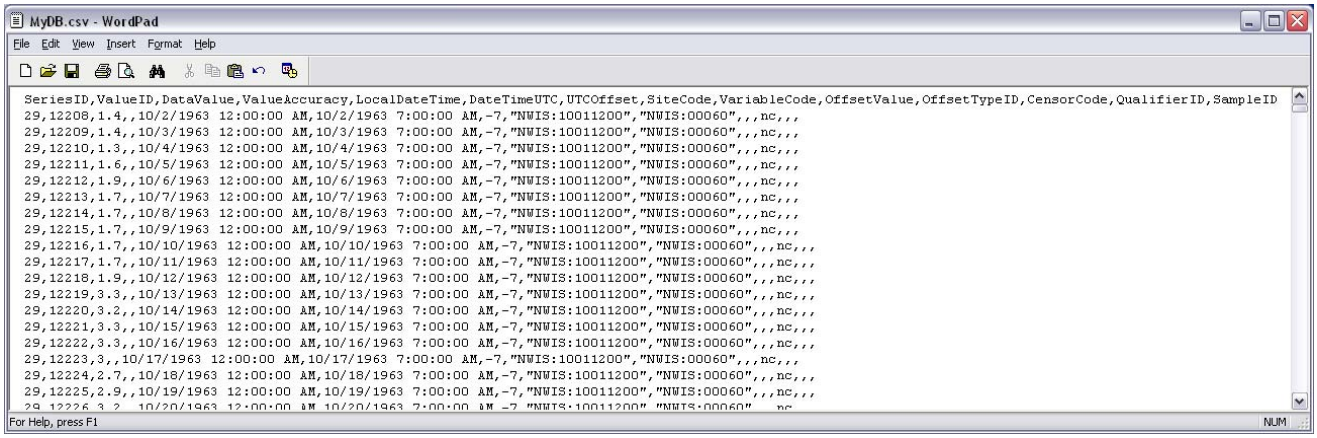
3. Click the “Export Checked Data” button. The following window will appear.



4. Navigate to the location on disk where you want to save your MyDB file. Give your MyDB file a name and then select a file type – either comma delimited format (.csv) or tab delimeted (.txt) format. Click the "Save" button to save your MyDB file. The following windows will appear.



5. Click on the "OK" button to return to the ODM Tools application. The following is an example of a MyDB file exported to a comma separated values text file.



NOTE: The above example shows a MyDB file with a minimum number of export columns. The MyDB file exported by ODM Tools can be customized to include additional attributes of the data values using the data export options described in Section 4.4 below.

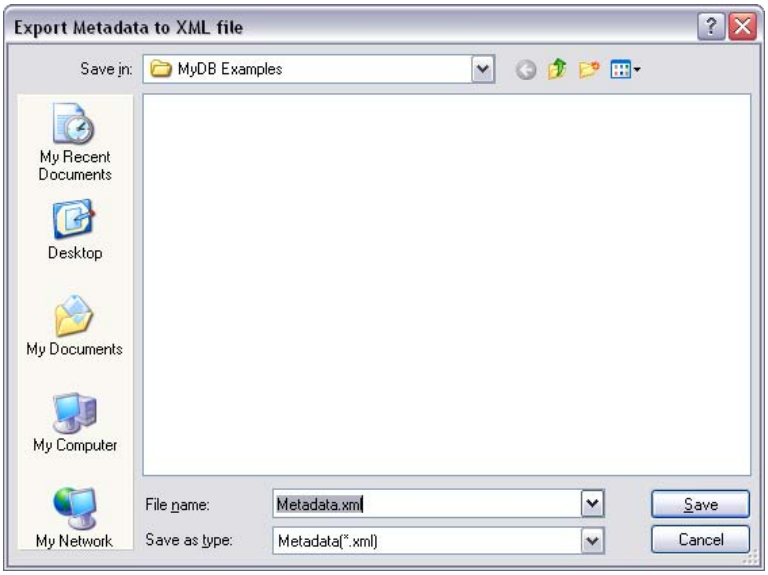
NOTE: Data series can also be exported by right clicking on them in the data grid at the bottom of the form and selecting "Export Single Data" from the context menu. This will pop up the same file dialog shown above that will allow you to save the MyDB file for the data series that you right clicked.

4.3 Exporting Metadata

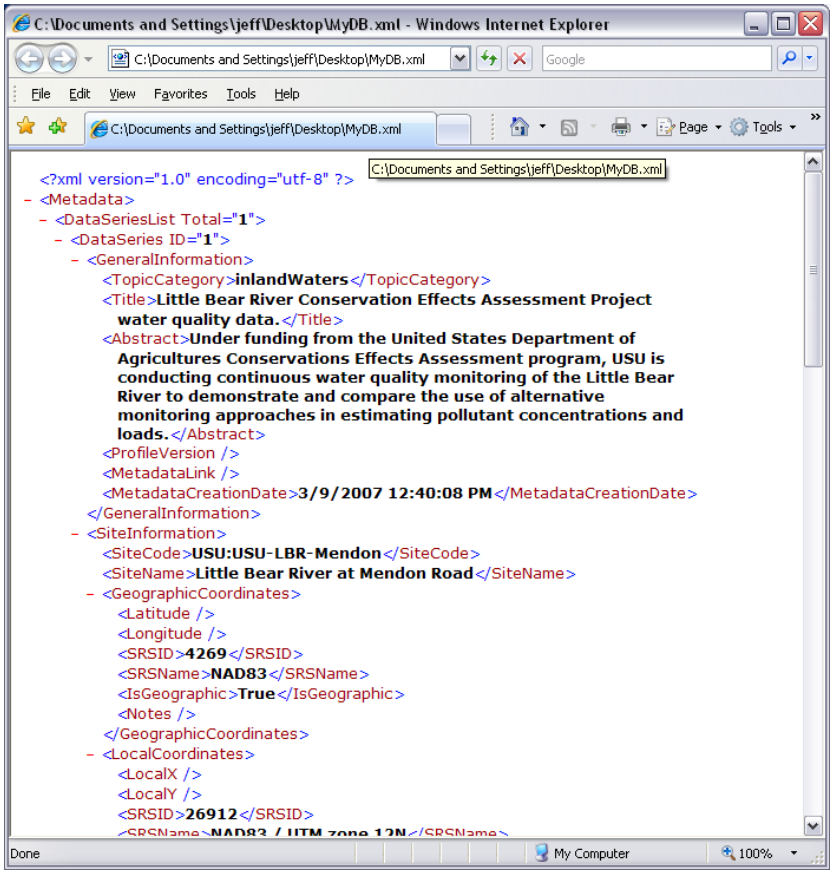
ODM provides the functionality to store a substantial amount of metadata about each data value and each data series contained in the database. This metadata includes information about the site at which the observations were made, information about the variable that was measured and the methods that were used to make the measurements, information about the organization that collected the data, and other supporting information. Much of this information is not included in the MyDB table format; however, in order to maintain the integrity of data series exported from ODM, ODM Tools includes functionality to view and export the metadata associated with one or more selected data series. The exported metadata file contains a snapshot of all of the metadata stored in ODM for the data series that are being exported. See Appendix B for a detailed description of the MyDB metadata file. Exporting a metadata file can be accomplished using the following steps:

1. Follow steps 1 and 2 above for exporting data series to a MyDB file.

2. Click the "Export Checked Metadata" button. The following window will appear.



3. Navigate to the location on disk where you want to save your metadata file. Give your metadata file a name and then click the "Save" button to save your metadata file. The following is an example of the metadata file.



NOTE: Once data series have been identified using the query options, you can also export the metadata for a single data series by right clicking on it in the data grid at the bottom of the form

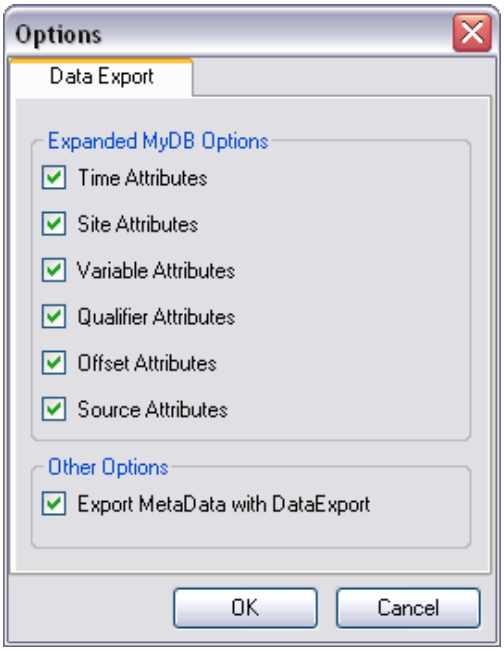
and selecting “Export Single Metadata” from the context menu. If you wish to preview the metadata prior to saving it to disk, you can right click on a data series in the data grid and select “View Metadata” from the context menu. This will launch the metadata file for the currently selected data series to the default XML file viewer on your machine (most likely Internet Explorer).

NOTE: The format and contents of the metadata file are described in detail in Appendix B of this document.

4.4 Data and Metadata Export Options

ODM Tools provides users with options for customizing the contents of the export MyDB file and for customizing the behavior of the metadata export. The data and metadata export options can be accessed using the following steps.

1. Click on the “Tools” pull down menu and select “Options”. The following window will appear.



2. By clicking on the check boxes next to attribute groups under “Expanded MyDB Options,” you can control the attribute groups that are written to MyDB files when they are exported. Choose your MyDB export options by clicking any check boxes next to the attribute groups that you are interested in. The attribute groups selected on the “Options” form will be written to any subsequent MyDB table exports.

3. The “Options” screen also allows you to control the behavior of the metadata export. If you click the checkbox next to “Export MetaData with DataExport,” ODM Tools will by default export a metadata file with every MyDB file that is exported. When this box is checked, each time a MyDB file is exported, an accompanying metadata file with the same file name (with a .xml extension) is written to the same location as the MyDB file.

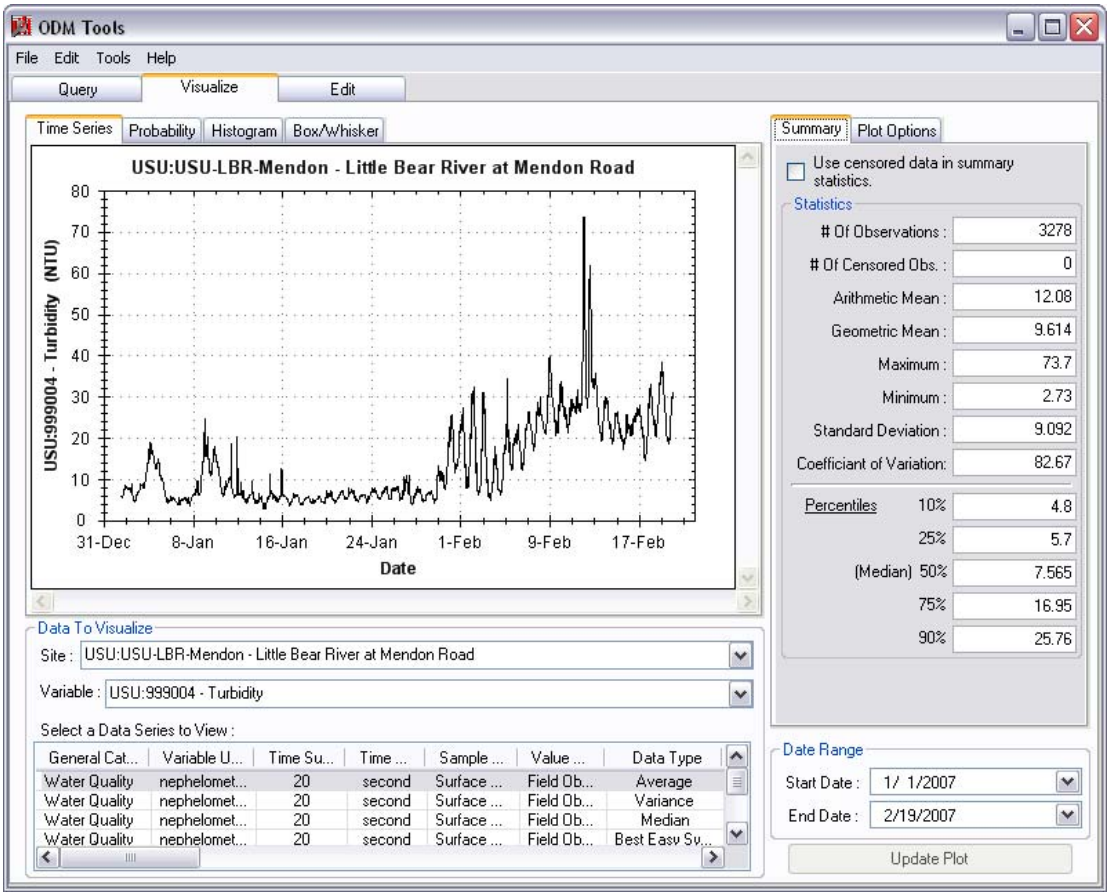
This ensures that the full data series specification (consisting of the MyDB table AND the metadata file) is preserved in the export. Choose your export options and then click the OK button. This will return you to the main ODM Tools application.

5.0 Visualizing and Summarizing Data Series

ODM Tools provides users with the capability to visualize data series using a variety of plot types and to generate simple descriptive statistics for data series. The following sections describe the data series visualization and summary statistics generation capabilities of the ODM Tools application.

5.1 Plotting Data with ODM Tools

The data series visualization and statistical summary tools are contained within the "Visualize" tab of the ODM Tools application. The following is a screen shot of the ODM Tools "Visualize" tab.

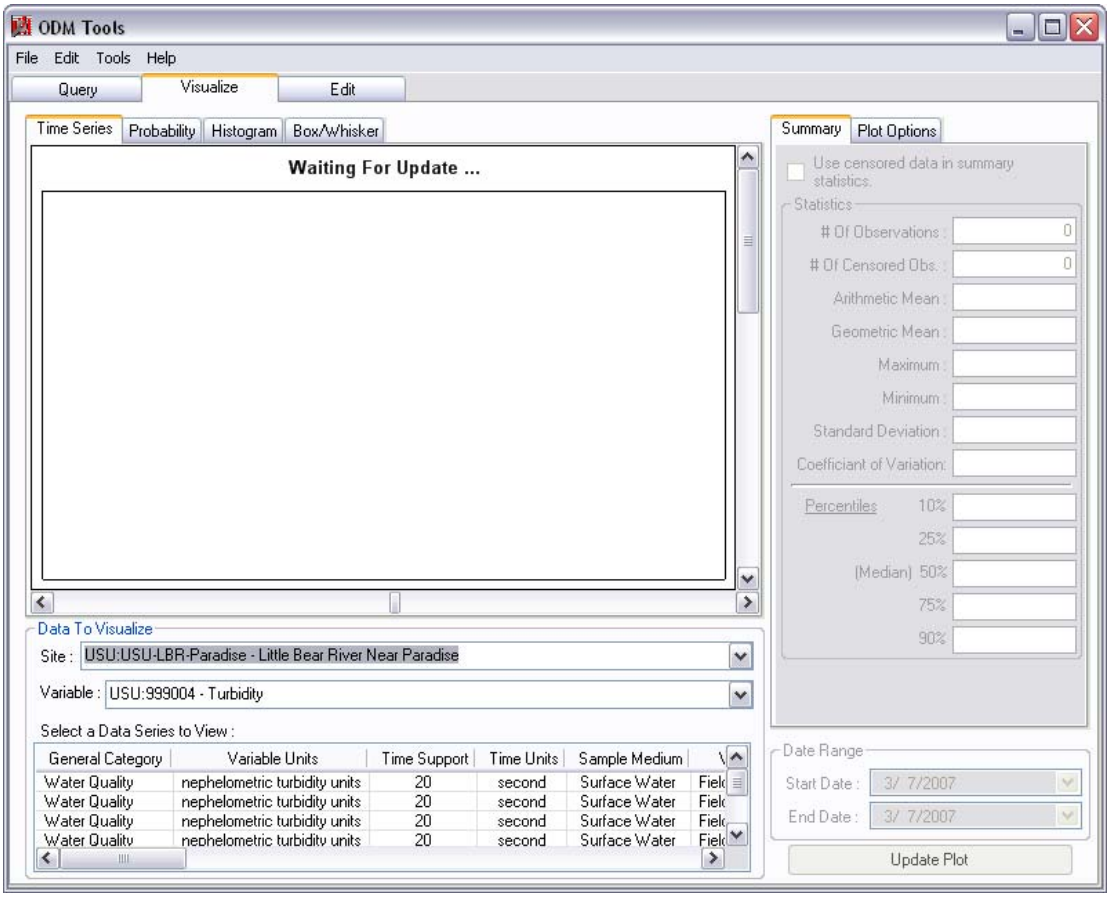


The "Visualize" tab can be accessed using one of two methods. First, users can access this functionality by clicking on the "Visualize" tab at the top of the ODM Tools application window. Alternatively, users can right click on a selected data series in the data grid at the bottom of the "Query" tab and select "Plot Data" from the context menu.

5.1.1 Selecting a Data Series for Plotting

Use the following steps to use the data visualization capabilities of ODM Tools.

1. If you haven't already, click on the "Visualize" tab at the top of the ODM Tools application window. This will take you to the "Visualize" tab. You will notice that much of the functionality on this form is disabled to begin with and no plot is shown (see the following figure).

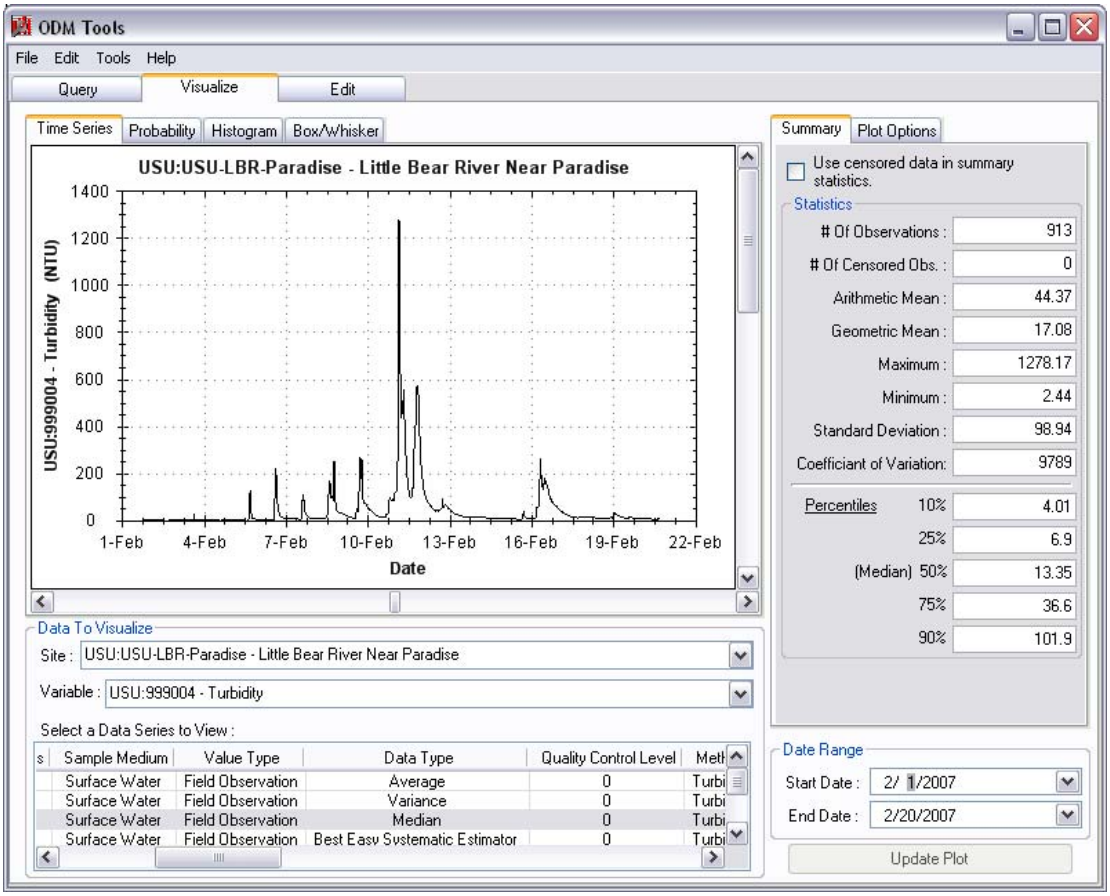


2. Select a site from the "Site" pull down menu, which is located near the bottom of the form.

3. Select a variable from the "Variable" pull down menu, which is located just beneath the "Site" pull down menu. You will notice that the data grid at the bottom of the form is populated with the list of data series available for the site and variable combination that you have chosen. Data series are listed in a data grid because it is possible to have multiple data series for a single site/variable combination (an example would be where at a single site a raw data series and a quality controlled data series exist for the same variable).

4. Once you have selected a site and variable, select a data series in the list at the bottom of the form by clicking on its row in the data grid. You can view all of the attributes for each data series in the grid by using the scroll bars at the bottom and right of the data grid.

5. Click the "Update Plot" button, which is located at the bottom right of the form, to generate the plot and descriptive statistics. Your screen may look something like the following.



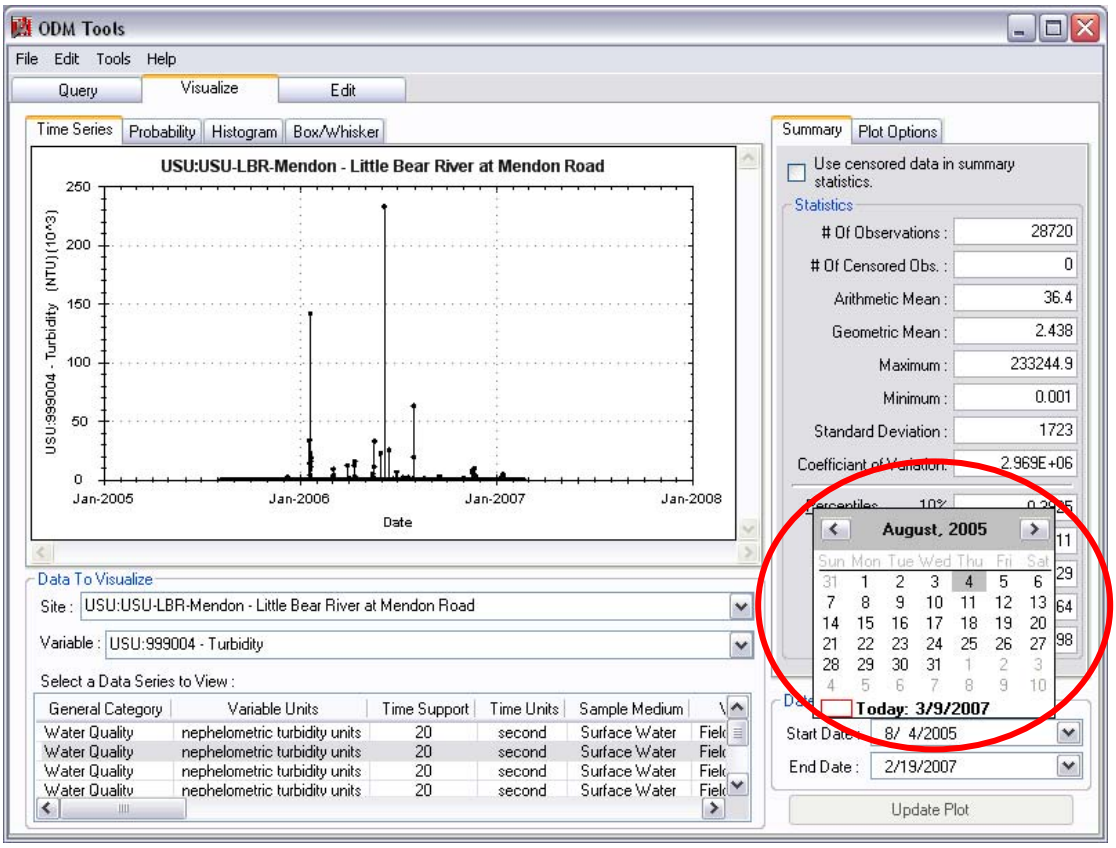
NOTE: Each time a new data series is selected, the plot window is cleared. You must then click the "Update Plot" button to generate a new plot of your newly selected data series.

5.1.2 Restricting the Date Range

ODM Tools provides the functionality to limit the date range for data shown in the plot window and summarized in the descriptive statistics. The "Start Date" and "End Date" for the data series are located at the bottom right of the form. These dates are automatically populated for each data series when they are selected in the data grid. The data shown in the plot and the statistics in the "Summary" tab are reflective of the data within the date range shown. If you wish to restrict the date range for a data series, you can click on the down arrows to the right of each date and select a new one from the calendar (see below), or you can simply click on the year, month, or day of one of the dates and use the up and down arrows to adjust the dates.

NOTE: If you restrict the date range for a data series, the plot will be cleared. Similar to the Note above, you must click the "Update Plot" button to regenerate the plot. Once it is regenerated, the plot will show the data within the date range that you select and the "Summary"

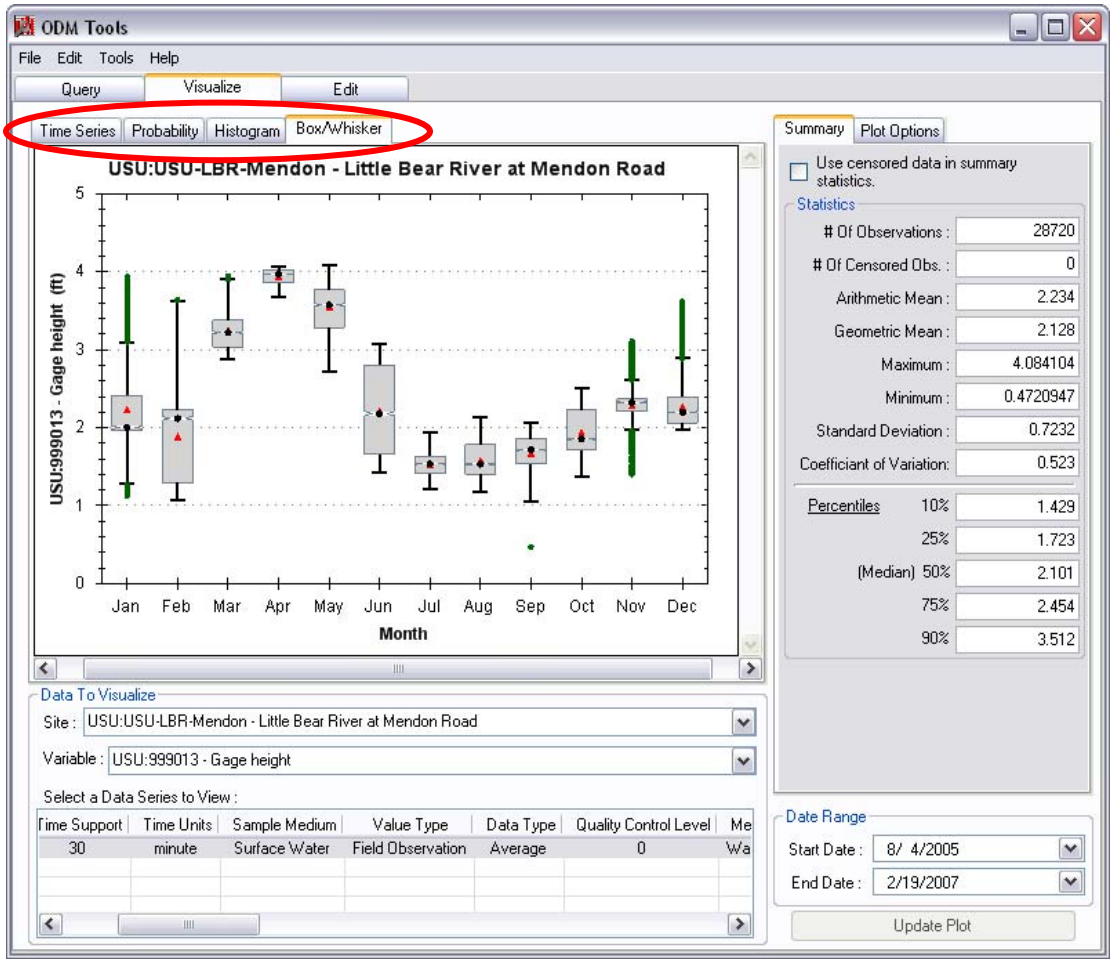
tab will show summary statistics that are representative of the data that fall within the range that you select.



5.1.3 Switching Between Plot Types

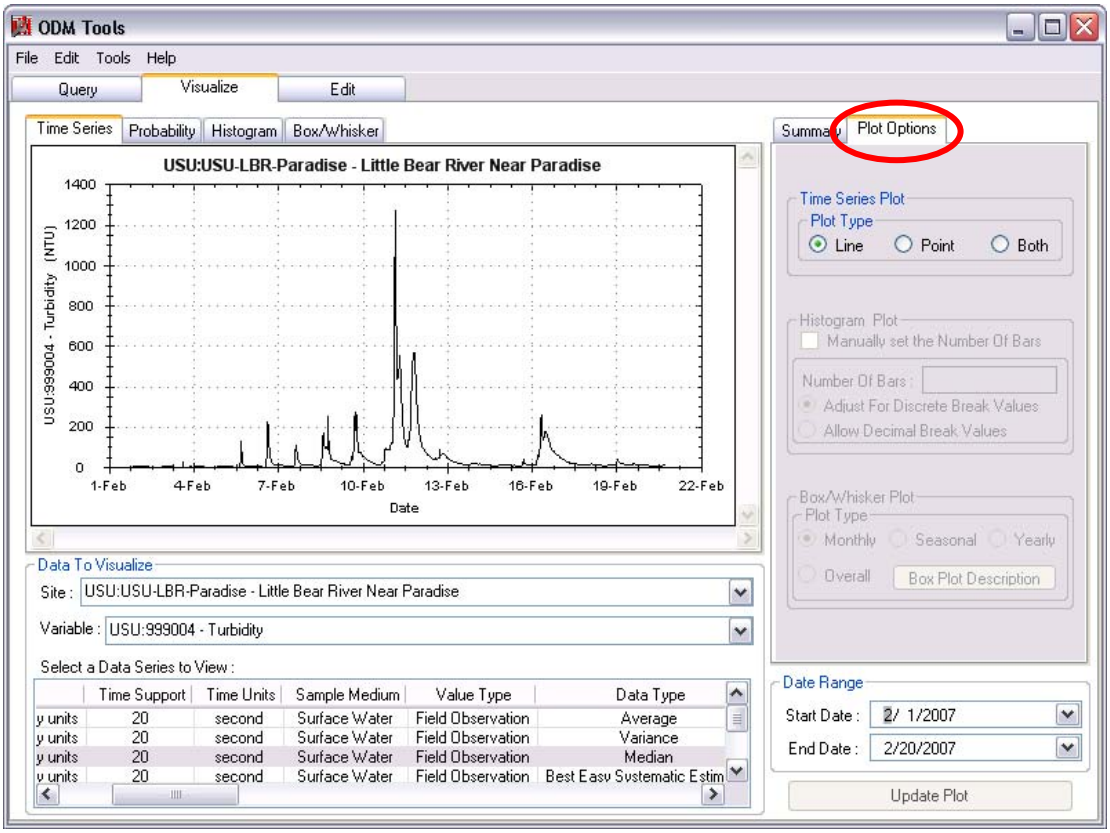
ODM Tools provide several different plot types that can be used to visualize the data. You can switch the plot types by clicking on the tabs at the top of the plot window (see figure below). Plot types currently included in ODM Tools include time series, probability, histogram, and box and whisker plots.

NOTE: Each time you select a new plot type by clicking on its tab, you must click the "Update Plot" button to generate the plot.



5.1.4 Changing Plot Options

Each of the plot types has one or more customization options that can be accessed by clicking on the “Plot Options” tab at the top right of the form (see below).



When you click on the "Plot Options" tab, only the options for the currently selected plot will be available. The plot options are automatically applied to the current plot when you click on them. The following is a list of all of the available plot options by plot type.

Time Series Plot – You can switch between "Line" (only the line is shown), "Point" (only the points are shown), or "Both" (both the line and the points are shown).

Histogram – You can control the number of bars shown in the histogram by clicking on the check box next to "Manually set the Number of Bars" and then inputting a number into the "Number of Bars" text box. There is an upper limit of 20 bars on the histogram plot. You must select a number less than 20. You can also choose between discrete break values or decimal break values for the bars.

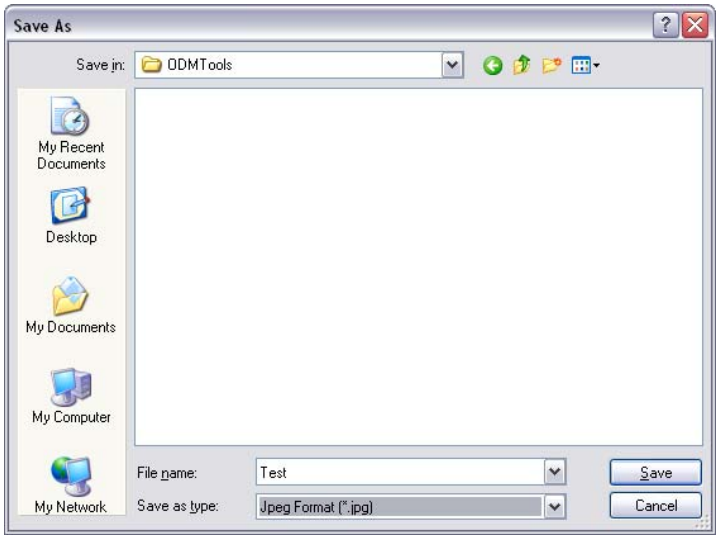
Box and Whisker Plot – You can choose the time period for the boxes from monthly (all data values are grouped by month), seasonal (all data values are grouped by season), yearly (all data values are grouped by year), or overall (all of the data values in one group) by clicking on the radio button next to each of these options. You can also access a description of the statistics included in the plots by clicking on the "Box Plot Description" button.

5.1.5 Exporting Plots

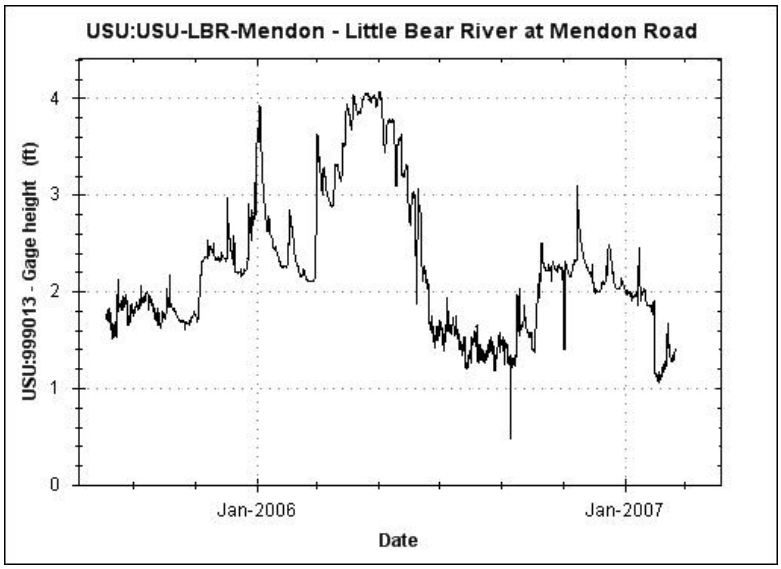
The plots generated by ODM Tools can be exported for use in documents, presentations, etc. To export a plot, use the following steps.

1. Select a site, variable, data series, time period, and plot type using the steps outlined above and then click the "Update Plot" button. ODM Tools will export whatever plot is shown in the plot window.

2. Right click on the plot window and select "Save Image As" from the context menu. The following form will appear.



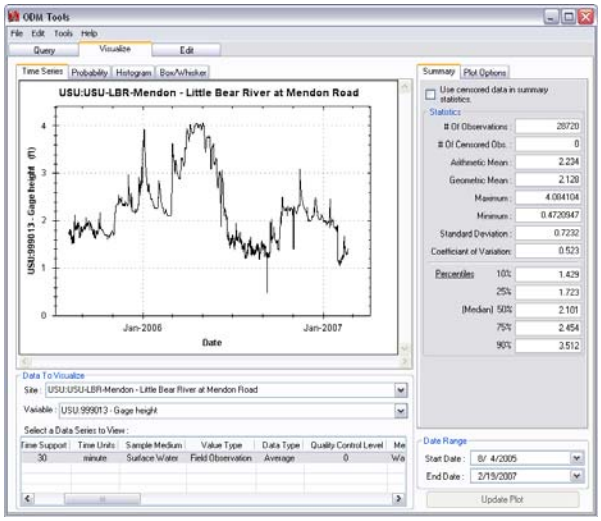
3. Give the plot a file name in the "File Name" text box and then select a file type from the "Save as type" pull down menu. Click the "Save" button to save the file to disk. The following is an example of an ODM Tools plot exported as a jpeg file and then imported to this document.



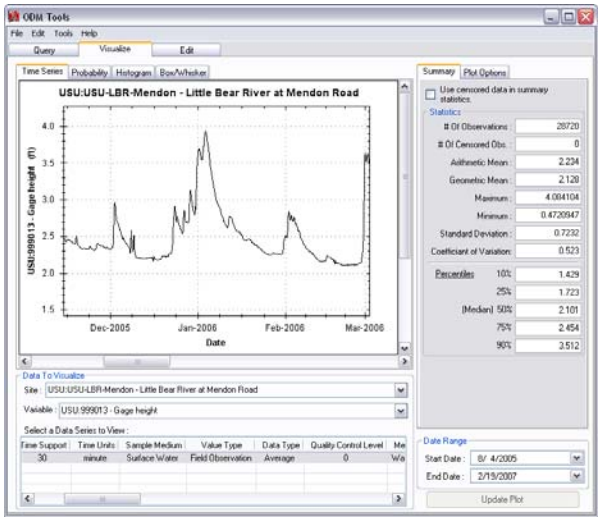
NOTE: You can also copy the current plot to the clipboard by right clicking on the plot window and selecting "Copy" from the context menu. The image on the clipboard can then be pasted into another application such as Microsoft Word or PowerPoint.

5.1.6 Plot Zooming

The plot window within ODM Tools allows you to zoom in on the data that are shown. You can do this by left clicking on the plot window and while holding down on the mouse button dragging a box around the area that you are interested in. The plot will be zoomed to the area that you selected. See the following figures.



View of Full Data Series



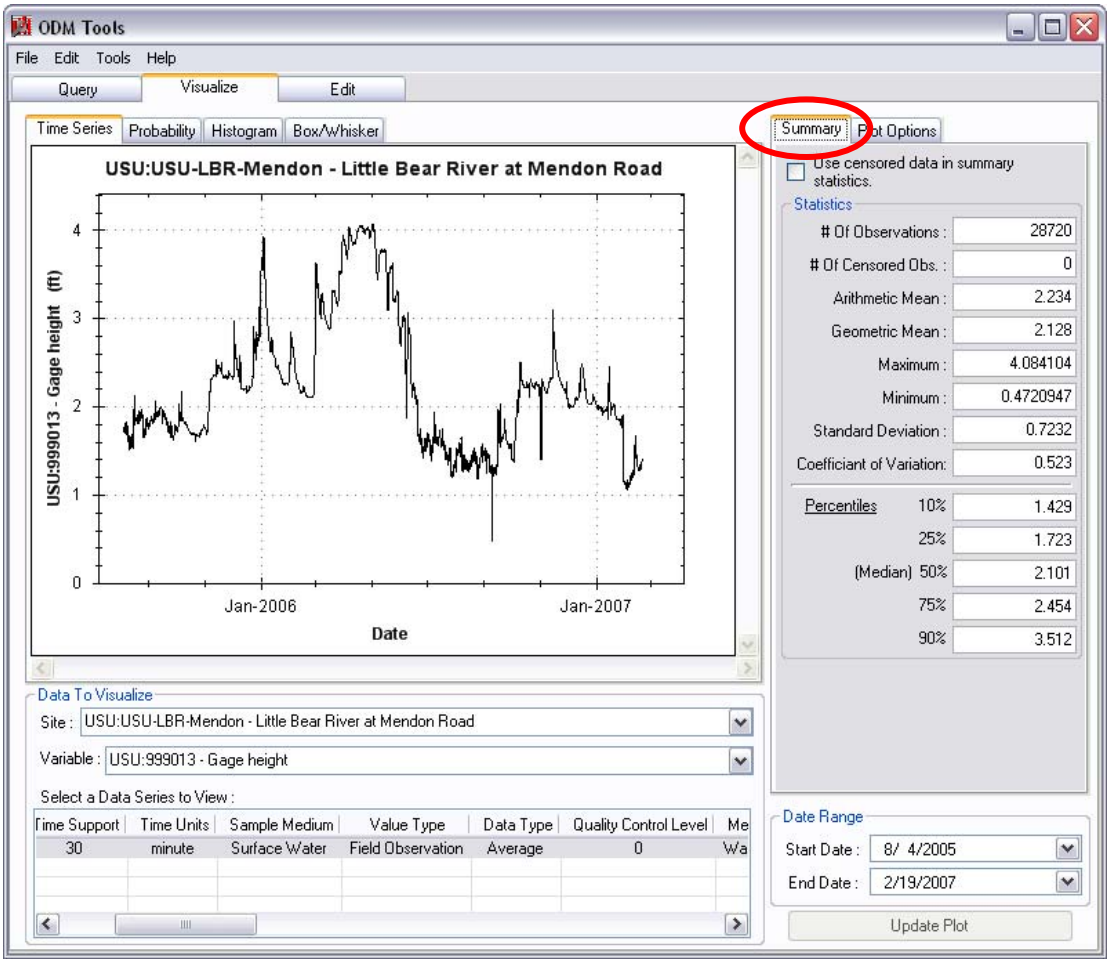
Plot zoomed to a portion of the data series

NOTE: When you zoom in on a plot, the date range of selected data is unaffected. Because of this, you will notice that the summary statistics do not change. The summary statistics only change when the date range is restricted at the bottom of the form.

You will notice when you zoom in that scroll bars become available on the bottom and right of the plot window. On the time series plot, you can use these scroll bars to scroll through time with a fixed time window on the plot. To undo the zoom on the plot, right click on the plot window and select "Un-Zoom" from the context menu. This will take you back one zoom level, or to your original position of you have only zoomed in once. To return the plot to the extents of the entire data series, right click on the plot window and select "Set Scale to Default."

5.2 Viewing Descriptive Statistics

ODM Tools provides summary statistics for data series on the "Summary" tab. The "Summary" tab can be accessed by clicking on it at the top right of the form (see figure below).



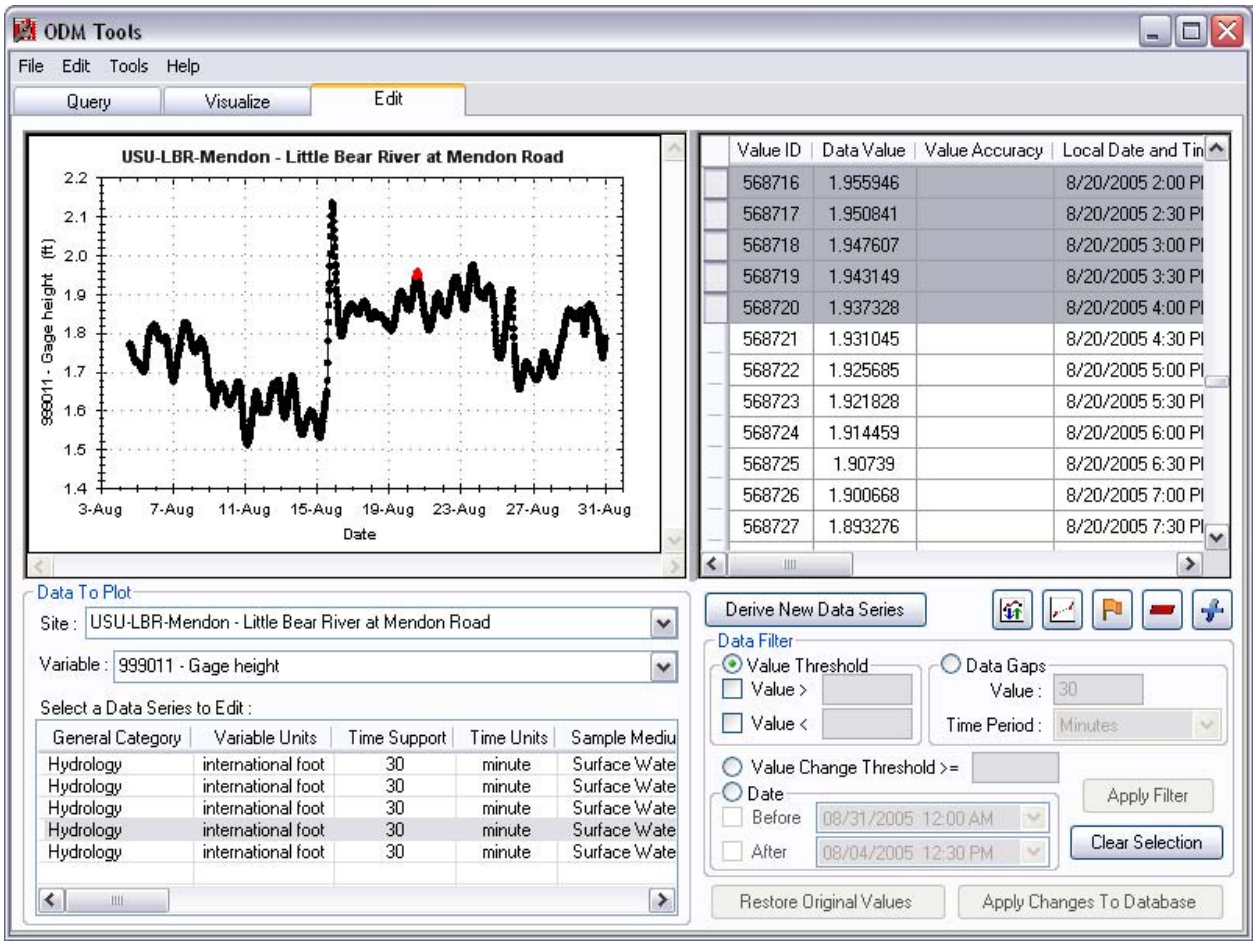
The "Summary" tab provides descriptive statistics for the selected data series and date range. If you restrict the date range for a data series and click the "Update Plot" button, both the plot and the summary statistics on the "Summary" tab are limited to the date range that you have chosen.

For data series that have censored data values, you can click the check box next to the "Use censored data in summary statistics" option to include these values in the computation of the summary statistics. Summary statistics for data series with censored data values are calculated using robust methods described in Helsel and Hirsch (2002). The summary statistics presented are subject to the following constraints:

- Censored data statistics are calculated only for a single censoring level. Multiple censoring levels are not currently supported.
- Censored data statistics are calculated only for datasets with observations below a censoring level. Datasets with values above a censoring level are not currently supported.

6.0 Editing Existing and Deriving New Data Series

ODM Tools includes functionality to edit the data values and some of the attributes of the data values within data series stored within an instance of the ODM. This is useful, for example, in performing manual quality assurance and quality control of data series, where some data values may need to be deleted, adjusted, or interpolated. In addition, ODM Tools provides functionality to derive new data series from existing data series. For example, daily average data values can be derived from more frequent observations using ODM Tools' aggregate functions. All of the data series editing and creation tools are on the Edit tab of the ODM Tools application. The following figure shows the ODM Tools Edit tab, and the following sections describe the data series creation and editing capabilities that are available on this tab.



6.1 Deriving New Data Series

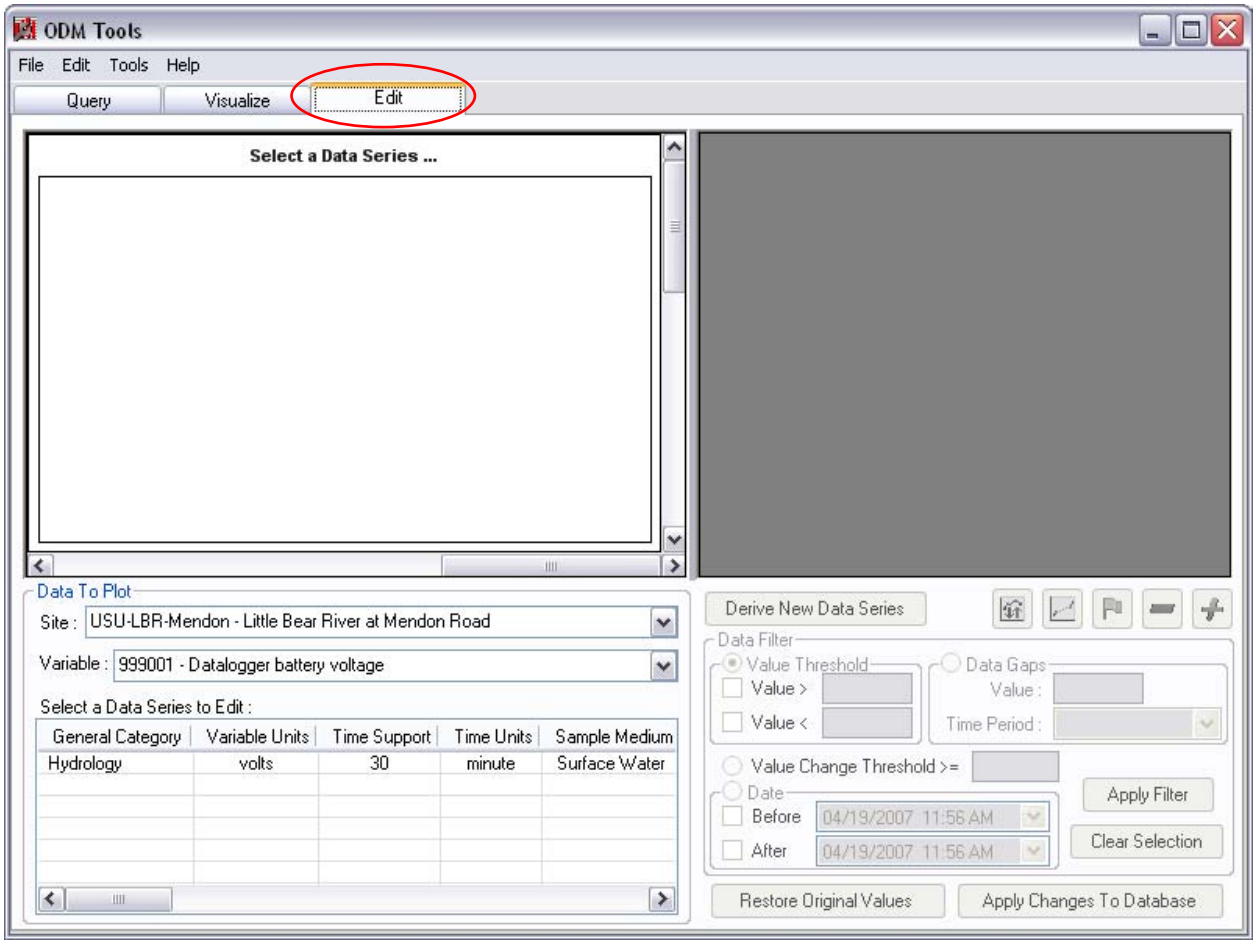
ODM Tools provides functionality to create entirely new data series (derived data series) from data series that are already stored in an instance of the ODM. Examples include creating a quality controlled data series from a raw data series, creating a daily aggregate (daily min, max, or average) data series from a data series with more frequent (i.e., hourly) observations, and deriving a new data series from an existing data series by using an algebraic transformation. This functionality is accessed by clicking on the Derive New Data Series button, which is located

just below the table view of the data, after a data series has been selected. The following sections describe the specific data series derivation functionality supported by ODM Tools.

6.1.1 Creating a Quality Controlled Data Series from a Raw Data Series

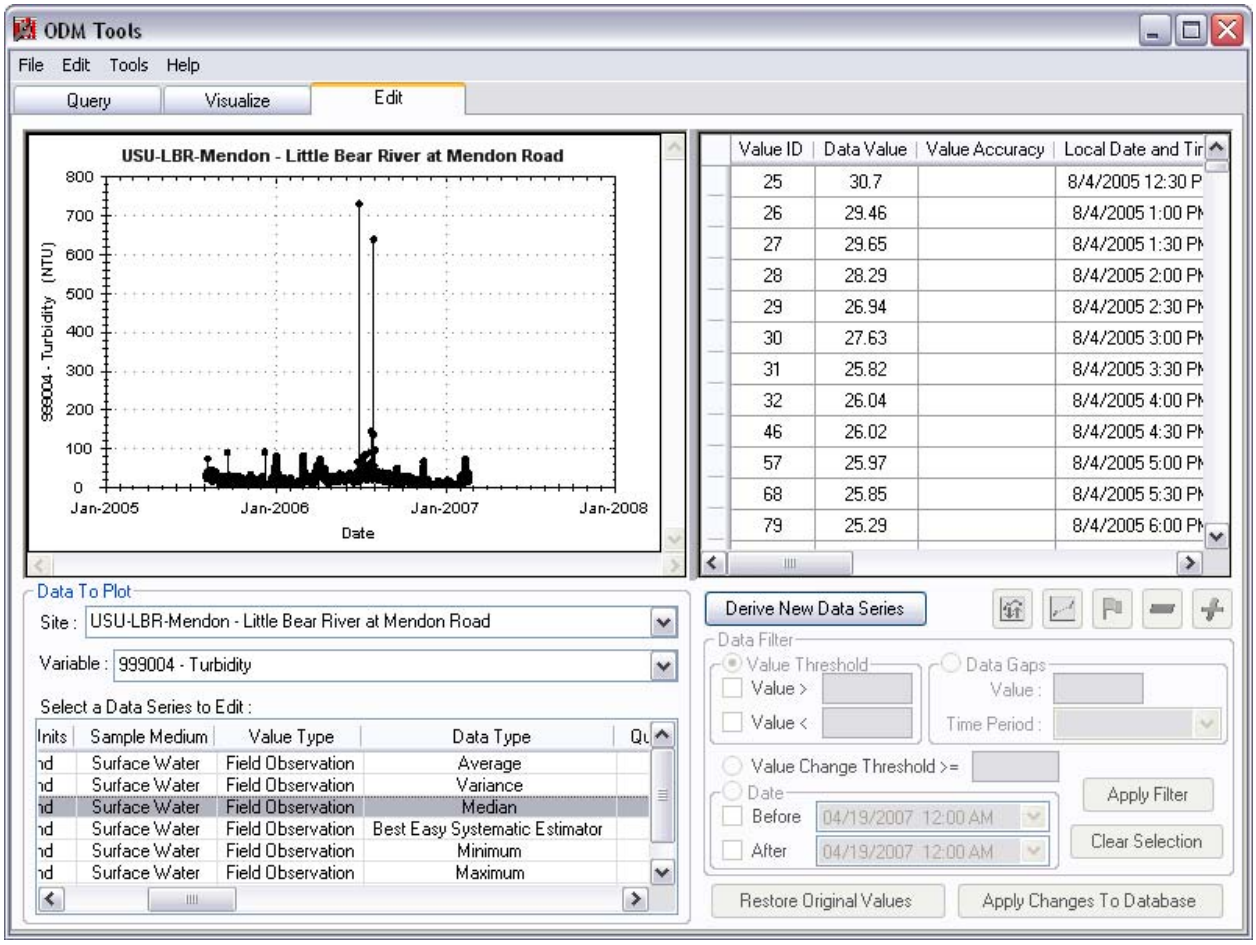
ODM Tools does not allow raw data series to be edited. In order to use all of the data editing functionality of ODM Tools to perform quality assurance and quality control for a raw data series, a copy of the data series must first be created. All data editing is then performed on the copy. Within ODM, raw data series are specified with a Quality Control Level of 0, and quality controlled data series are specified with a Quality Control Level of 1. For more details about how data versioning and editing are managed within ODM, refer to Appendix C of this document. The following steps illustrate how to create a Quality Control Level 1 data series from a Quality Control Level 0 data series so that the editing functionality of ODM Tools can be used.

1. In ODM Tools make the Edit tab active by clicking on it at the top of the application. Your screen will look like the following:

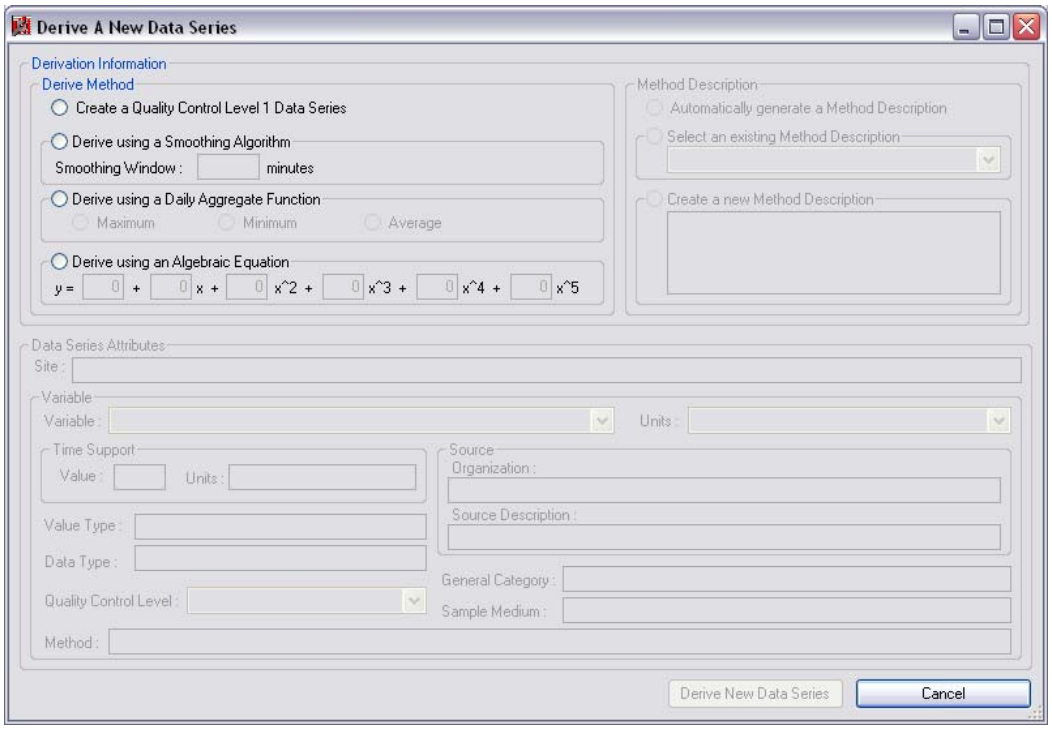


2. Select the raw data series for which you want to create a quality controlled data series by selecting a site from the Site drop down menu, a variable from the Variable drop down menu and a raw (Quality Control Level = 0) data series from the list of available data

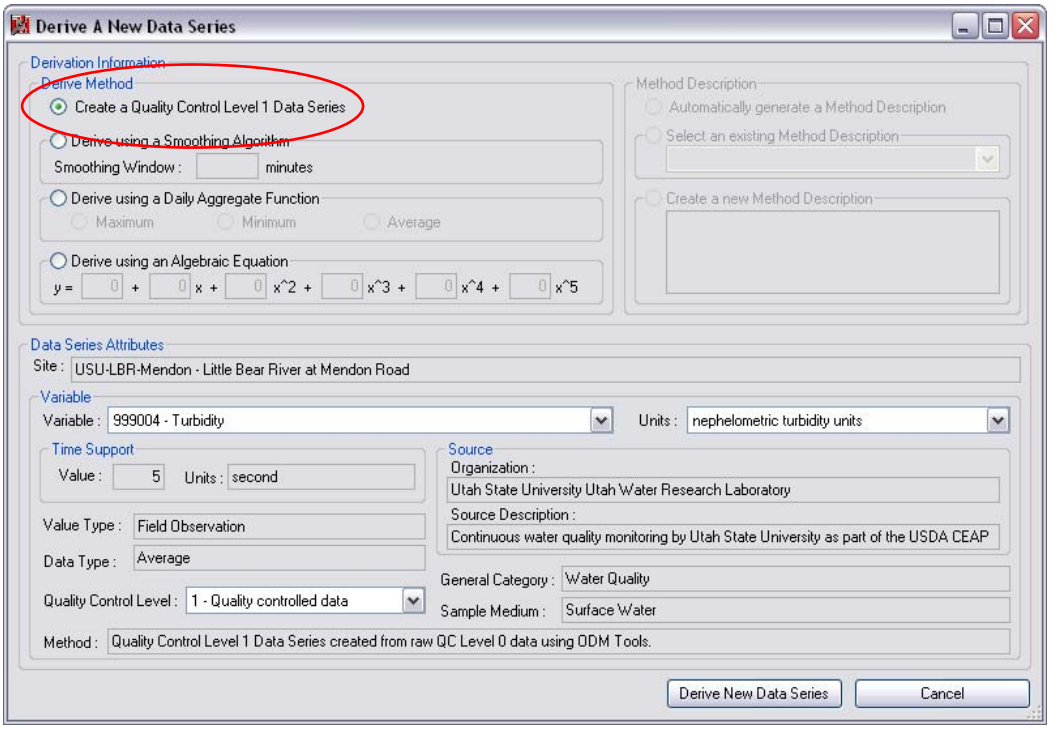
series for that site/variable combination. When you click on a data series in the list, the plot and table on the Edit tab will automatically be populated with the data for the data series that you have selected. For large data series, this can take some time. Your screen should look something like the following:



3. Now that you have selected a data series, you will notice that the Derive New Data Series button, which is located just below the table view of the data, is activated. Also, notice that since you have selected a QualityControlLevel 0 data series, none of the data editing tools are available. Click on the Derive New Data Series Button. The following window will open:



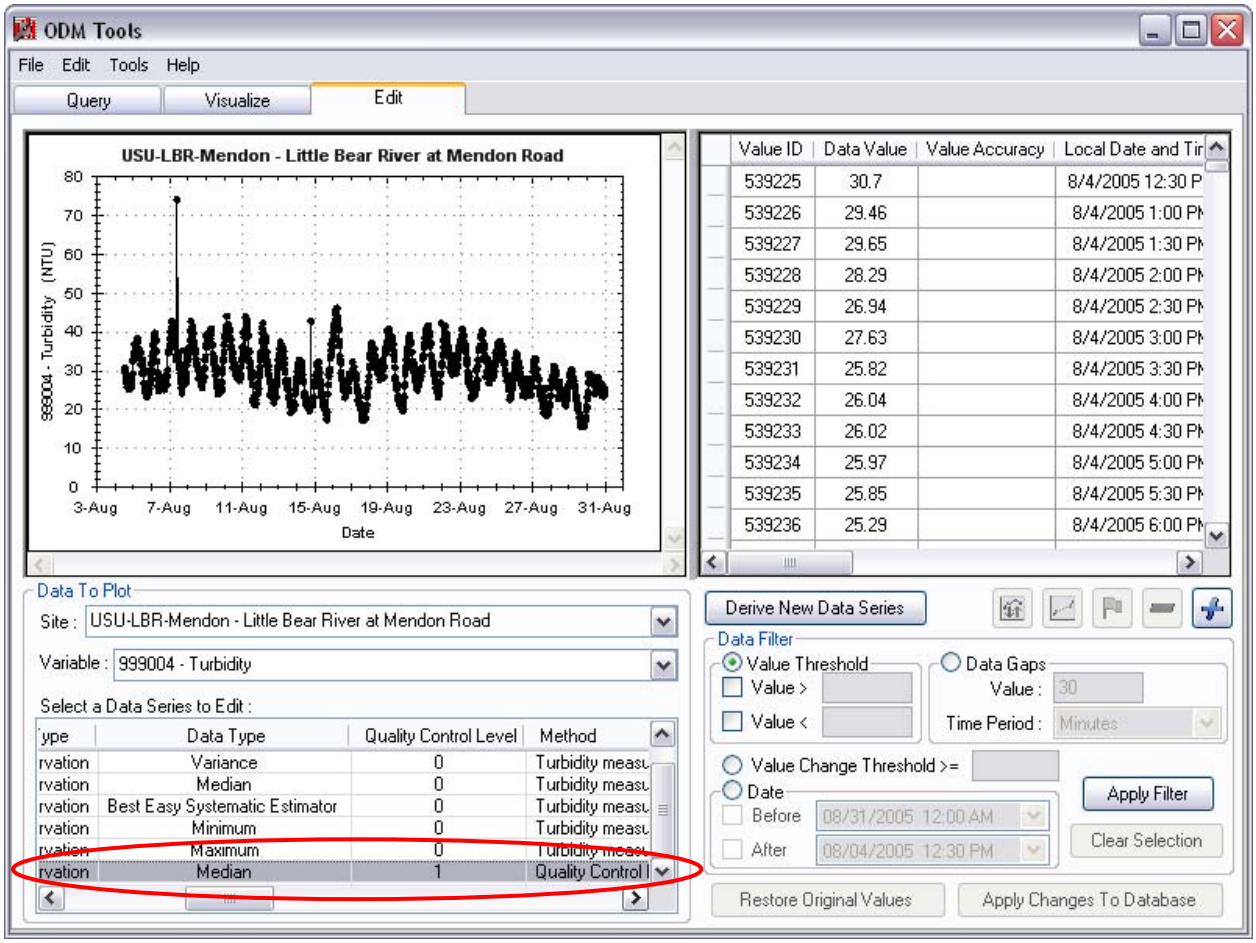
4. Click on the radio button next to "Create a Quality Control Level 1 Data Series" under the Derive Method group. This will activate the Data Series Attributes section at the bottom of the form.



5. You will notice that all of the Data Series Attributes have been selected for you. This is because you are creating an exact copy of your Quality Control Level 0 data series for editing, and the only thing that changes is the Quality Control Level (notice that a Quality Control Level of 1 has been selected). Click the Derive New Data Series button at the bottom of the form to create the new data series.

NOTE: For data series with a large number of data values, this may take some time.

6. When creation of the new data series is complete, you will be returned to the main ODM Tools Edit tab. You will notice that a new data series appears in the list at the bottom of the Edit tab, and that it is selected and shown in the plot and table views.



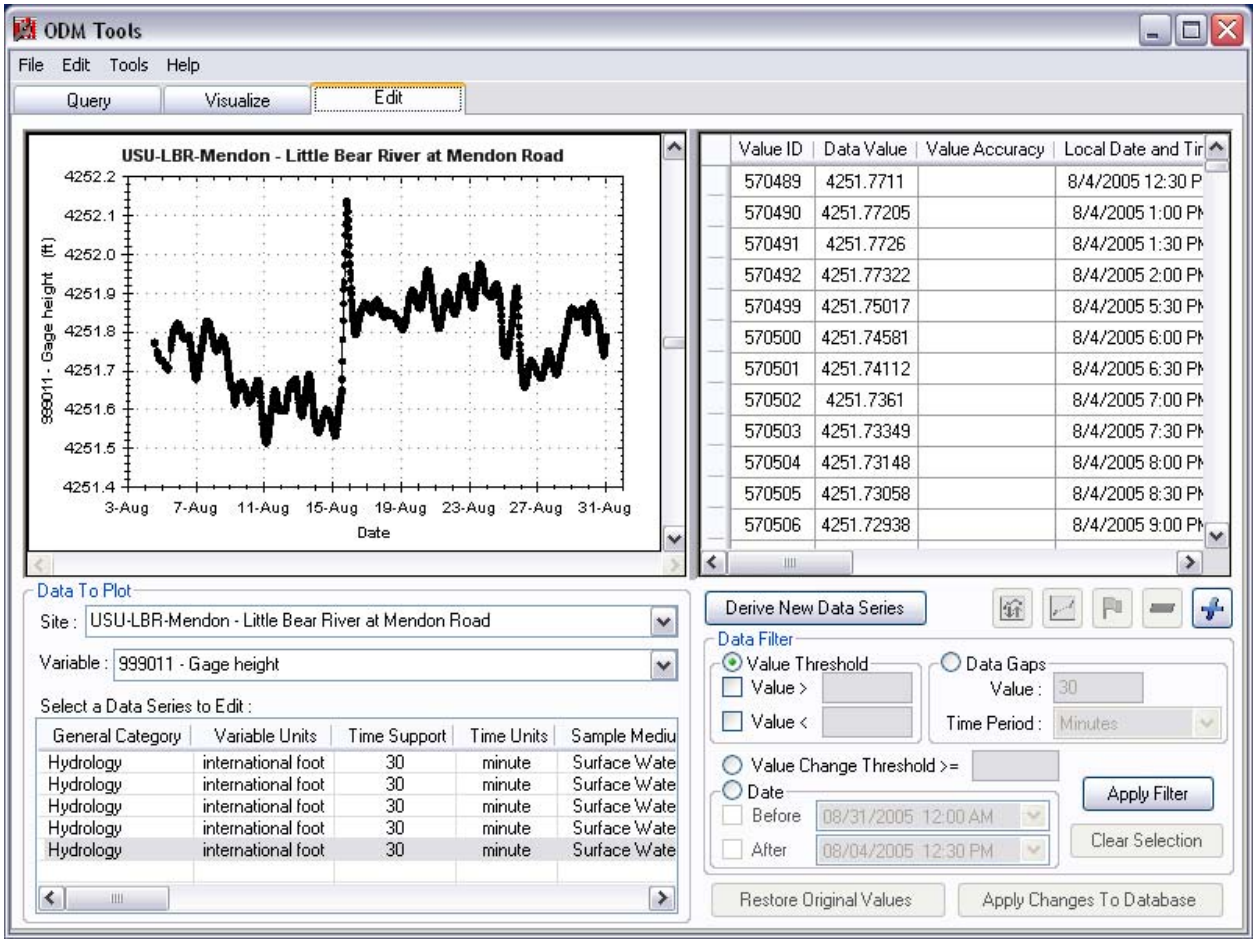
NOTE: This process only creates a Quality Control Level 1 copy of your data series. It does not do any editing or filtering of the data values. Once you have completed the above steps and have created a Quality Control Level 1 data series, you can then use all of the data editing functionality of ODM Tools to perform any data edits required (see the sections on editing data series below).

NOTE: Only one Quality Control Level 1 data series can exist in the database for each Quality Control Level 0 data series. See Appendix C for details.

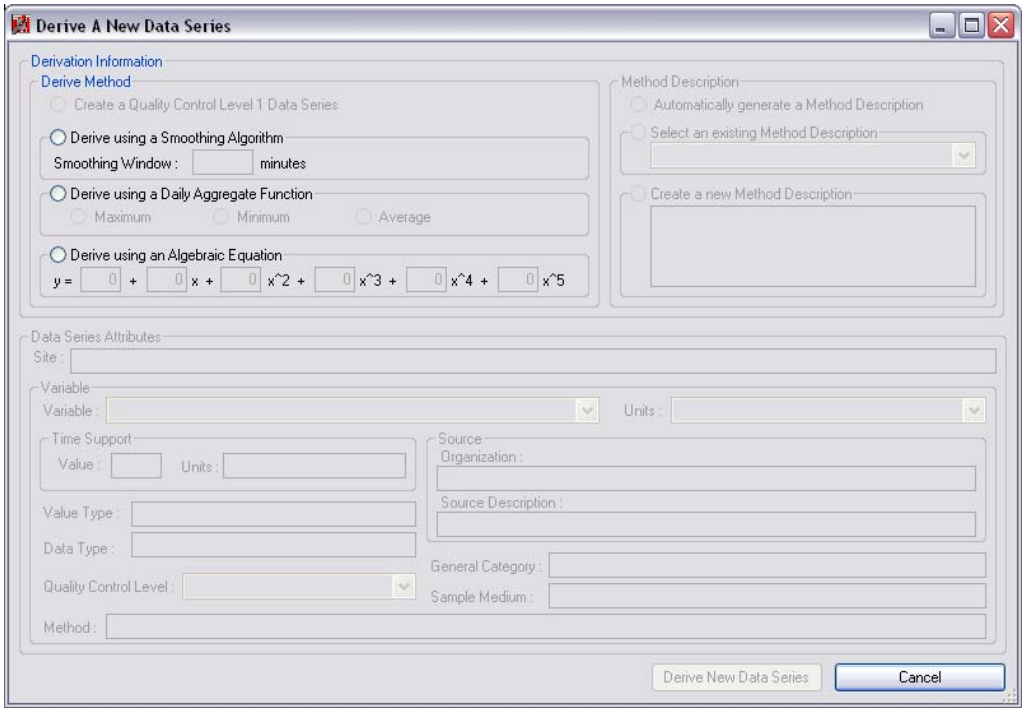
6.1.2 Creating a Smoothed Data Series

ODM Tools allows you to create a new data series by applying a smoothing algorithm to an existing data series already stored within the ODM. Smoothing is useful where the scatter in the data blurs a fundamental underlying pattern. For example, smoothing can be used to provide an estimate when artificial variability has impacted the recorded values. This functionality is only available for data series that are continuous and that do not have large data gaps. The following steps illustrate how to create a smoothed data series.

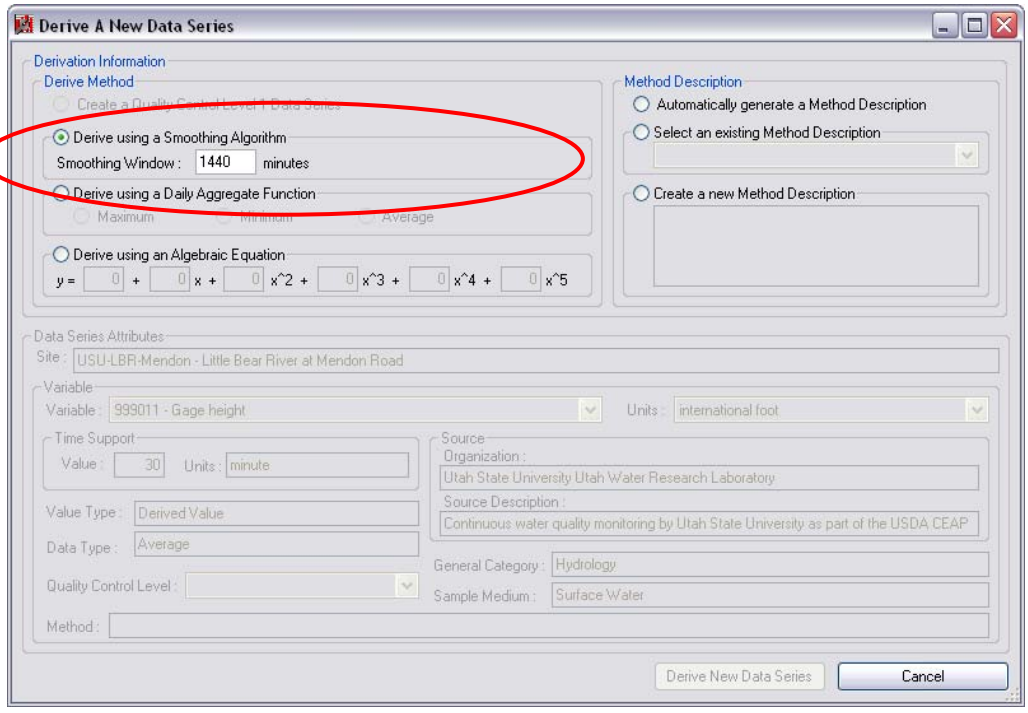
1. Select a data series for which you want to create a smoothed data series by selecting a site from the Site drop down menu, a variable from the Variable drop down menu and a data series from the list of available data series for that site/variable combination. When you click on a data series in the list, the plot and table on the Edit tab will automatically be populated with the data for the series that you have selected. For large data series, this can take some time. Your screen should look something like the following:



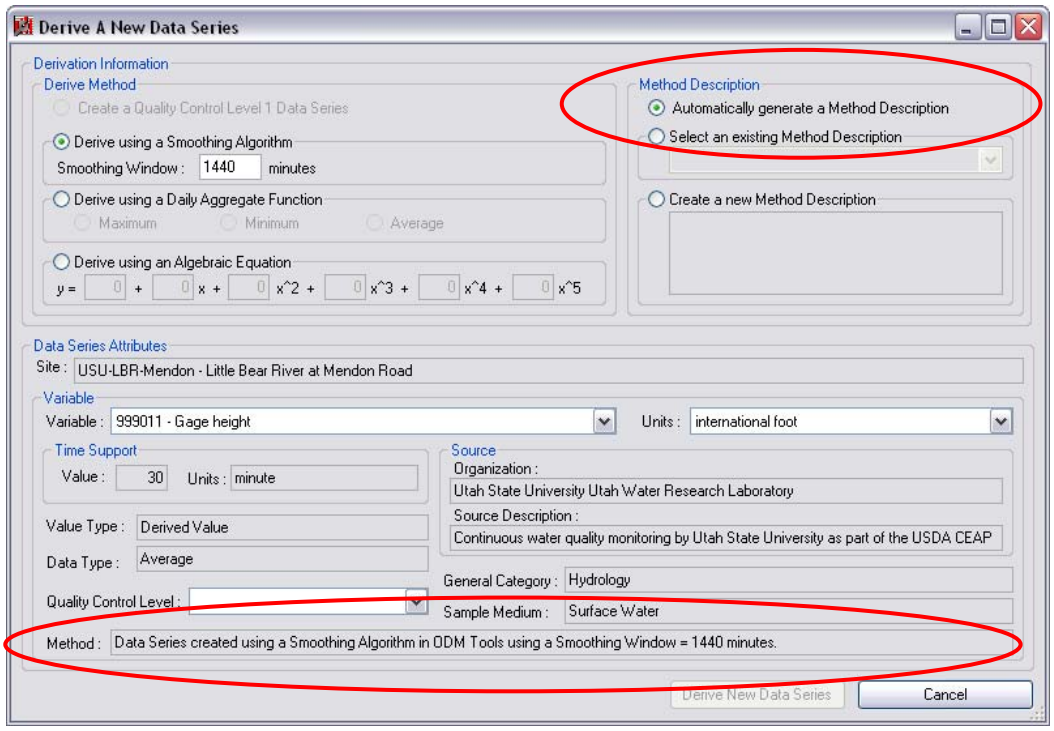
2. Click the Derive New Data Series button. The following window will open:



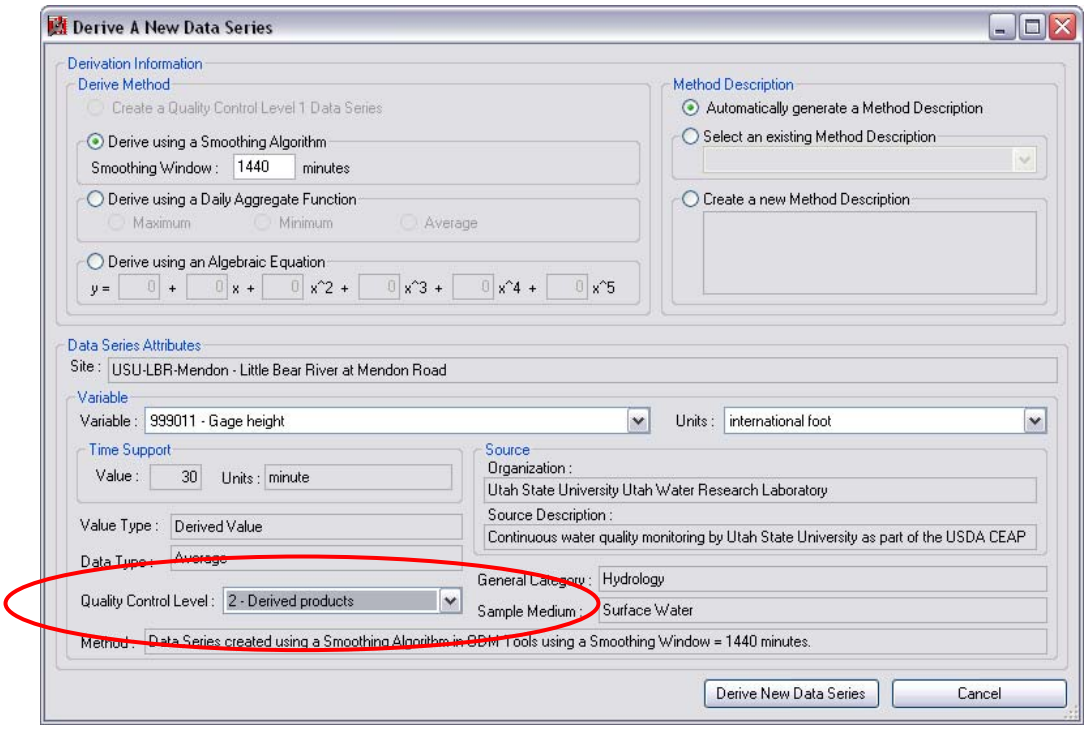
3. Click the radio button next to Derive using a Smoothing Algorithm. Enter a smoothing window for the smoothing algorithm. The smoothing algorithm is based on local linear regression, with weights given to each data value that determine their importance in the local regression. The smoothing window is the time period over which the weights are applied (i.e., the weight for data values outside of the window is 0). Your screen should now look like the following:



4. You will notice that the Method Description group is now activated on the right hand side of the Derive New Data Series form. When deriving data series, you can choose one of three options for specifying the method that will be associated with the new data series that you are creating: 1) let ODM Tools automatically generate a method description, 2) select a method description from the method descriptions already in your ODM database, and 3) create a new method description by typing it into a text box on the form. For this example, we will use the method description that ODM Tools automatically generates. Click the radio button next to Automatically generate a Method Descrtiption. You will now notice that the Data Series Attributes group has been activated on the form and that a method description has been filled in the Method text box at the bottom of the form. Your screen should look something like the following:



5. You will notice that most of the attributes within the Data Series Attributes section have been filled in for you. You will notice, however, that the Quality Control Level may not be filled in. This will only be the case if you are deriving your aggregate data series from a data series with Quality Control Level of 1 or greater. If you are deriving from a Quality Control Level 0 data series, the resulting data series will also be Quality Control Level 0 (see Appendix C for details). If it has not already been done for you, select the appropriate Quality Control Level for your derived data series. In the following figure, a Quality Control Level of 2 has been selected because the new data series is being derived from a Quality Control Level 1 data series:



6. You have now specified all of the inputs needed to derive the new smoothed data series, and you will notice that the Derive New Data Series button has now been activated. Click on it to derive your new smoothed data series.

NOTE: For data series with a large number of data values, this may take some time.

7. When creation of the new data series is complete, you will be returned to the main ODM Tools Edit tab. You will notice that a new data series appears in the list at the bottom of the Edit tab, and that it is selected and shown in the plot and table views.

NOTE: The smoothing algorithm employed in ODM Tools is a LOWESS algorithm that uses locally weighted least squares regression with bi-square weights. The code for this algorithm was adapted from code contributed by Brent Aulenbach of the United States Geological Survey and uses the methods of Cleveland and McGill (1984) and Cleveland (1979).

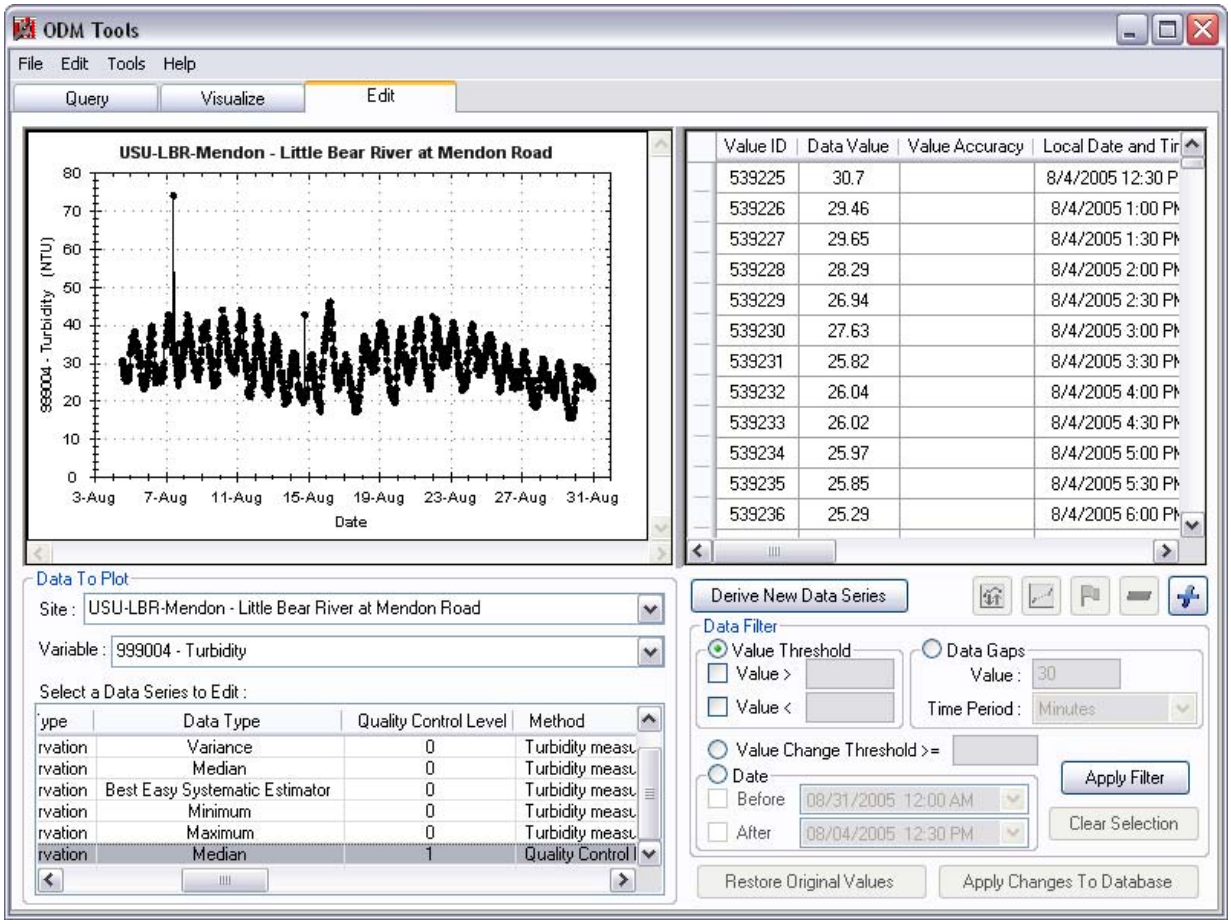
6.1.3 Creating a Daily Aggregate Data Series

Within ODM Tools, you can create a derived data series by generating daily minimum, maximum, or average value statistics from a data series containing more frequent observations already stored within the ODM. An example of this is computing daily average discharge values for a stream gage that measures discharge continuously on a 15-minute interval. This functionality is available for all data series, but is particularly useful for data series that have regular spacing between data values less than one day. ODM Tools calculates aggregate values for each day within the data series and assigns the aggregate value (the minimum, maximum, or average) to midnight at the beginning of the day in which the data values occurred. This is

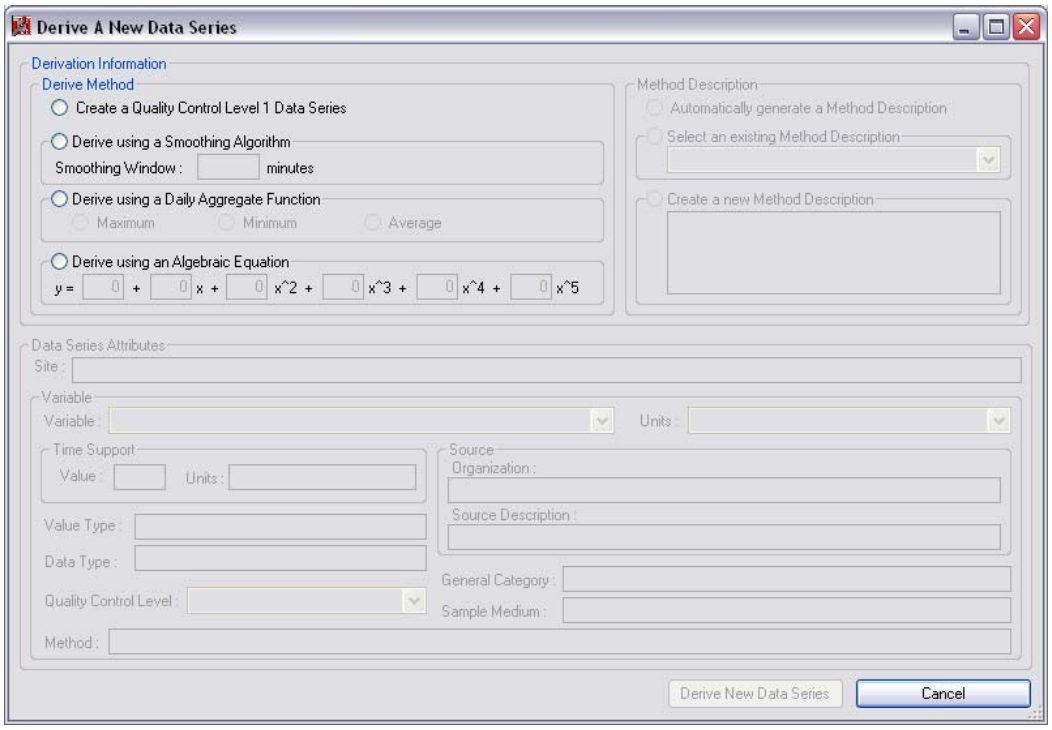
consistent with the beginning of interval reporting convention found in the ODM Design Specifications Document.

The following steps illustrate how to create a daily average data series from a data series with more frequent observations.

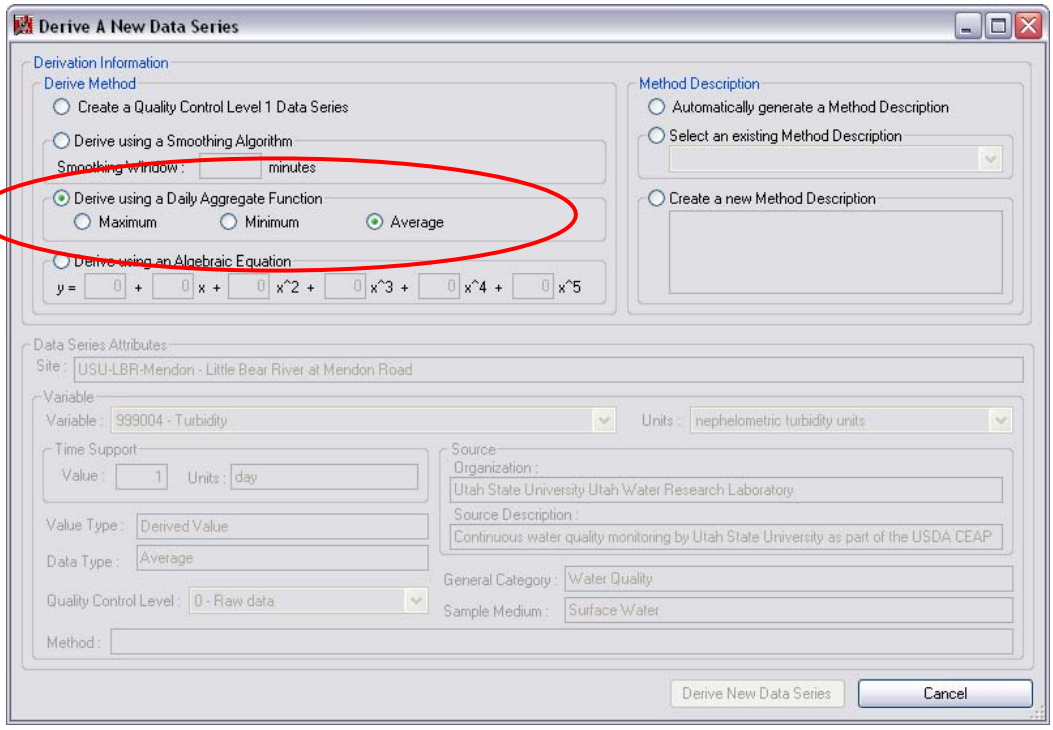
1. Select the data series for which you want to create a daily aggregate data series by selecting a site from the Site drop down menu, a variable from the Variable drop down menu and a data series from the list of available data series for that site/variable combination. When you click on a data series in the list, the plot and table on the Edit tab will automatically be populated with the data for the data series that you have selected. For large data series, this can take some time. Your screen should look something like the following:



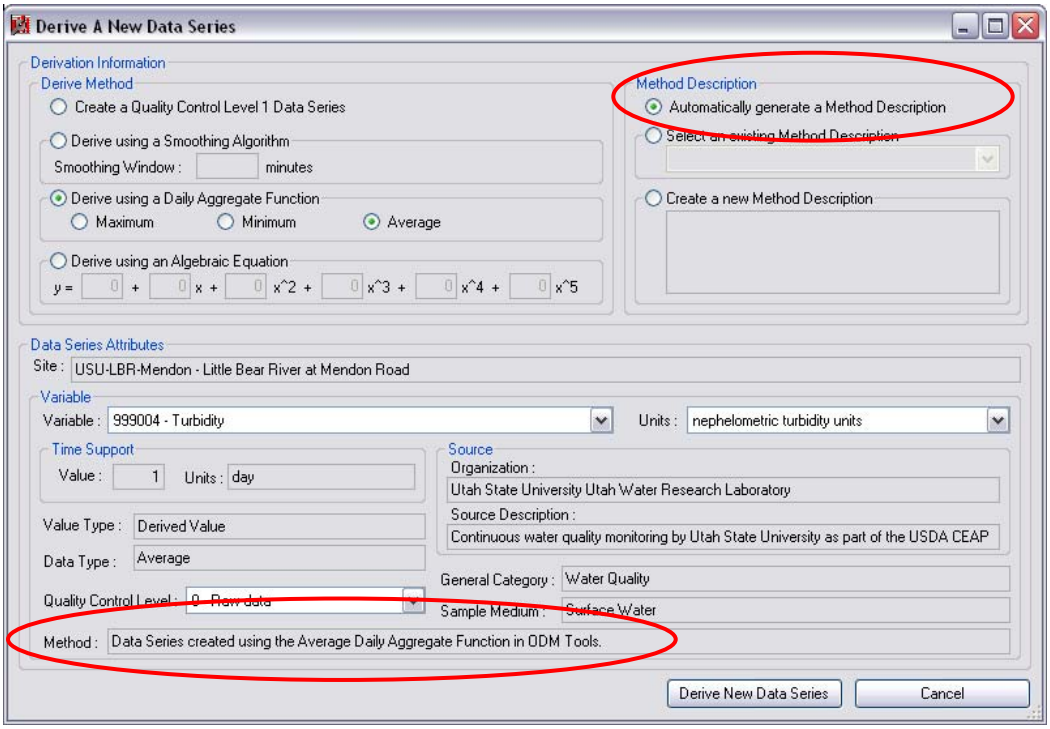
2. Click the Derive New Data Series button. The following window will open:



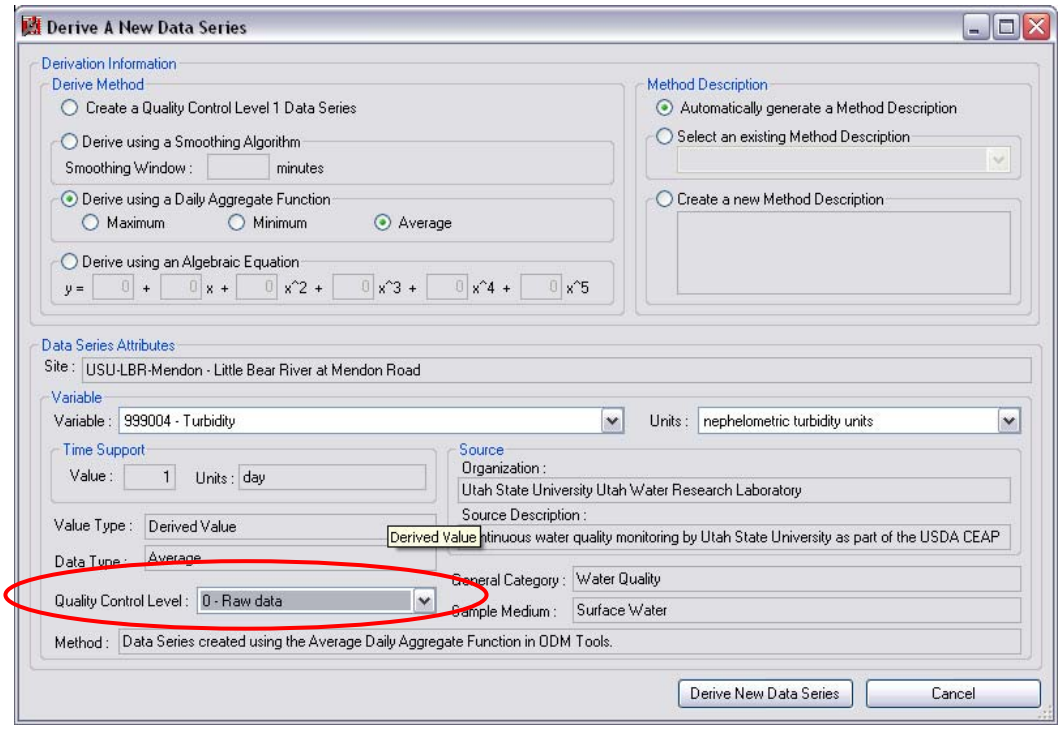
3. Click the radio button next to Derive using a Daily Aggregate Function. Select Average from the Daily Aggregate Function options by clicking on the radio button if it is not already selected. Your screen should now look like the following:



4. You will notice that the Method Description group is now activated on the right hand side of the Derive New Data Series form. When deriving aggregate data series, you can choose one of three options for specifying the method that will be associated with the new data series that you are creating: 1) let ODM Tools automatically generate a method description, 2) select a method description from the method descriptions already in your ODM database, and 3) create a new method description by typing it into a text box on the form. For this example, we will use the method description that ODM Tools automatically generates. Click the radio button next to Automatically generate a Method Description. You will now notice that the Data Series Attributes group has been activated on the form and that a method description has been filled in the Method text box at the bottom of the form. Your screen should look something like the following:



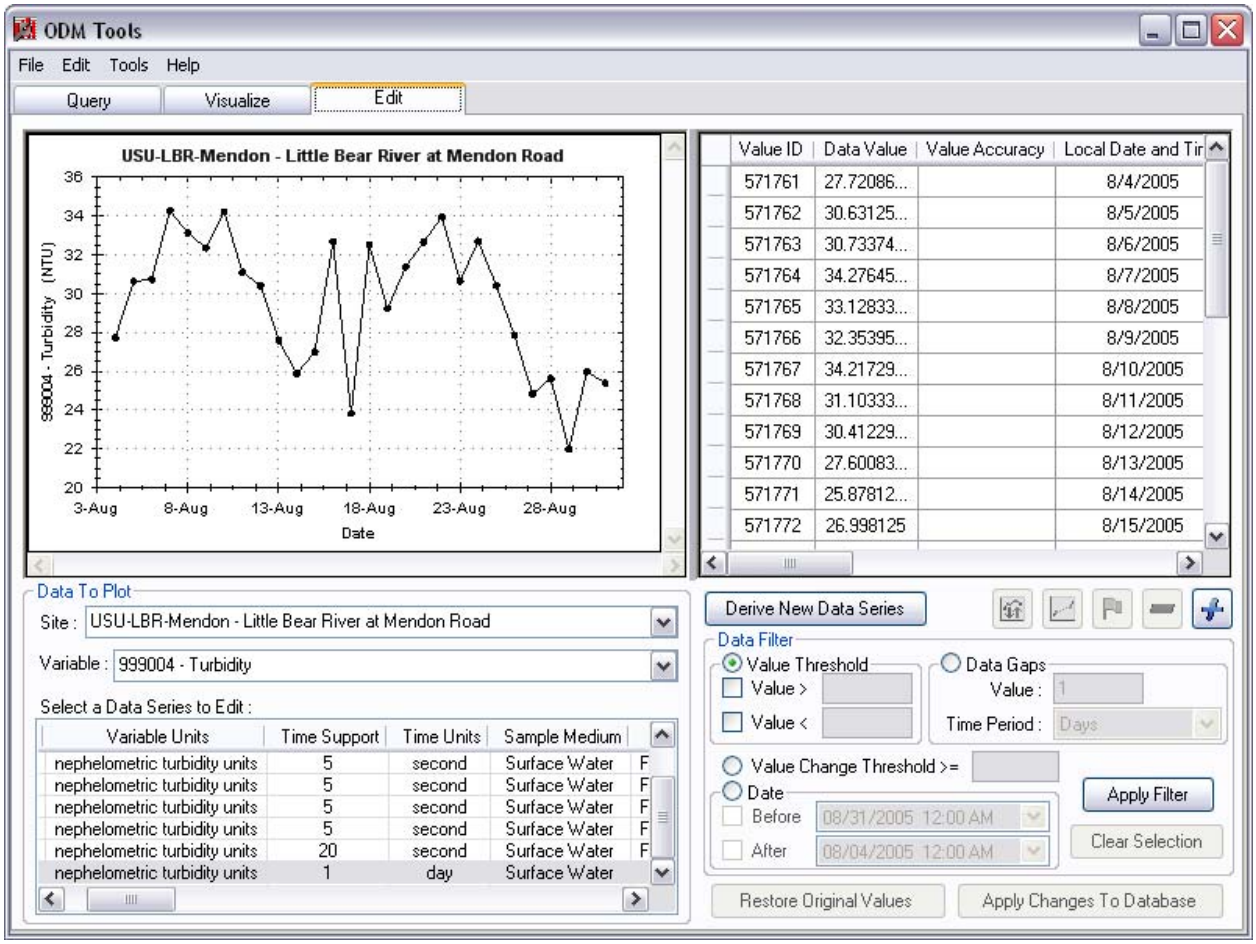
5. You will notice that most of the attributes within the Data Series Attributes section have been filled in for you. Because you are creating a daily aggregate data series, most of the attributes for your data series will not change. You will notice, however, that the Time Support attribute is set to one day, and that the Quality Control Level may not be filled in. This will only be the case if you are deriving your aggregate data series from a data series with Quality Control Level of 1 or greater. If you are deriving from a Quality Control Level 0 data series, the resulting data series will also be Quality Control Level 0 (see Appendix C for details). If it has not already been done for you, select the appropriate Quality Control Level for your derived data series. In the following figure, a Quality Control Level of 0 is selected because the new data series is being derived from a Quality Control Level 0 data series.



6. You have now specified all of the inputs needed to derive the new aggregate data series, and you will notice that the Derive New Data Series button has now been activated. Click on it to derive your new daily average data series.

NOTE: For data series with a large number of data values, this may take some time.

7. When creation of the new data series is complete, you will be returned to the main ODM Tools Edit tab. You will notice that a new data series appears in the list at the bottom of the Edit tab, and that it is selected and shown in the plot and table views.



6.1.4 Creating a Derived Data Series Using an Algebraic Function

ODM Tools supports creation of a derived data series by performing an algebraic calculation on an existing data series already stored within the ODM. An example of this is computing a discharge data series based on a stage data series using a site specific rating curve. This functionality applies the same algebraic conversion to each data value within the input data series, resulting in an output data series with the same number of data values and same spacing between data values. ODM Tools currently supports a limited number of algebraic equation formats, each of which must be specified as a 5th degree polynomial:

$$y = a + bx + cx^2 + dx^3 + ex^4 + fx^5$$

Where:

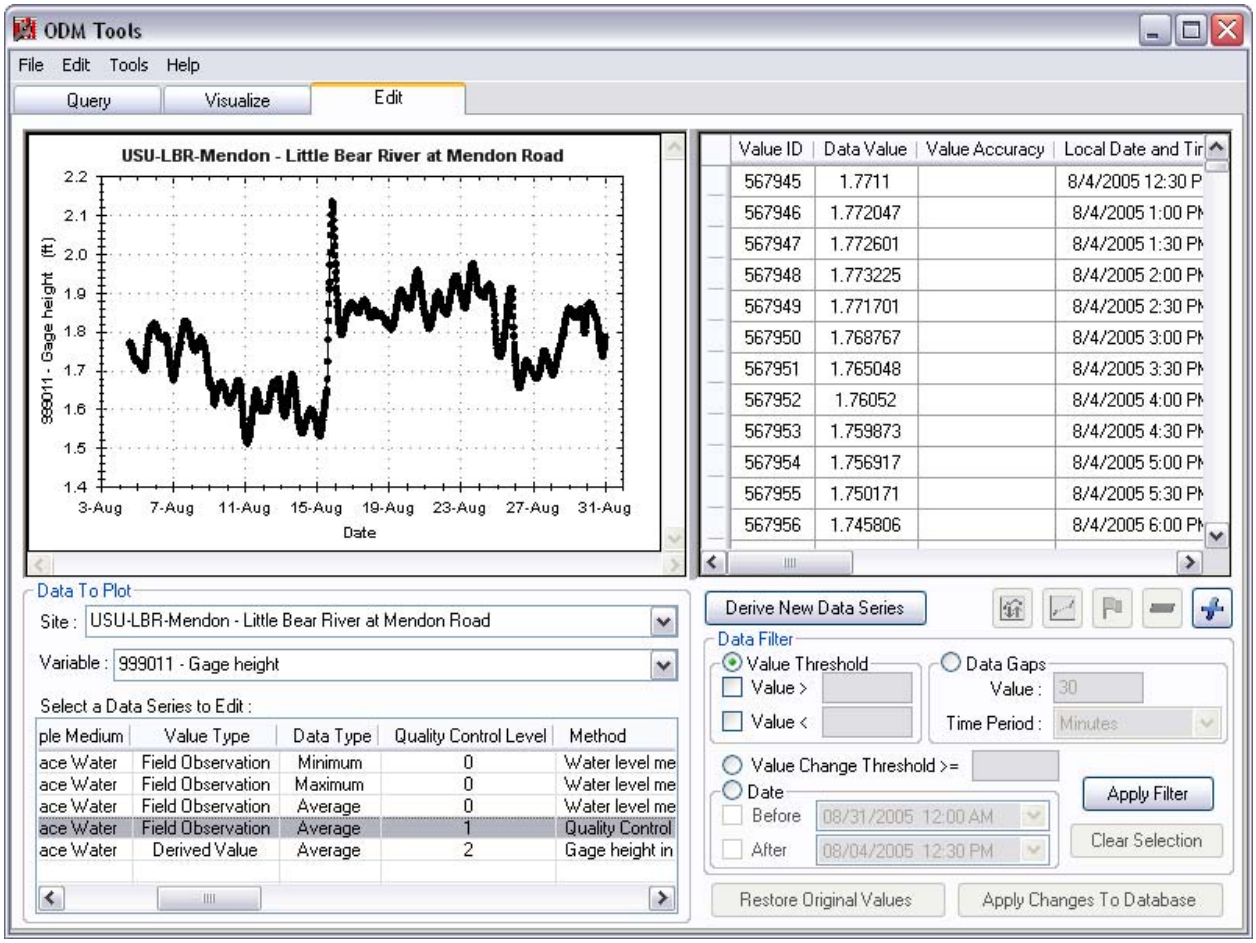
y = the new derived quantity
x = the derived from quantity
a, b, c, d, e, and f = coefficients that can be input by the user

For example, to do a simple linear transformation of a data series, the slope and intercept of the linear equation would be input using the a and b coefficients, and all other coefficients would be set to zero. Algebraic transformations can be used to convert from one variable to another (i.e.,

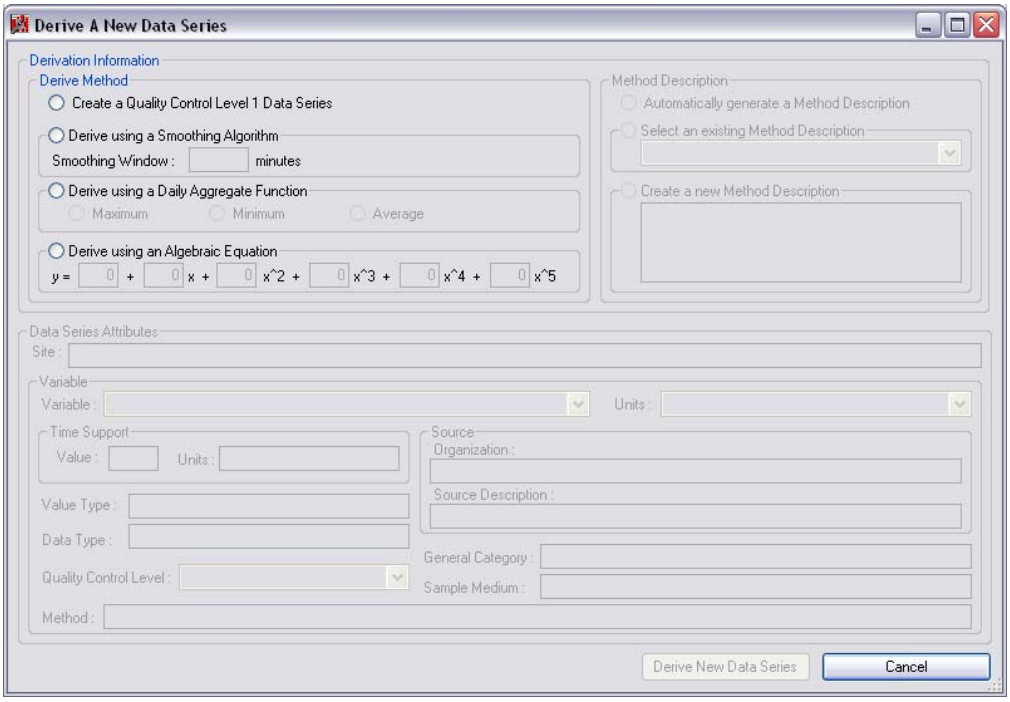
convert stage to discharge), to change the units of the values in a data series (i.e., convert feet to meters), and to adjust the values within a data series (i.e., add a constant to change a datum).

The following steps illustrate how to create a new data series from an existing data series using an algebraic equation. For this example, we will do a very simple adjustment of stage values measured by a stage recording device so that the stage values are expressed as an elevation with reference to a fixed vertical datum rather than an elevation with reference to the location of the water level sensor.

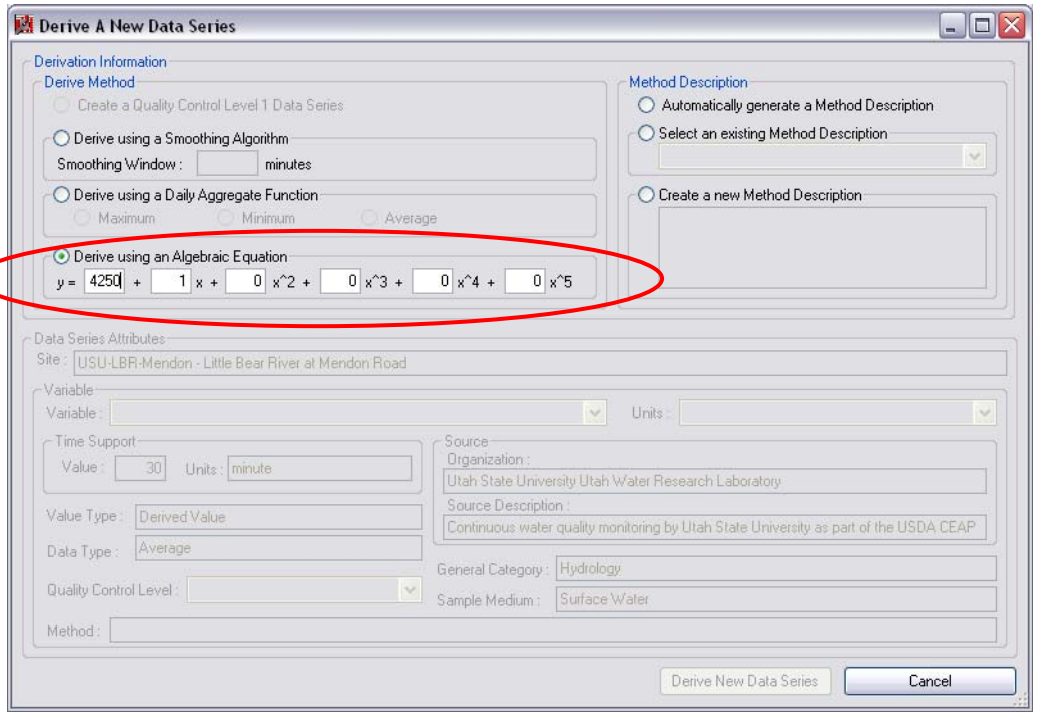
1. Select the data series for which you want to derive the new data series by selecting a site from the Site drop down menu, a variable from the Variable drop down menu and a data series from the list of available data series for that site/variable combination. Your screen should look something like the following:



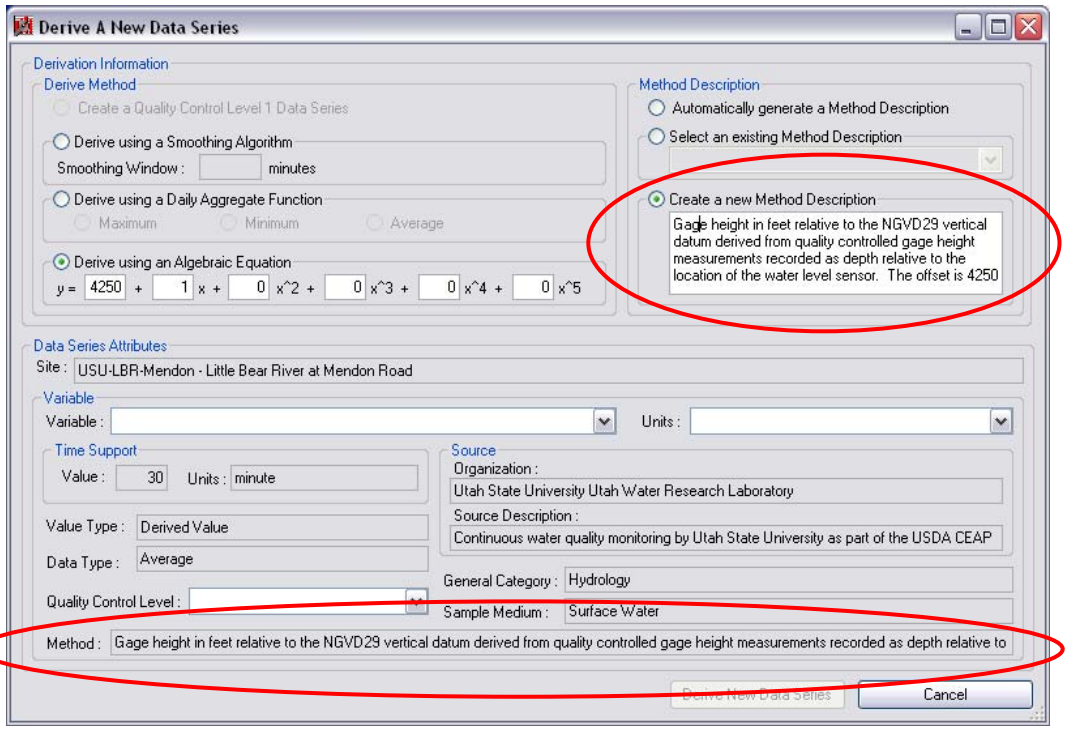
2. Click the Derive New Data Series button. The following window will open:



3. Click the radio button next to Derive using an Algebraic Equation. For this example, we are simply adding an offset to the stage values recorded by the water level sensor so that they will be relative to a known vertical datum (i.e., feet relative to the NGVD29 vertical datum) and not to the fixed location of the level sensor (i.e., feet above the sensor). This offset is equal to 4250 ft, so we will set the *a* coefficient to 4250, the *b* coefficient to 1, and all of the other coefficients to 0. See the following figure:

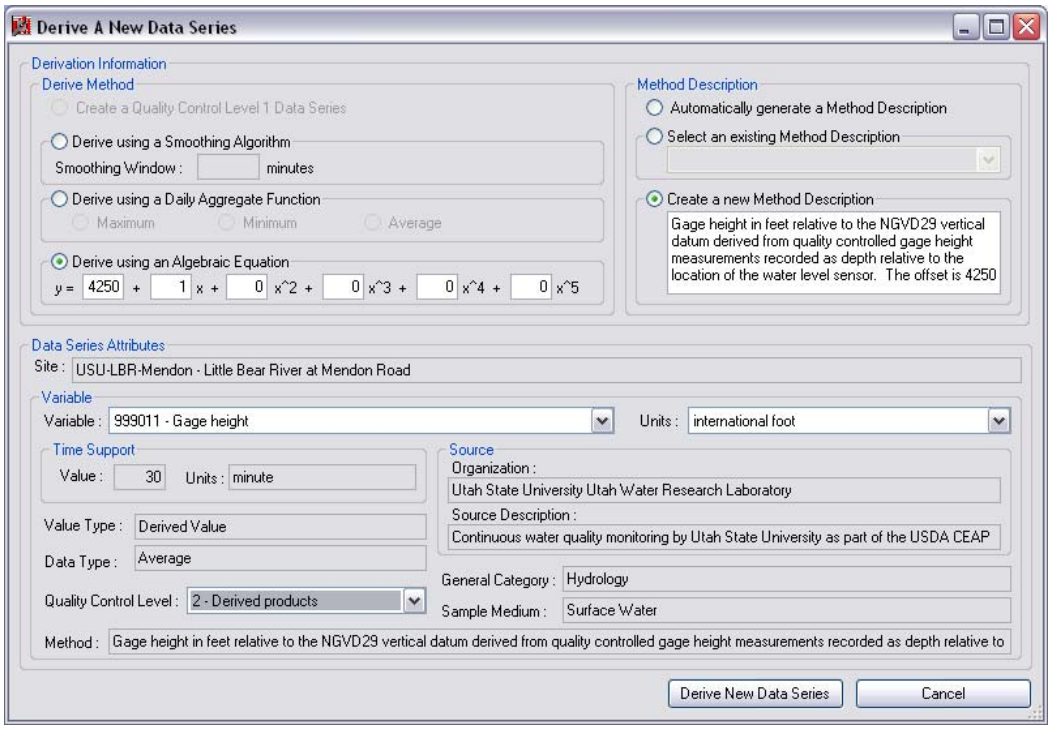


4. You will notice that the Method Description group is now activated. For this example, we will create a new method description that details the conversion that we are performing. This is done by clicking on the radio button next to the Create a new Method Description option and then typing the new method description into the text box. When the method description is complete, the Enter key is pressed. This activates the Data Series Attributes group and fills in the Method text box at the bottom of the form:



NOTE: It is important in this case that we enter a good description of the method used to create the derived data series because all of the other attributes will remain the same.

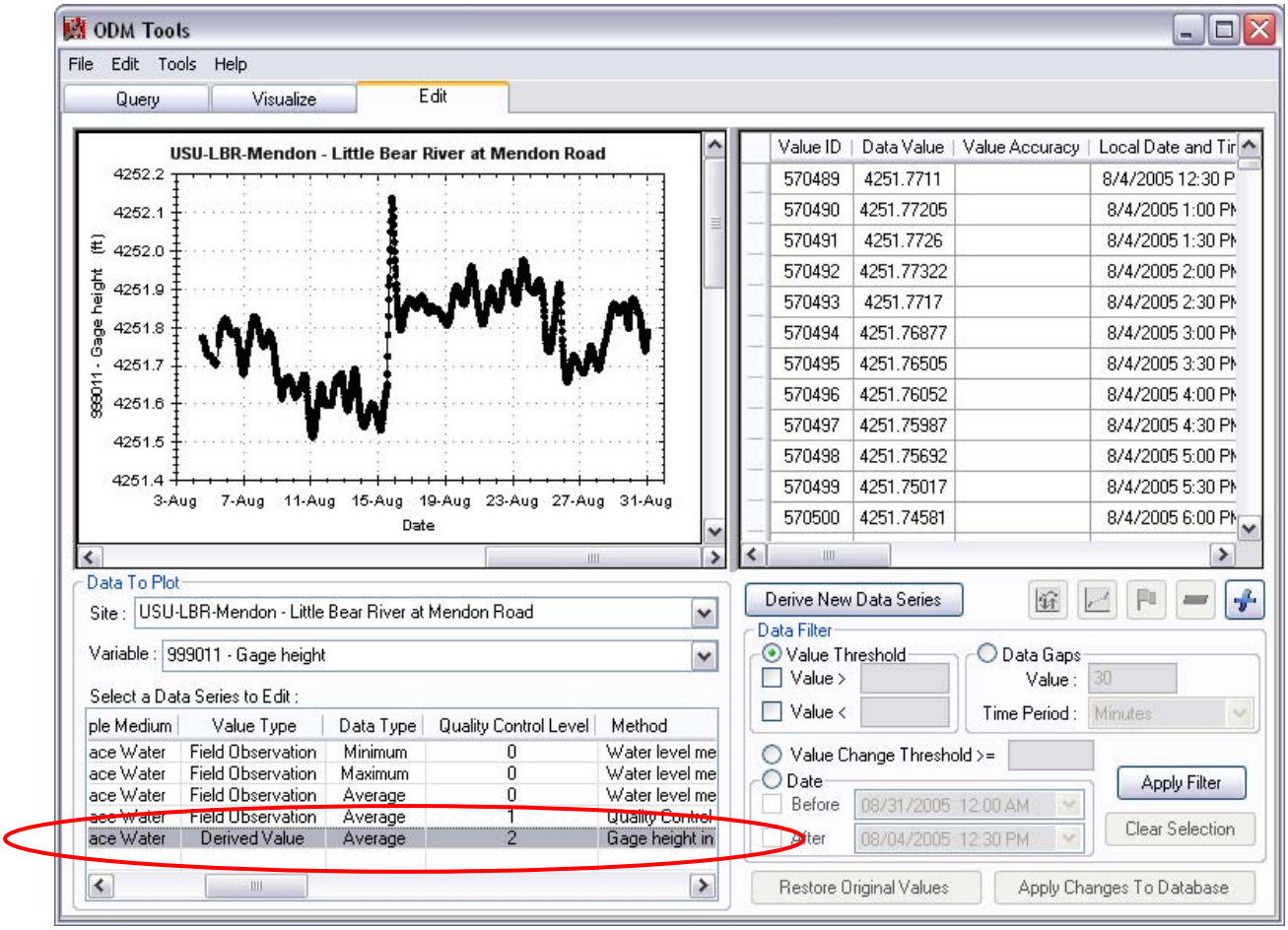
5. With algebraic conversions, it is possible to create data series for which the variable and units may not be consistent with the input data series. Because of this, you will notice that the Variable and Units selection boxes are empty. In addition, you will notice that the Quality Control Level attribute may be blank (unless you are deriving from a level zero data series). Since we are not changing the variable or the units, we will make the appropriate selections from the drop down lists. We will also select a Quality Control Level of 2, indicating that the new data series is a derived data product. See the following figure:



6. You have now specified all of the inputs needed to derive the new derived data series, and you will notice that the Derive New Data Series button has now been activated. Click on it to derive your new daily average data series.

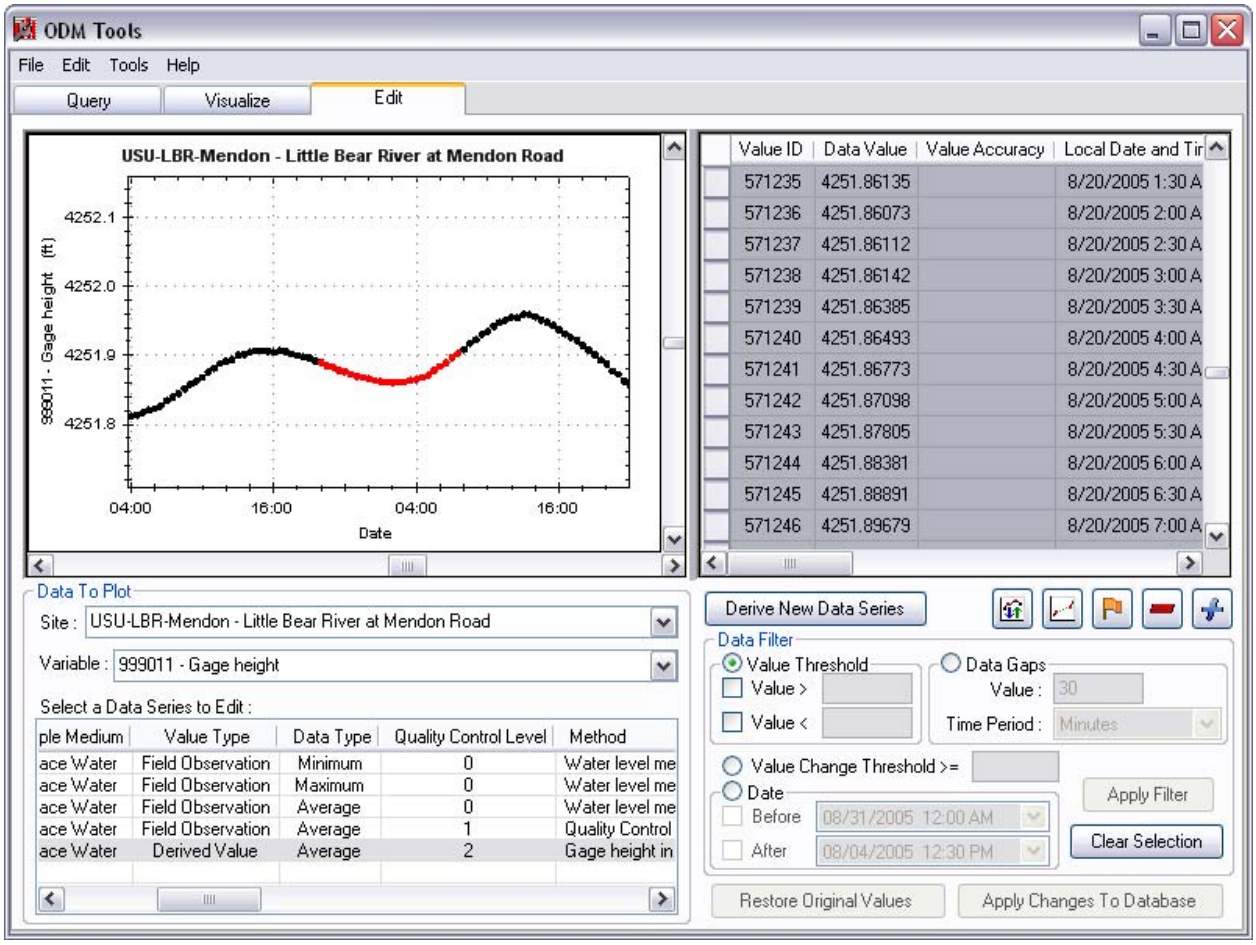
NOTE: For data series with a large number of data values, this may take some time.

7. When creation of the new data series is complete, you will be returned to the main ODM Tools Edit tab. You will notice that a new data series appears in the list at the bottom of the Edit tab, and that it is selected and shown in the plot and table views.



6.2 Editing Existing Data Series

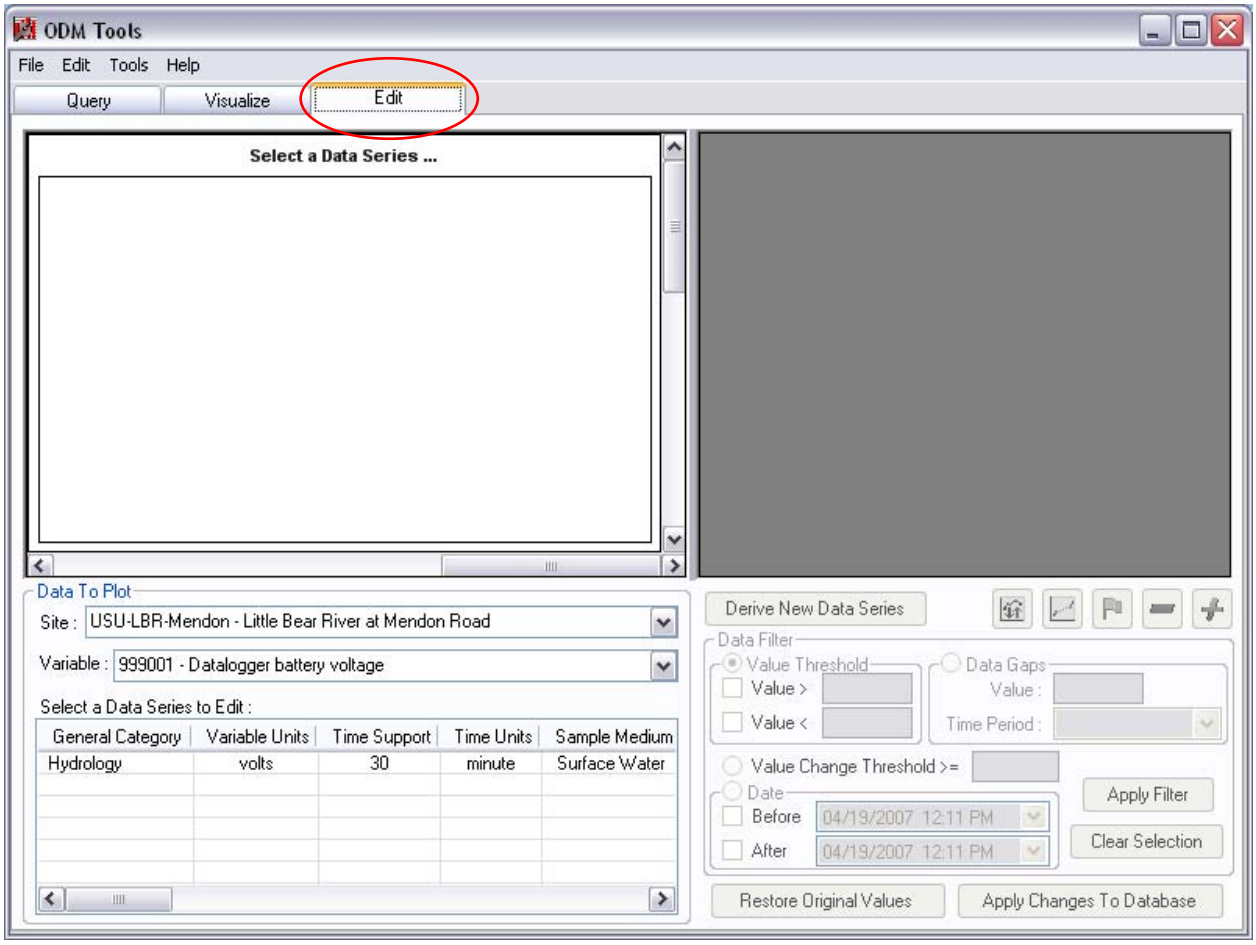
ODM Tools contains a variety of functionality that allows you to edit the data values within a data series. Only data series of Quality Control Level 1 or higher can be edited using ODM Tools. Editing is based on a visual representation of the data in both plot and tabular formats. Data selected for editing are highlighted on both the plot and within the table. All data editing is held in memory until editing for the entire data series is complete, at which point you can choose whether to commit all of the changes you have made to the database and replace the existing data values. The following sections describe the data series editing functionality of ODM Tools. All of this functionality is located on the ODM Tools Edit tab, which is shown in the following figure.



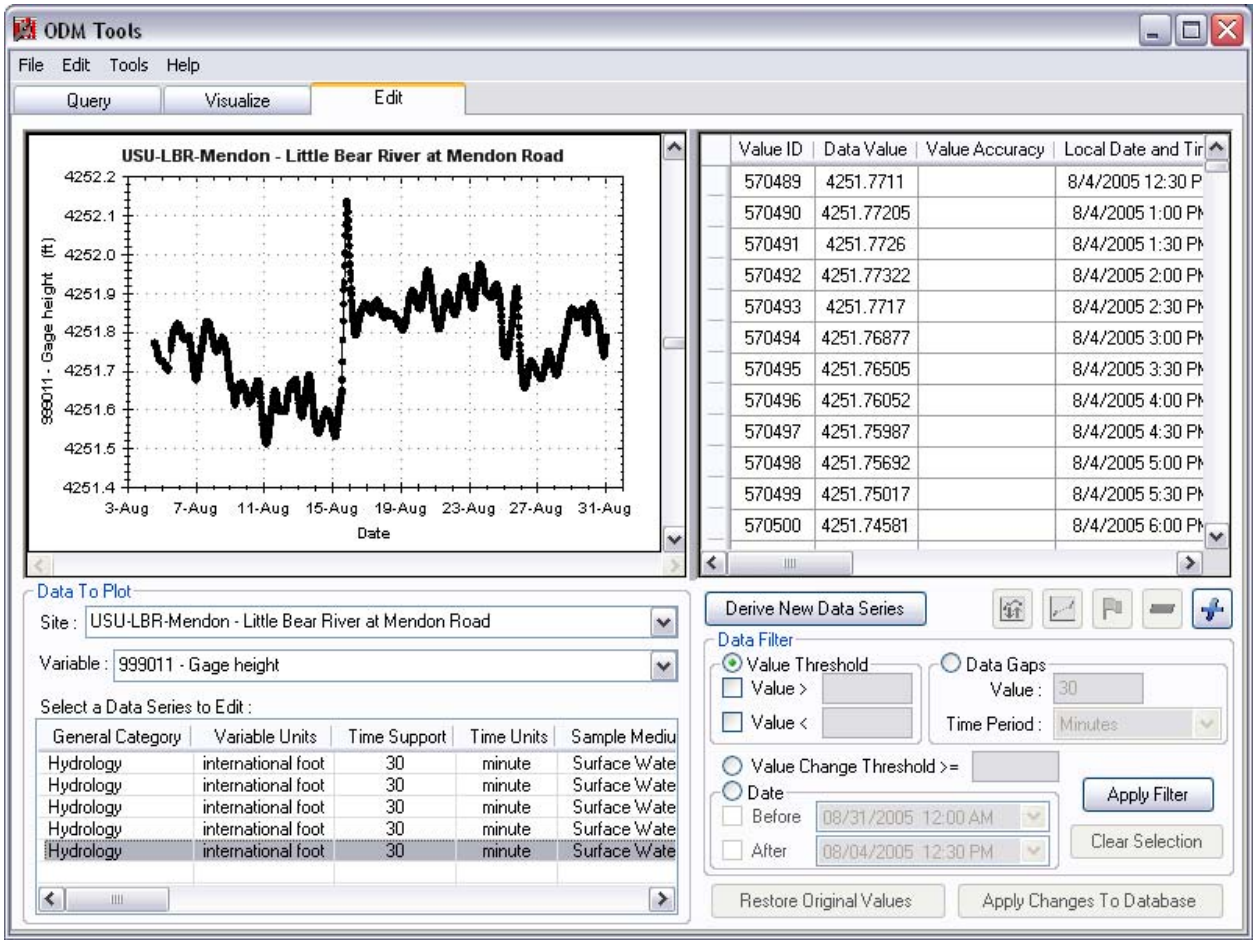
6.2.1 Selecting a Data Series for Editing

Within ODM Tools, a data series must first be selected before it can be edited. Selecting a data series essentially opens it within an editing session inside the ODM Tools Edit tab. Once a data series has been selected, all of the ODM Tools data series editing functionality is activated. Use the following steps to select a data series for editing:

1. If it isn't already, make the Edit tab active by clicking on it at the top of the ODM Tools window. Your screen will look something like the following:



2. Select the data series that you want to edit by selecting a site from the Site drop down menu, a variable from the Variable drop down menu and a data series from the list of available data series for that site/variable combination. When you click on a data series in the list at the bottom of the form, the plot and table on the Edit tab will automatically be populated with the data for the data series that you have selected. For large data series, this can take some time. Your screen should look something like the following:

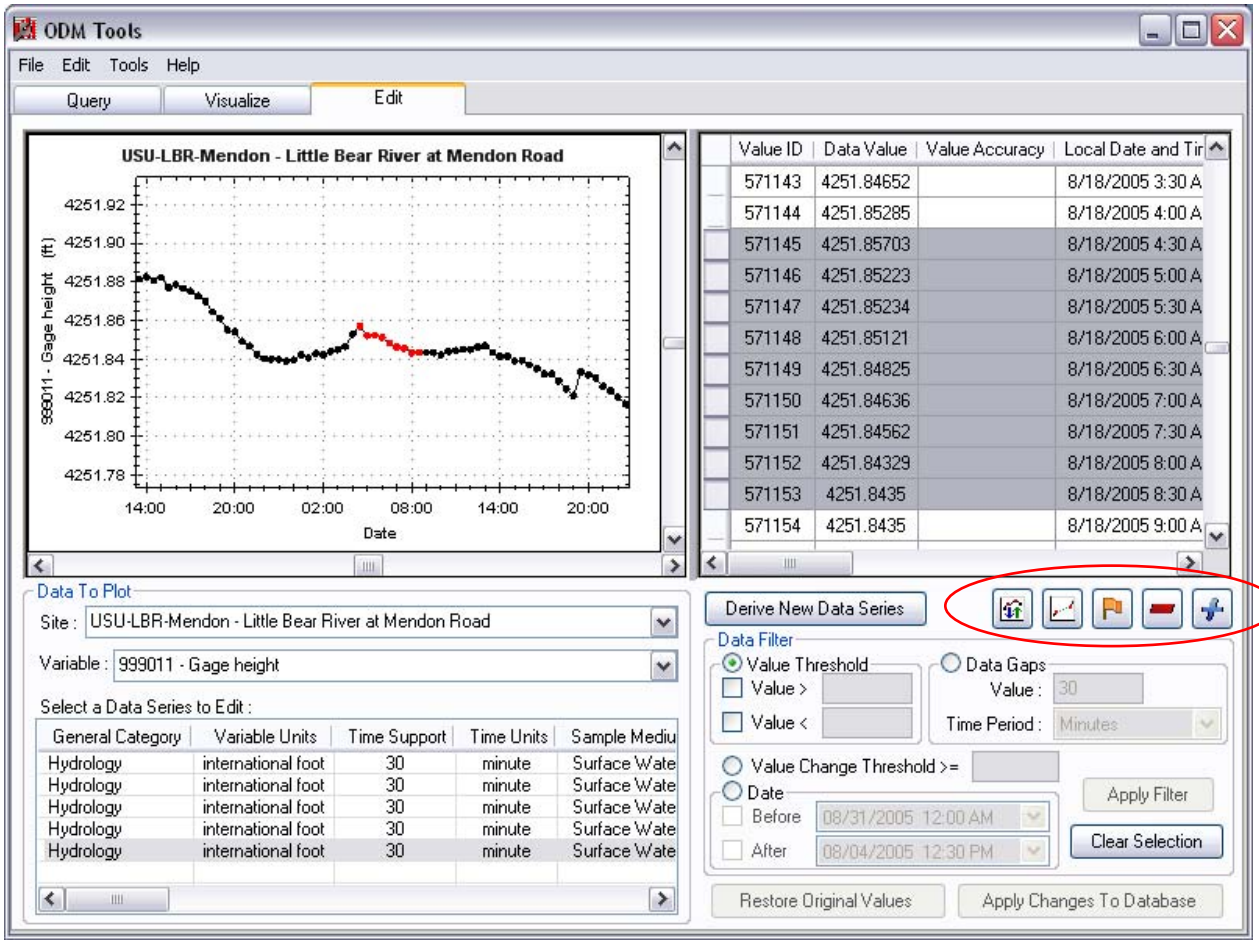


3. Once you have selected your data series and it has been loaded into the plot and table, it is ready to be edited using the ODM Tools data series editing tools.

NOTE: Once selected, you can use the ODM Tools data series editing functionality to edit the selected data series. Any edits you perform on the data series will be held in memory until you choose to commit them to the database (see sections below).

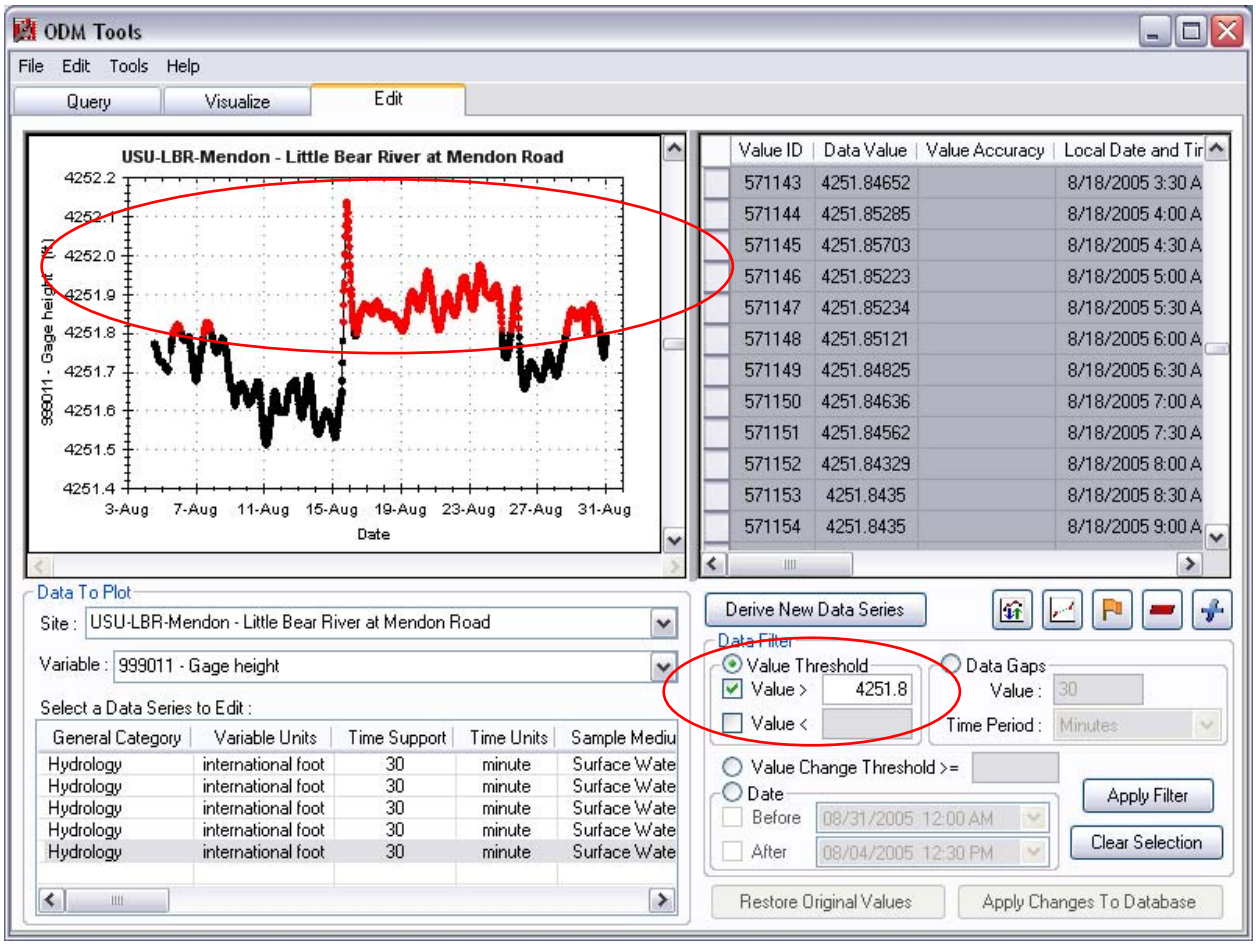
6.2.2 Selecting Data Values to Edit Using the Data Filters

Once you have selected a data series for editing, you can select individual data values within that data series for editing. Data values must be selected before they can be edited. Selecting data values can be done using one of two methods. First, you can select data values for editing by selecting their rows in the table view at the right of the form. This is done by clicking on an individual row or by holding down the left mouse button while dragging over several records to select more than one. You can also use shift and click to select multiple records or control and click to select multiple records that are not continuous. You will notice that when you select a set of records in the table view the corresponding points on the plot view are highlighted and shown in red (see the following figure). You will also notice that when you have selected a set of data values, all of the ODM Tools data editing tools are activated.

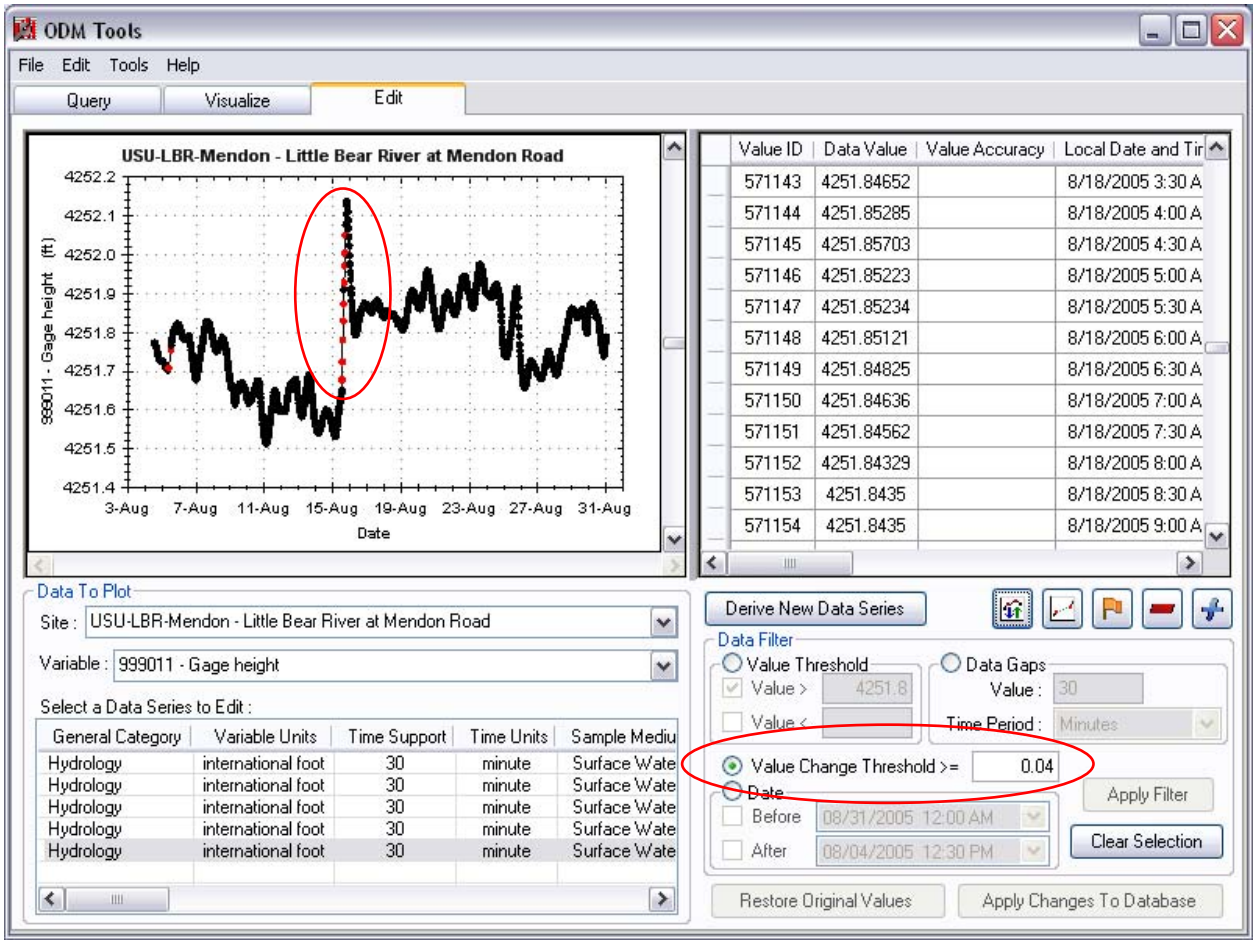


You can also select data values for editing using the ODM Tools data filters. These data filtering tools are located at the bottom right of the Edit tab. These tools are designed to enable you to identify and select data values that may need to be edited. The data filter tools allow you to select, within the selected data series, data values that are above or below a value threshold, data values where the change from one data value to the next is greater than some threshold, data values within a specific date range, and data values where there is a data gap. The following sections describe each of these data filters.

Selecting Data Values Using a Value Threshold – Using the Value Threshold filter, you can select all of the data values within your selected data series that are above a certain value, below a certain value, or between a set of values. This is useful for identifying data values that fall outside of an expected range for a variable, indicating that they may need to be edited. To use this filter, click on the radio button next to the Value Threshold option and enter the threshold values that you wish to use as criteria. You must check the boxes next to the criteria for them to be evaluated. For example, if you wish to select all data values greater than a threshold, you would check the box next to "Value >" and then input your threshold value into the corresponding text box. Click the Apply Filter button to perform the selection. The following figure is an example. You will notice that the data values are selected in both the plot (highlighted in red) and in the table view (highlighted in grey). To clear the selection, click on the Clear Selection button.

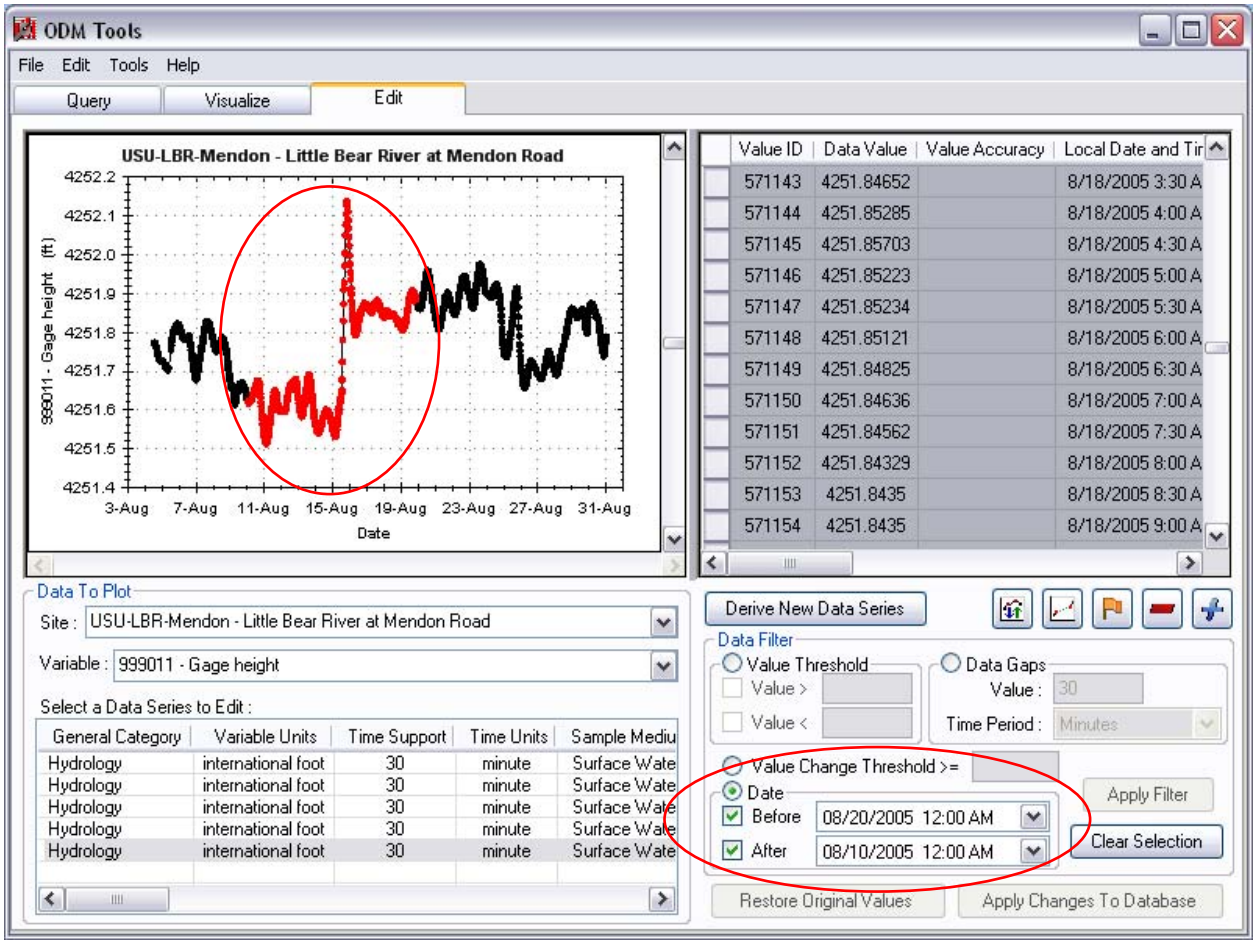


Selecting Data Values Using a Value Change Threshold – The value change threshold filter allows you to select data values where the change from the previous data value to the current data value is greater than some threshold. This is useful in identifying periods where the data values change rapidly, potentially indicating anomalous data values. These anomalies may be natural occurrences, but they may also be spurious data values that need to be edited. To use this filter, click the radio button next to the Value Change Threshold option and then enter a threshold value in the text box. Click the Apply Filter button to run the filter, and the Clear Selection button to clear the selection. The following figure shows an example of the value change filter.

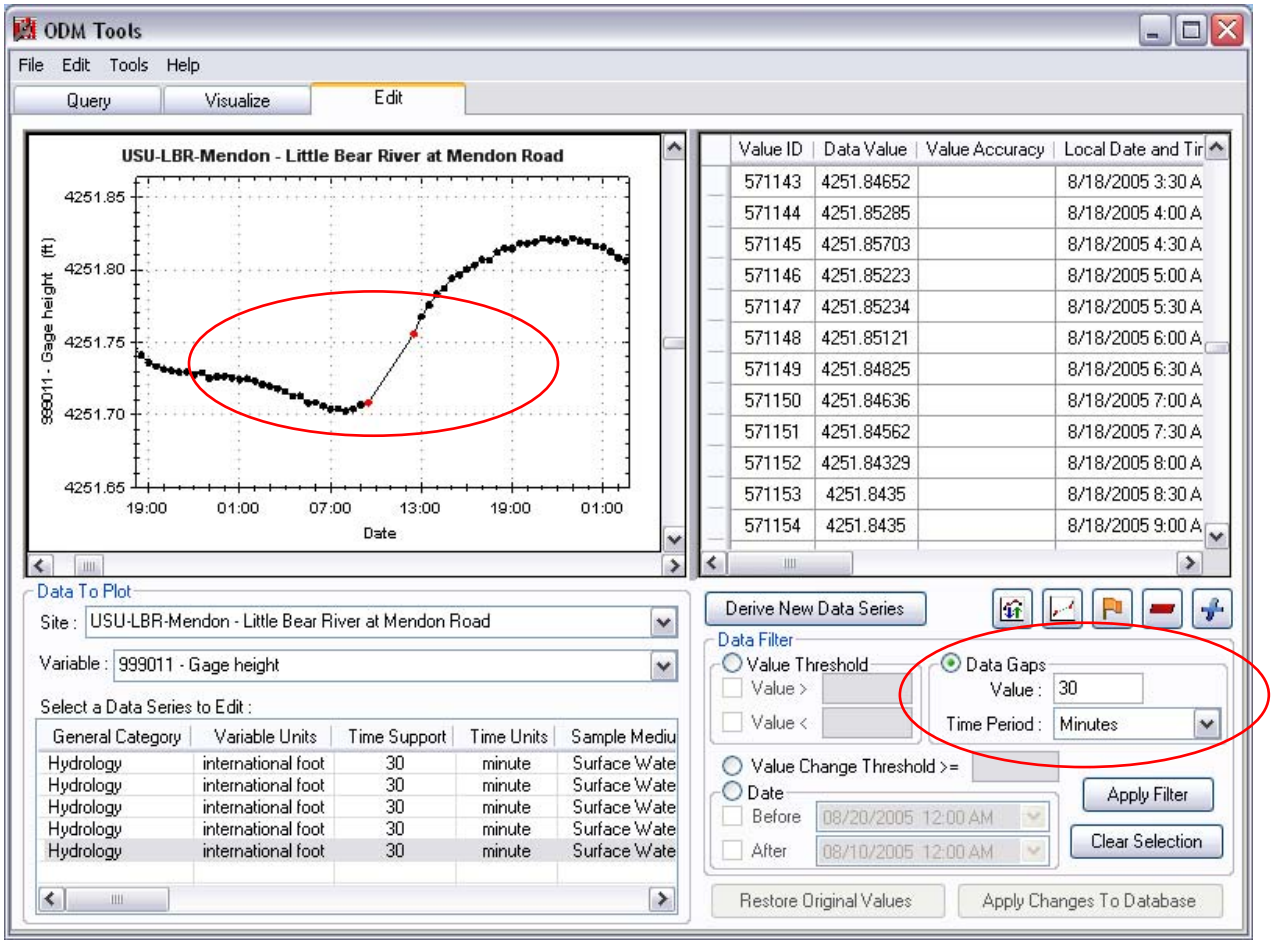


NOTE: This filter selects the data values where the change from the previous data value is greater than the value change threshold (i.e., it looks backward and not forward).

Selecting Data Values Using a Date Range – The date range filter allows you to select a set of data values that fall within a specific date range. This is useful for isolating data values for editing based on their date. For example, it may be discovered that a water temperature sensor was buried in sediment for a period of time. The data values within the time period that the sensor was buried would need to be either edited or deleted since they are not representative of water temperature. To use this filter, click on the radio button next to the Date option and then click on the check boxes next to the date criteria that you wish to use (i.e., Before, After). Enter a date into the date selection box(es) and then click the Apply Filter button to perform the selection. You can use this filter to select data values before a certain date (Before), after a certain date (After), between two dates (Before and After), as well as before a certain date AND after a certain date (i.e., exclude a date range from selection). The following is an example of the Date filter.

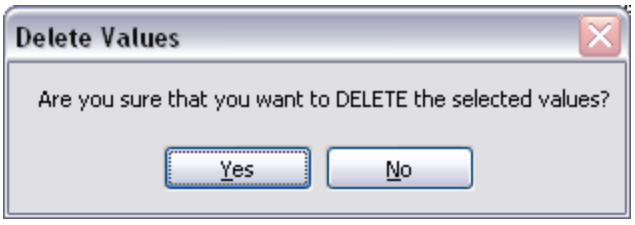


Selecting Data Values with Data Gaps – The data gap filter allows you to select data values that occur at the beginning and end of a data gap. This is useful in identifying gaps in a data series that are caused by instrument malfunctions, etc. In many cases, especially where the data series is regular with a high data collection frequency, these gaps may be small enough that they are not easily discovered through visual inspection of the data on a time series plot. The data gap filter uses a value and a time period to search for data gaps. The value is the numeric value of the size of the gap, and the Time Period of the gap is the units of the gap size (i.e., minutes, hours, days, etc.). ODM Tools uses the input value and time period units to search for gaps in the data series. This filter selects data values at the beginning and end of data gaps. To use this filter, click the radio button next to the Data Gaps option, input a value for the gap size in the Value text box, and then select a time period from the Time Period pull down menu. Click the Apply Filter button to perform the selection. The following figure is an example of the data gaps filter.



6.2.3 Deleting Data Values

ODM Tools allows you to select data values within a data series and permanently delete them from the database. This is useful for removing erroneous values from the database, for example deleting values from a water temperature data series for time periods when the temperature sensor was out of the water or buried in sediment. Data values can be selected for deletion using the data filters and methods described above. To delete selected data values, click on the Delete Data Values button . When you click on this button, the following form will appear asking you whether you whether you really want to delete the values from the selected data series.

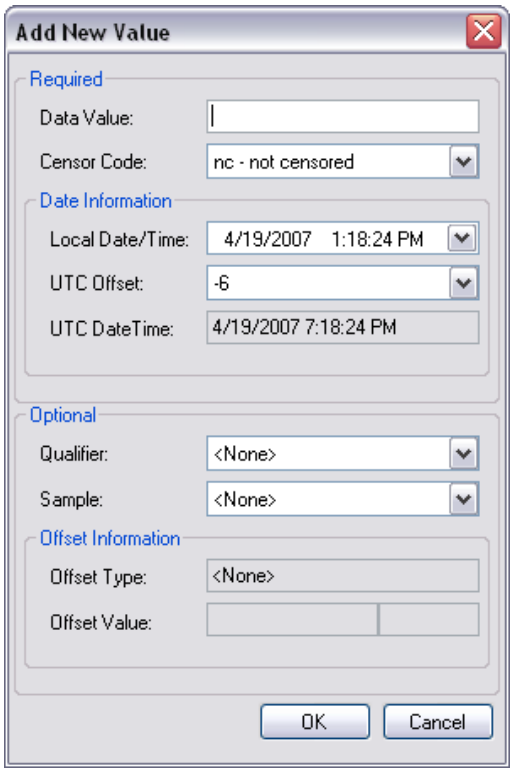


Click Yes to approve the delete and remove the selected data values from the data series.

NOTE: The ODM Tools Delete Data Values button deletes whatever data values are selected in the table and plot views.

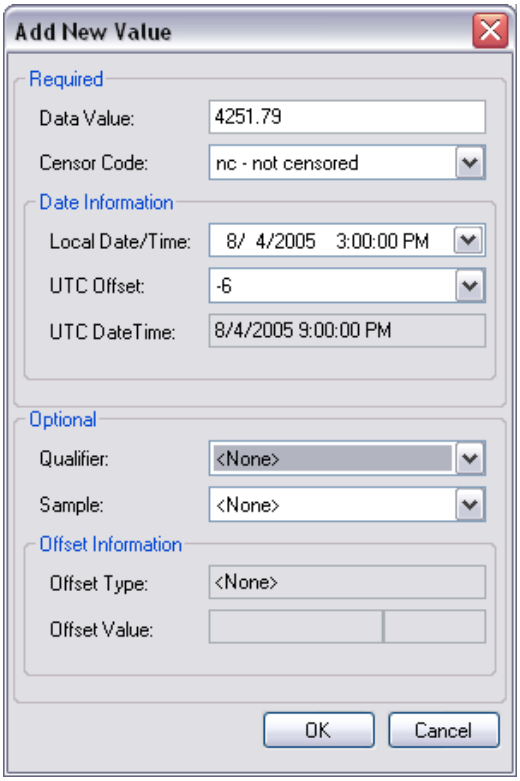
6.2.4 Inserting Data Values to Fill Data Gaps

Where *brief* gaps occur in a data series, ODM Tools allows you to fill them by inserting data values. Data values must be added individually using the Add Data Values Button,  . When you click on the Add Data Values button, the following form appears:



The Add New Value form is divided into Required information and Optional information. In the Required information group, you must first fill in the numeric value of the data value to be inserted in the Data Value text box. Next, you must assign the new data value a Censor Code from the Censor Code drop down list. Assign your new data value a Local Date/Time and a UTC Offset using the Local Date/Time dialog and the UTC Offset drop down. The UTC DateTime will be automatically populated based on your Local Date/Time and UTC Offset. ‘

In the Optional group, you can assign a Qualifier to your new data value using the Qualifier Drop down and you may assign a Sample to your new data value using the Sample drop down. In the Offset information group, you have the option of specifying an Offset Value for your new data value. The Offset Type is automatically populated with the Offset Type for the data series that you are editing. Once populated, the Add New Value form should look something like the following:



Click the OK button to add the new data value to the data series. You will notice that a new data value has been added to the data series and that it is selected in the plot and table views.

NOTE: When selecting a Qualifier for your new data value, you can either select an existing qualifier from the drop down list, or you can click on <Create New> in the list to create a new qualifier. Creating a new Qualifier will add it to the list, after which you can assign it to your new data value.

NOTE: You can only assign a sample to a new data value if it already exists in the database.

6.2.5 Interpolating Data Values

ODM Tools allows you to replace existing data values with linearly interpolated values. Interpolation is based on the data value before the first selected data value and the data value immediately after the last selected data value. Any existing data values within the selected range are interpolated. This function is useful for fixing *brief* data gaps in continuous data series caused by datalogger or sensor malfunctions where interpolation is appropriate.

The interpolation function of ODM Tools will only interpolate selected data values. This means that if a data gap exists that you wish to fix by interpolating the data values, you must first insert data values to fill the gap and then use the interpolation function to interpolate the data values. The data values that you insert (see section 6.2.4 above) can be arbitrary since you are going to replace them by interpolating. Use the following steps to use the interpolation functionality of ODM Tools.

1. Select a set of data values to interpolate by using the data filters or by clicking on them in the table view. Your screen should look something like the following:

2. Click on the Interpolate Data Values button,  . ODM Tools will then interpolate the data values that you have selected by using the first data value before your selection and the next data value after your selection. Your screen should now look something like the following:

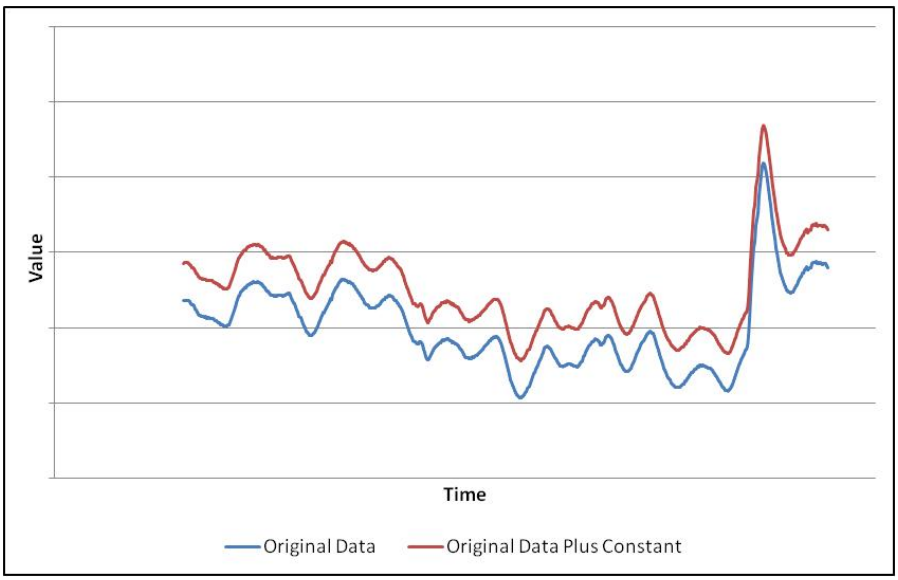
NOTE: The functionality to linearly interpolate data values that is implemented within ODM Tools is not appropriate for filling data gaps in all situations. You are left to your own discretion as to whether data values can be replaced by interpolation.

6.2.6 Adjusting Data Values

There are multiple ways to adjust selected data values using ODM Tools. First, if only a single data value is selected, you can adjust its value by simply double clicking on the DataValue cell for the selected record in the table view. This will make that cell editable, and you can type in a new value. When you leave that cell by clicking somewhere else on the table or Edit tab, the new data value that you entered will be saved in the place of the previous data value.

When multiple data values are selected, you can click on the Adjust Data Values button to adjust their values. ODM Tools currently allows you to adjust multiple selected data values within a data series using the following options:

Adding a Constant Value – This function adds a user supplied constant value to all of the selected data values. This may be useful for adjusting all or portions of a data series where an incorrect offset was used in a datalogger program. The following is an example of the result of adding a constant value.



Multiplying by a Constant Value - This function multiplies all of the selected data values by a user supplied constant value. This may be useful for adjusting all or portions of a data series where an incorrect multiplier was used in a datalogger program. The result of multiplying by a constant value is very similar to the example above.

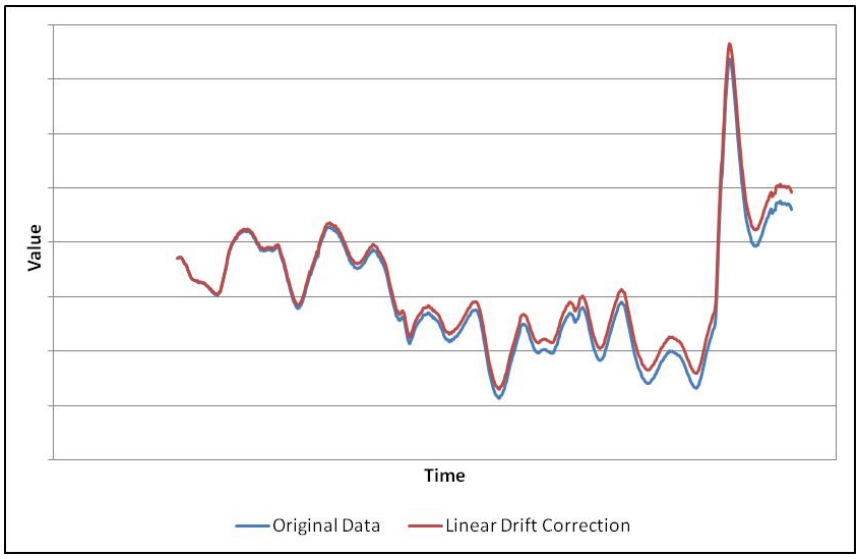
Applying a Linear Drift Correction – This function applies a linear drift correction to all of the selected data values. This may be useful in adjusting data series that were measured using a sensor that is prone to drifting over time. An example is a water level data series that is calibrated to a stage plate. If after a certain period of time the water level measurements made by the water level sensor diverge from observations made using the stationary stage plate, the two can be brought back into sync using a drift correction. The amount by which each data value is adjusted is calculated using the following equation:

$$y_n = y_o + G\left(\frac{x_i}{x_t}\right)$$

Where:

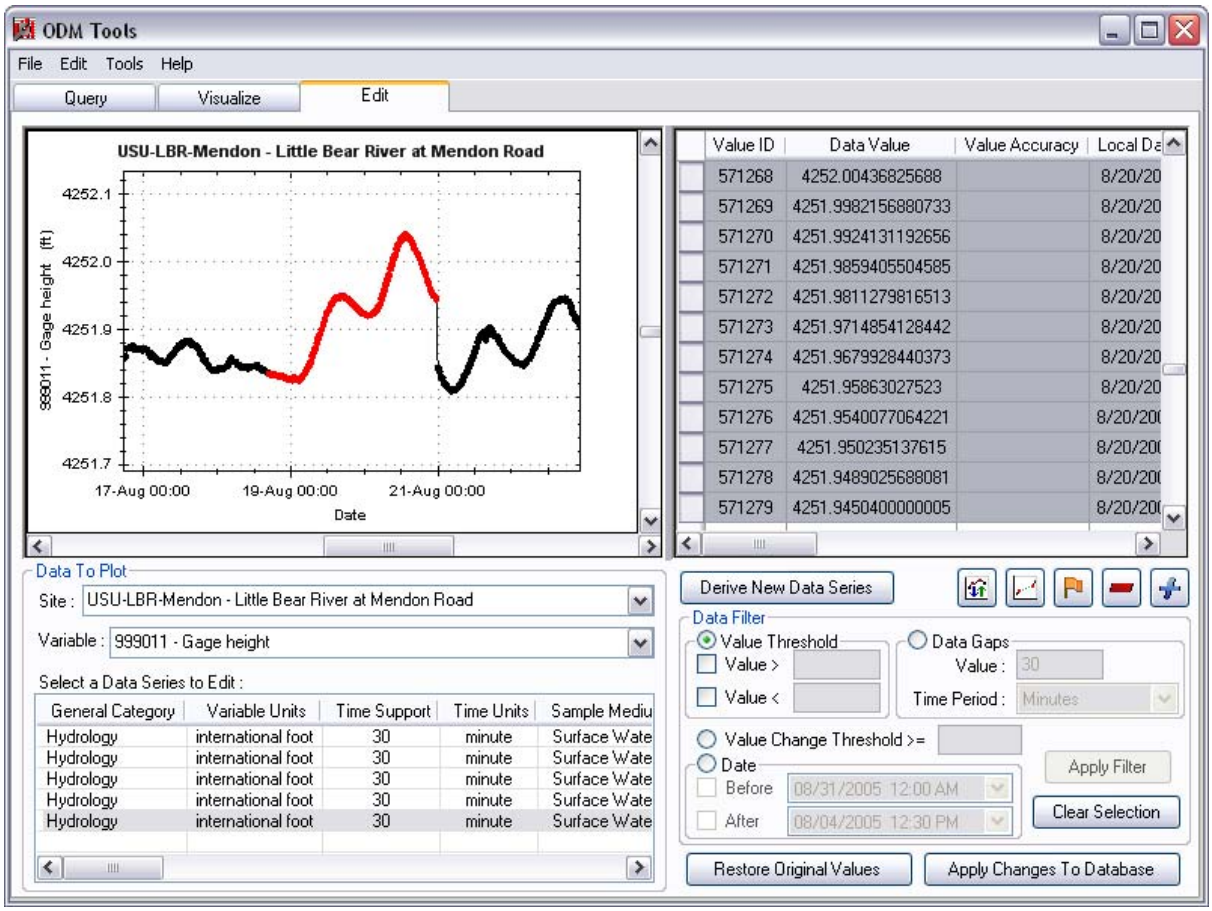
y_n = New adjusted value
y_o = Old value
G = Final gap width (i.e., the difference between the sensor reading and the stage plate reading)
x_i = Distance (in time units) of y_n from the beginning of the linear drift correction
x_t = Total length (in time units) of the linear drift correction

The following is an example of the result of applying a linear drift correction.

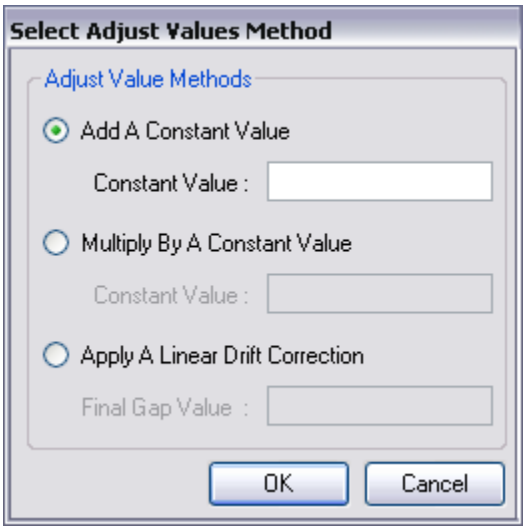


The following steps provide an example of how to use the Adjust Data Values functionality of ODM Tools.

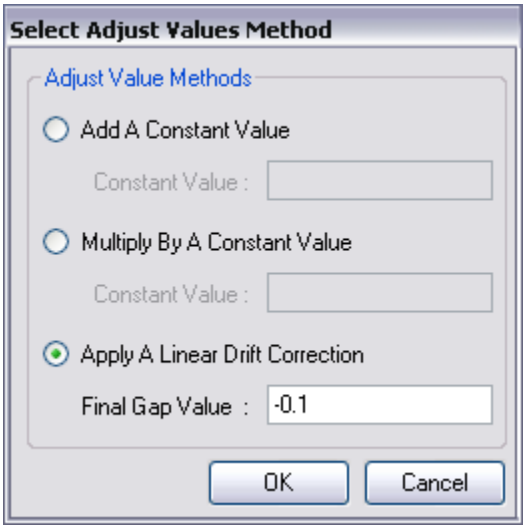
1. Select a set of data values by using the ODM Tools data filters or by clicking on them in the table view. Your screen should look something like the following:



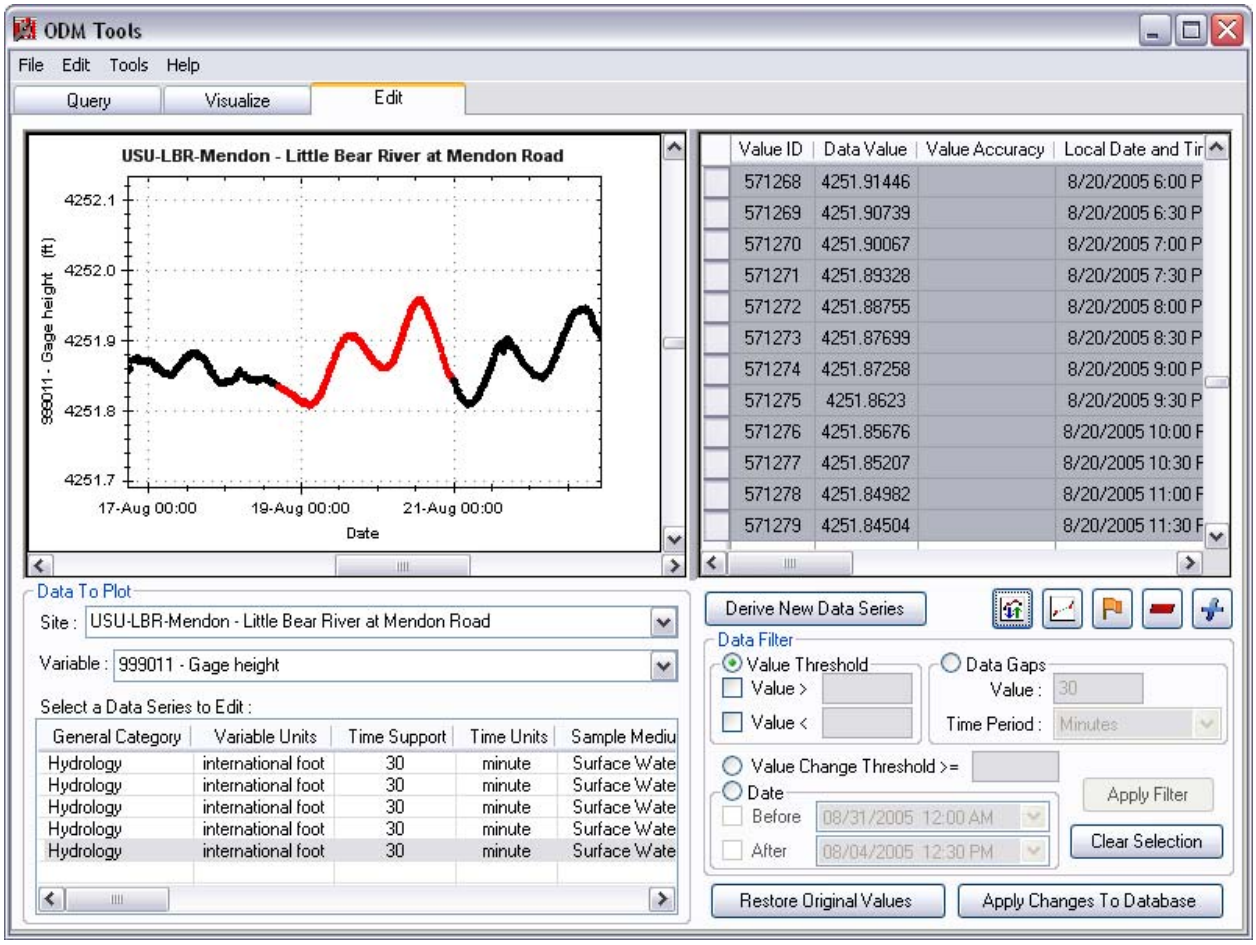
2. Click on the Adjust Data Values Button,  . The following form will open.



3. Click the radio button next to the option that you wish to implement. For this example, we will apply a linear drift correction to a set of data values. We have noticed that our sensor has drifted up by approximately 0.1 ft over the selected time period, so we will use a final gap value of -0.01. The minus sign denotes that we want to adjust the data values down to match the subsequent data. The Select Adjust Values Method form should now look like the following:



4. Click OK to apply the drift correction. ODM Tools will then adjust each of the selected data values. After applying the linear drift correction, your screen should now look something like the following:

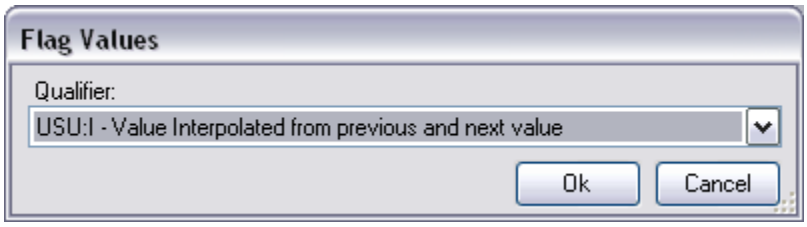


6.2.7 Flagging Data Values with a Data Qualifying Comment

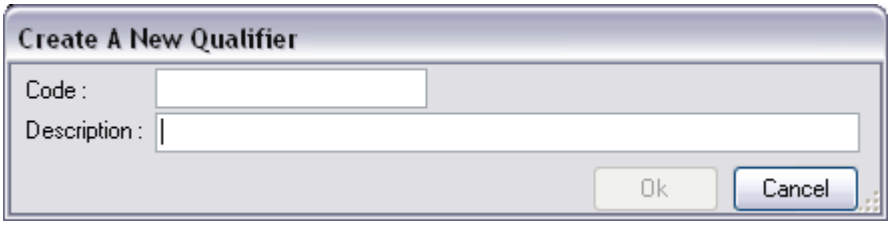
When data values are in any way questionable or when they are edited, it may be critical that they are flagged with a data qualifying comment so that when they are used for analysis it is clear that they are questionable or have been modified. A good example would be when data gaps are filled with interpolated values. Users of the data may wish to know which data values are real observations and which have been created by interpolation. When creating interpolated values, you can flag them with a data qualifying comment that specifies that they were interpolated. Use the following steps to add data qualifying comments to selected data values.

1. Select a set of data values using the ODM Tools data filters or by clicking on them in the table view.

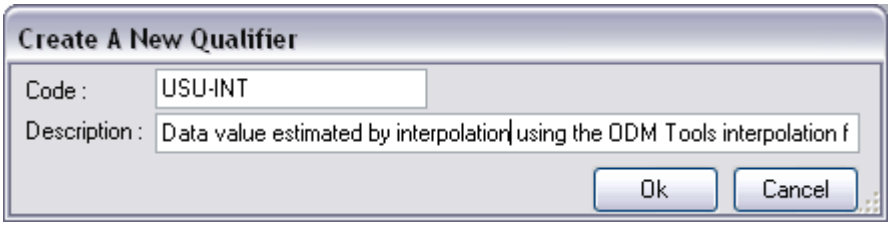
2. Click on the Flag Data Values button , which will open the following form and allow you to add a data qualifying comment to the selected data values.



3. You can either select a qualifying comment from the list of qualifiers already in the database, or you can choose <Create New> from the Qualifier drop down list to create a new qualifier. For this example, we will create a new qualifier. Click on <Create New> in the drop down list. The following form will appear:



4. Assign your new qualifier a code and a description by typing in the appropriate text boxes. Your screen may look something like the following:



5. Click on the OK button. This will add the new qualifier to the Qualifiers table of your ODM database, and you will now notice that the qualifier that you just created has been added to the Qualifier drop down list. Make sure that your new qualifier is selected in the drop down list and then click OK. This will assign the selected qualifier to the selected data values.

NOTE: When you select a data qualifying comment, it will be applied to all of the selected data values. In addition, it is important to note that this will overwrite any existing data qualifying comments that these data values may already have. This is because ODM currently supports only one data qualifying comment per data value.

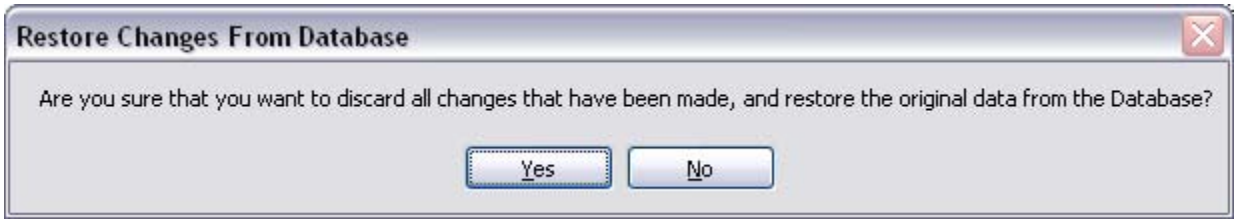
6.2.8 Applying Edits to the Underlying Database

NOTE: It is important to understand before reading the following that once you commit your data series edits to the database you CANNOT undo them using the ODM Tools application. Your changes will be permanently applied to the database. If you have made a mistake while editing the data series and wish to restore the original data values, see Section 6.2.9 below.

ODM Tools holds all data series edits in memory until you choose to commit them to the database. When all of your data series edits are complete and you are satisfied that they are correct, click the Apply Changes to Database button located at the bottom right of the Edit tab. This will commit all of the changes that you have made to the selected data series to the database. As stated in the note above, this action is final, and you cannot undo your changes using the ODM Tools application once you have committed your changes to the database.

6.2.9 Restoring Original Data Values

Since ODM Tools holds all data series edits in memory until you choose to commit them to the database, you can restore the original data values at any time during your edit session up to the point that you commit them to the database using the functionality described above. If you wish to discard all of the changes that you have made to the data series during the current editing session, click on the Restore Original Values button. The following form will pop up:



Click Yes to discard all of the changes that you have made and restore the original data from the database.

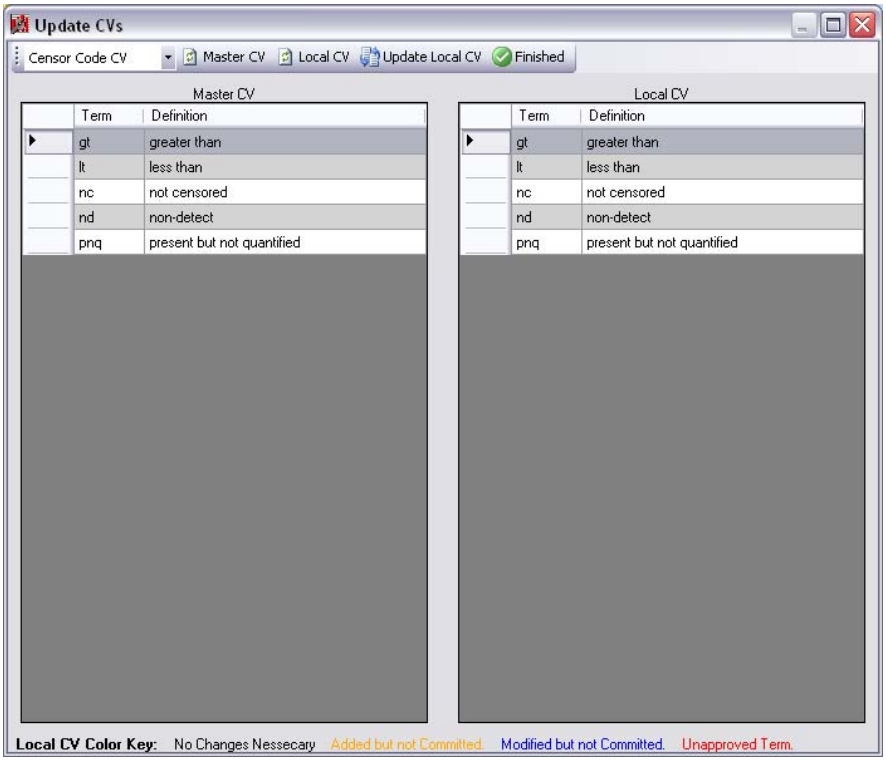
NOTE: This functionality will only undo data series edits that have been made in the current editing session and have not been committed to the database. Once the data series edits are committed to the database using the Apply Changes to Database button they cannot be undone using ODM Tools.

7.0 Updating the ODM Controlled Vocabularies

ODM Tools includes functionality to update the controlled vocabulary tables within a local ODM database. The controlled vocabulary tables define the valid values for several fields within several of the tables within ODM. These controlled vocabularies change over time as new values are added, and a master set of controlled vocabulary tables is available at http://water.usu.edu/cuahsi/odm/cv.aspx. Users can request additions to the ODM controlled vocabularies at this website. The functionality for updating the ODM controlled vocabulary tables can be accessed via the ODM Tools "Tools" drop down menu. The following sections describe the process of updating the controlled vocabulary tables using ODM Tools.

7.1 Interactive CV Update

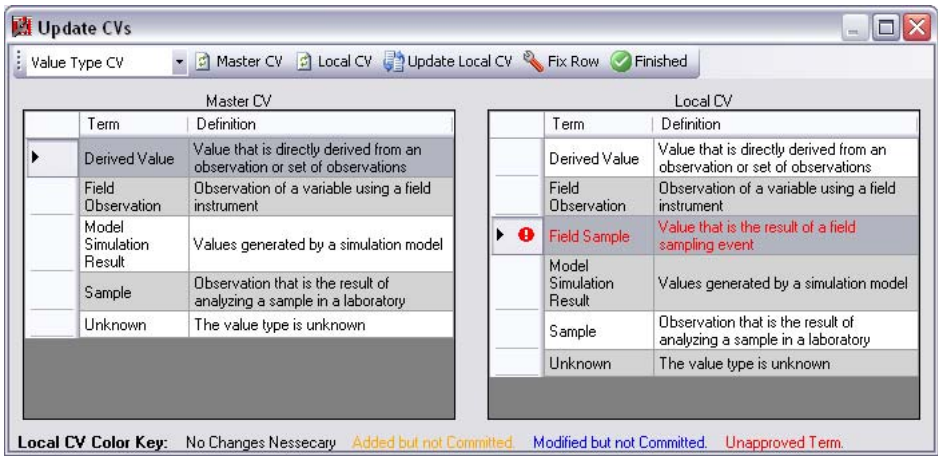
The interactive CV update allows you to compare, side-by-side, the values within your local CV tables with those in the master CV repository. You can then use the available tools to resolve any discrepancies that arise. To run the interactive CV update, click on "Interactive CV Update" in the "Tools" drop down menu. The following window will appear.



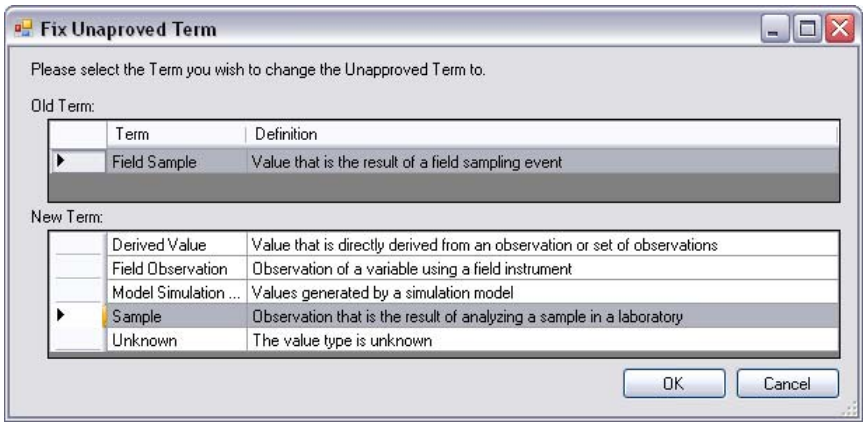
The terms in the master CV are shown on the left, and the terms in your local CV are shown on the right. You can automatically compare the terms from both sources and highlight any discrepancies by clicking on the "Update Local CV" button on the toolbar. This will add any new terms from the master CV to the local CV, and it will highlight in color any actions that have been taken. The following are the cases that can occur when you click on "Update Local CV".

1. *The term and its definition in your CV table is the same as the term and definition in the master repository*. In this case, no changes are required, and the term is shown in black on the right in your local CV.
2. *A term that was in the master repository but not in your CV table has been added to your CV table*. In this case, the term is highlighted in orange and you must commit this change to your database.
3. *A term or definition in your CV table is different than the term or definition in the master CV*. In this case, the term is highlighted in blue and you must commit this change to your database.
4. *A term exists in your local CV, but it does not exist in the master repository*. This is considered an "unapproved term" in that it has not been approved as part of the master list. This can happen if you add your own terms to your own database, or if you are using a term that is then removed from the master list. Unapproved terms are automatically removed from your CV table unless they are being used to describe the data in your database. If they are being used, they will remain in your CV table, and they are highlighted in red.

Cases 1 – 3 do not require any user intervention, and you can commit these to your database by clicking on the "Apply" button. When you click the "Apply" button, all of these changes will be committed to your database, including cascading updates to any tables that use the CV terms. Clicking on the "Apply" button will not have any effect on unapproved terms (case 4). These terms can either be left in the database while you request their addition to the master CV repository, or you can match them to a different term that has been approved. The following is an example of a CV with an unapproved term.



If you click on the unapproved term to select it, you will notice that an additional button has been added to the toolbar called "Fix Row." If you click on this button, the following window will appear.



On this form, you can select an approved term to replace your unapproved term. By doing so, this change will be cascaded to the tables within the database that use the controlled vocabularies. In the above example, I might choose to replace "Field Sample" with the approved term "Sample." If you click on the "OK" button and then the "Apply" button, the changes to unapproved term will be committed to the database. Click the "Finished" button when all of the changes have been made.

7.2 Quick CV Update

The quick CV update allows you to quickly add new terms from the master CV repository to your local ODM database. The quick CV update simply adds new terms that don't already exist in your CV tables. It does not resolve discrepancies between any terms that may have been changed in some way (i.e., either a term or its definition in your database does not match the master repository). These differences are resolved using the interactive CV update. To perform a quick CV update on your ODM database, select "Quick CV Update" from the "Tools" drop down menu. When you do so, a window will appear for a few moments indicating that your CV tables are being updated, after which the quick CV update is complete. When complete, any terms that existed in the master CVs but not in your local CVs will have been added to your database.

References

Cleveland, W.S., and R. McGill. 1984. The many faces of a scatterplot. Journal of the American Statistical Association. Volume 79, Number 388. P. 807-822.

Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association. Volume 74, Number 368. P. 829-836.

Helsel, D.R., and R.M. Hirsch. 2002. Statistical Methods in Water Resources. In: Techniques of Water-Resources Investigations of the United States Geological Survey, Book 4, Hydrologic Analysis and Interpretation. http://pubs.usgs.gov/twri/twri4a3/html/pdf_new.html

Tarboton, D.G., Horsburgh, J.S., and D.R. Maidment. 2007. CUAHSI Community Observations Data Model (ODM) Design Specifications Document: Version 1.0. http://www.cuahsi.org/his/odm.html

Appendix A
Instructions for Attaching the ODM Tools Test Databases to an Instance of Microsoft SQL Server 2005

Introduction

ODM Tools requires that you have an ODM database running in Microsoft SQL Server 2005 so that you can attach to it. ODM Tools can be used with any of the SQL Server 2005 versions (i.e., Express, Standard, Enterprise). In the event that you do not already have an ODM database, we have provided two test databases for download so that you can test the functionality of ODM Tools. These databases are available at http://water.usu.edu/CUAHSI/ODM/. The purpose of this Appendix is to show you how to attach these test databases to your instance of Microsoft SQL Server. The example instructions in this Appendix were completed in Microsoft SQL Server 2005 Express.

Installing SQL Server 2005 Express

If you do not already have an instance of Microsoft SQL Server running, you can download and install Microsoft SQL Server 2005 Express from Microsoft for free. It is recommended that you download and install both SQL Server 2005 Express and SQL Server Management Studio Express. You can get both of these products in a single installation (download and install the SQL Server 2005 Express Edition with Advanced Services SP2) as well as installation instructions at the following Microsoft URL:

http://www.microsoft.com/sql/editions/express/default.mspx

When you install SQL Server 2005 Express, you must enable mixed mode authentication (both SQL Server authentication and Windows authentication). ODM Tools relies on SQL Server authentication to connect to ODM databases, and SQL Server authentication is only enabled when you choose the mixed mode authentication during installation. When you enable mixed mode authentication during installation, you will be prompted to create a password for the "sa" logon. The "sa" login is the administrative account and allows you to manage all other logins. Choose a password that you will remember as you will use this account frequently.

Attaching the Test Databases to SQL Server

The following are the steps required to attach one of the test databases to an instance of Microsoft SQL Server. These steps were written using SQL Server Management Studio Express; however, the steps are similar regardless of which version of the Microsoft SQL Server Management Studio you are using.

1. Extract the test database and its log file from the zip file to a location on your hard drive using WinZip or some other equivalent software. It is suggested that you extract your

database to the default SQL Server data folder, which is located at the following location on disk: C:\Program Files\Microsoft SQL Server\MSSQL.1\MSSQL\Data\

NOTE: You can extract your test databases to any location on disk. However, if you do so and you have connected to SQL Server using SQL Server authentication and not Windows Authentication, you will have to give SQL Server access to read and write to the folder where you extracted your databases prior to attaching them. SQL Server already has access to its default data folder using either SQL Server or Windows authentication and so this is the easiest location in which to work.

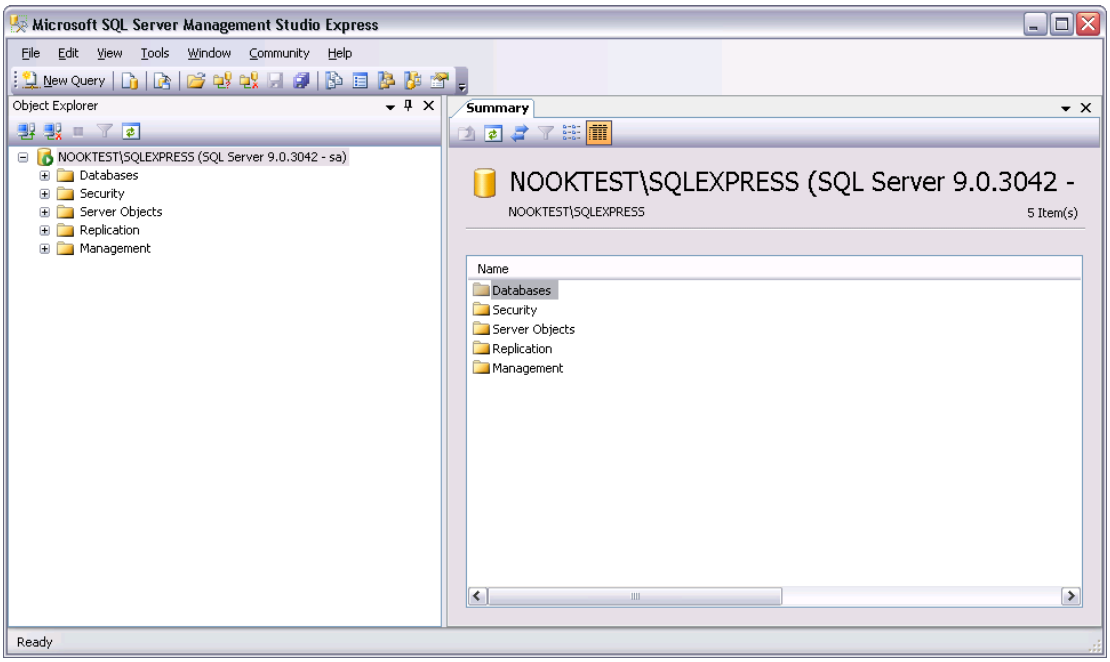
2. Open the Microsoft SQL Server Management Studio from the Start Menu by clicking on Start --- All Programs --- Microsoft SQL Server 2005 --- SQL Server Management Studio. The following window will appear.

NOTE: The path to your SQL Server Management Studio shortcut in the Start menu may be different depending on which version of SQL Server you have installed and where you chose to put the shortcut in the Start Menu.

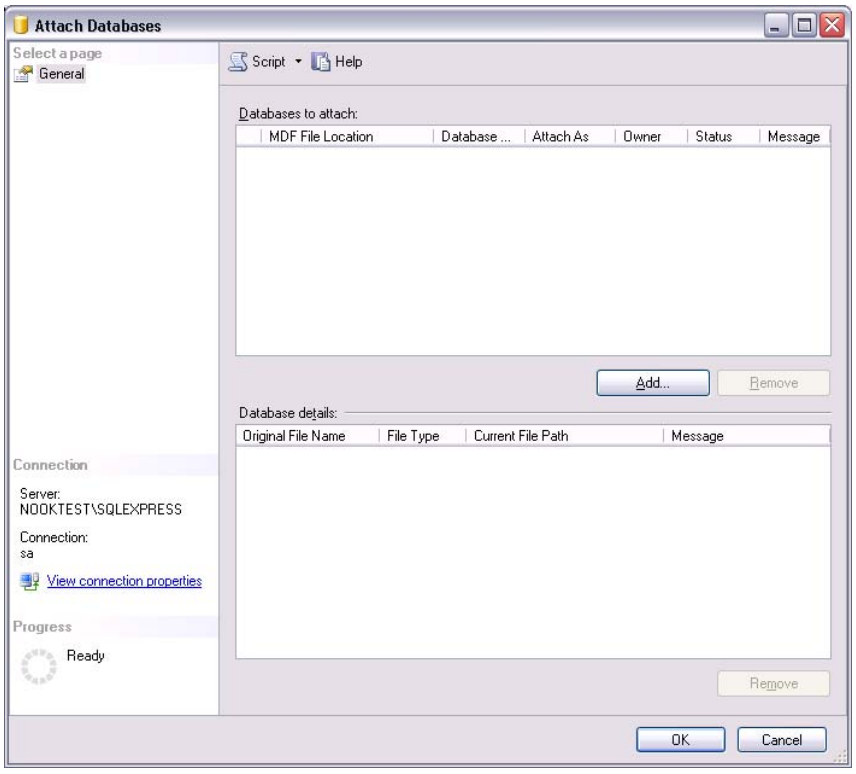
3. It is assumed that you are connecting to your local instance of SQL Server. You should see your computer's name followed by "\SQLEXPRESS" in the "Server Name" drop down. In the following figure, the computer's name is "NOOKTEST." Change the Authentication dropdown to "SQL Server Authentication," enter "sa" for your login, and then enter your administrative password in the "Password" text box. Your login screen should look similar to the following.

NOTE: You can complete these steps using Windows authentication rather than SQL Server authentication. However, ODM Tools requires that you have a SQL Server authentication login for the database that you are attaching.

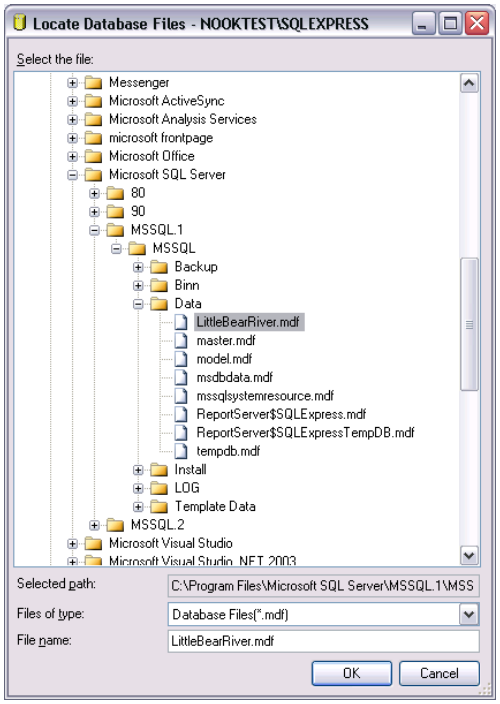
4. Click on the "Connect" button. This will connect the Management Studio to your local SQL Server instance. Your Management Studio window should look similar to the following.



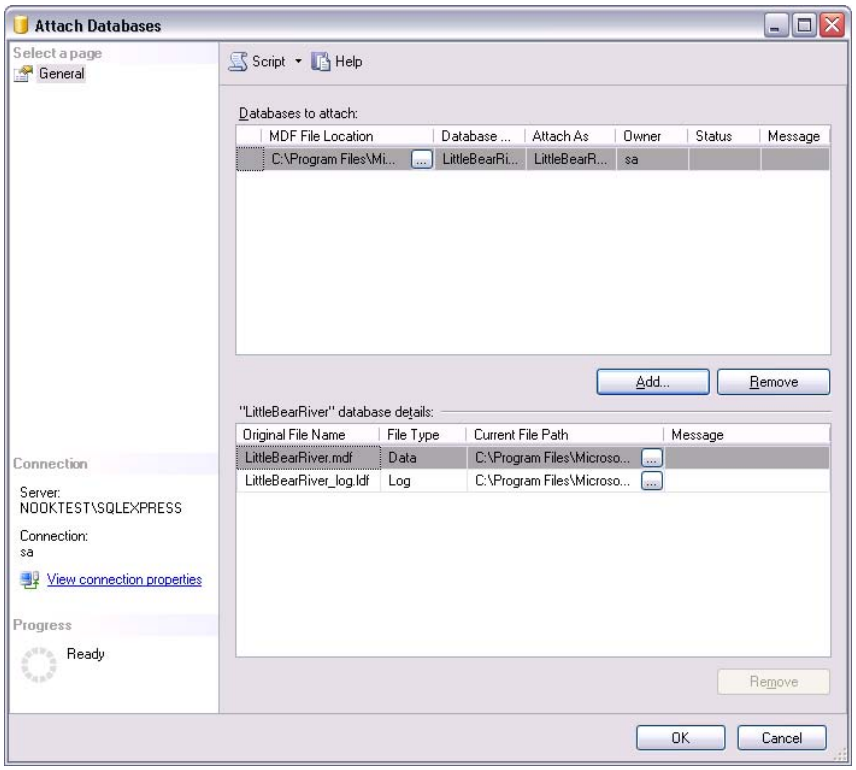
5. Right click on the "Databases" item under your server in the Object Explorer at the left of the window and choose "Attach" from the context menu. The following window will appear.



6. Click on the "Add" button near the center of the form. In the window that opens, navigate to the location on your hard drive where you extracted the ODM test database. Select the .mdf file associated with the database that you want to attach. See the following figure for an example.

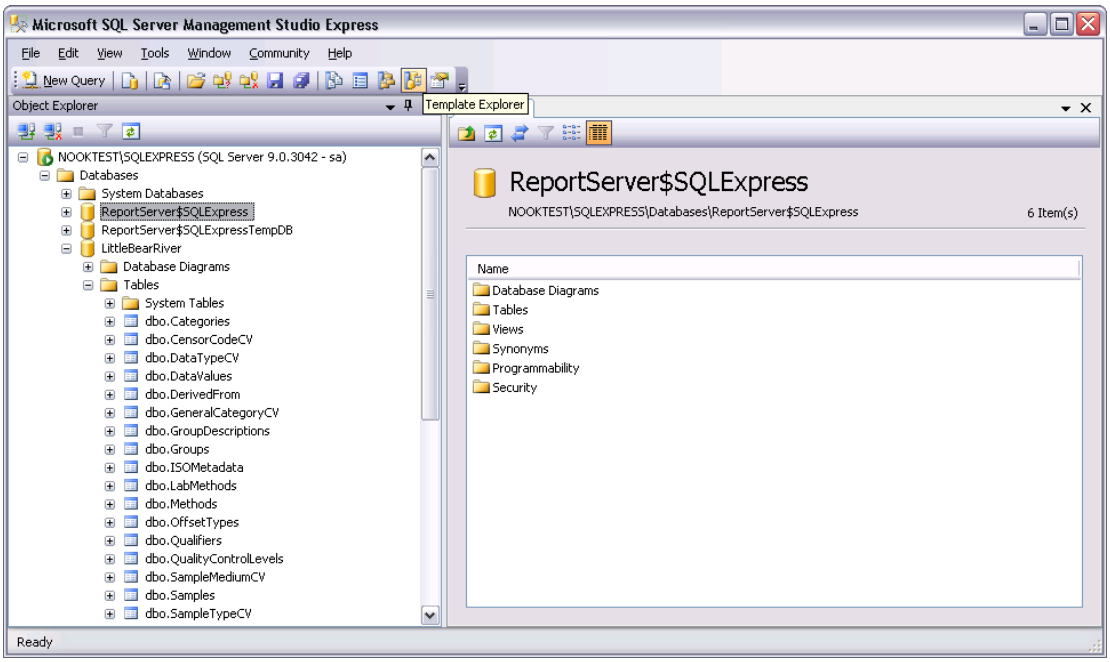


7. Click the “OK” button. This will return you to the “Attach Databases” form and will populate that form with the information needed to attach the database that you have selected. You will notice that your selected database is listed in the “Databases to attach” section and that the details of your database files are shown at the bottom of the form. See the following figure.

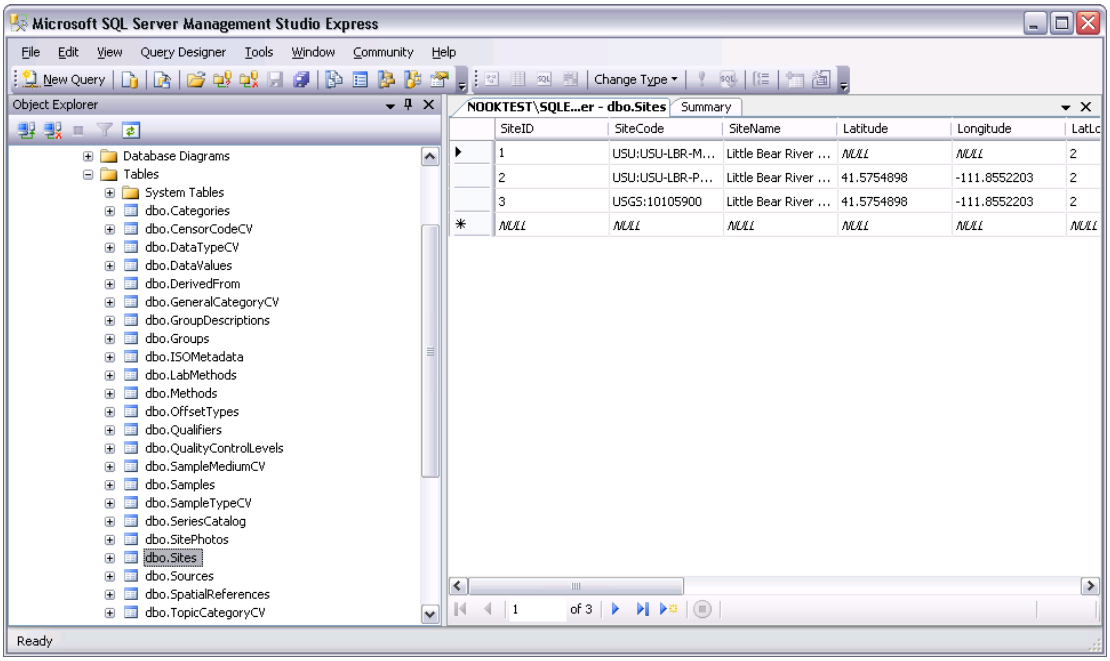


8. Click the “OK” button and then wait for a moment while your database is being attached.

9. Once your database has been attached, expand the “Databases” item under your server in the Object Explorer by clicking on the plus sign next to “Databases.” You should now see an item under “Databases” for the database that you just attached. If you do not see an item for your database, right click on the “Databases” item and choose “Refresh” from the context menu. You can further expand your database by clicking on the plus sign next to its name and then clicking on the plus sign next to “Tables” (see the following figure).



10. You can view the contents of the tables in your database by right clicking on them in the Object Explorer and choosing “Open Table” (see the following for the Sites table).



11. You are now ready to attach the ODM Tools application to your database using the steps outlined in the main text of this document. You can use the “sa” login for your database when connecting ODM Tools.

Appendix B
CUAHSI ODM MyDB Version 1.0
Design Specifications

Jeffery S. Horsburgh[1], David G. Tarboton[1], David R. Maidment[2], and Tim Whiteaker[2]

Introduction

The CUAHSI Hydrologic Information System (HIS) project is developing information technology infrastructure to support hydrologic science. One aspect of this is a data model for the storage and retrieval of hydrologic observations in a relational database. The Observations Data Model (ODM) is designed to store hydrologic observations and sufficient ancillary information (metadata) about the data values to provide traceable heritage from raw measurements to usable information allowing them to be unambiguously interpreted and used (Tarboton et al., 2007). A relational database format has been used in the design of ODM to provide querying capability and to allow data retrieval supporting diverse analyses.

It is intended that an instance of ODM will be the central repository for data and information related to hydrologic observations made in hydrologic and environmental observatories, test beds, and other scientific study areas. Because of this, there are several issues associated with providing user access to data stored within an ODM instance. First, since ODM is the central repository for data for a test bed or observatory, security is an issue. Users should be given unrestricted access to the data that they need, but this should be done in a way that protects the security and integrity of the central database. Second, users may wish to use the data stored in an ODM instance for complex or intensive analysis. This is best done with a copy of the data on their own machines rather than repeatedly querying the database to get the data. Last, many users do not wish to invest the time required to learn how to use a relatively complicated relational database structure when the analysis that they wish to do does not require the processing tools available within the database (i.e., they just want to get the data so they can work in a simple spreadsheet).

One solution to the issues listed above is to provide users with a set of Web Services that allow them to programmatically extract data from the database. The CUAHIS HIS team is currently developing Web Services for the ODM; however, it is beyond the scope of this document to describe those web services. Another solution to these issues is to allow users to export the data that they need from the ODM database and use it on their local machines. This Appendix describes MyDB, which is a simple data export format that can be used to supply data stored in an instance of ODM to users.

[1] Utah Water Research Laboratory, Utah State University
[2] Center for Research in Water Resources, University of Texas at Austin

MyDB Purpose

The purpose of MyDB is to provide a simple data structure for storing data series exported from an instance of ODM that preserves the entire context of the exported data series and does not require expensive software or database expertise to use.

Data Series Defined

In order to fully grasp the concepts that follow, the idea of a "data series" in the context of ODM must be clarified. A "data series" is an organizing principle that is present in the ODM. A data series consists of all of the values associated with a unique Site and Variable combination within the ODM database. For example, all of the water temperature (Variable) values measured in the Logan River near Logan, UT (Site) would be considered a data series. In addition to the site and variable, however, data series are also uniquely qualified by the method that was used to generate the data values and the quality control level of the data values. The full specification for a data series, then, would be something like "all of the raw unchecked (QualityControlLevel) water temperature (Variable) values measured in the Logan River near Logan, UT (Site) using a field temperature sensor (method)."

MyDB Version 1.0 Structure and Function

MyDB Table

MyDB is a single table that is essentially a much simpler version of the full ODM and stores a subset of all of the types of information available in ODM. The MyDB table can be easily interpreted since there are no relationships to navigate. MyDB is saved as a tab or comma delimited text file, which can easily be consumed by a number of client applications, including most spreadsheet and data analysis applications. Text files are generally operating system and software independent and are not limited in the number of records that can be stored, whereas some spreadsheet programs are.

Core MyDB Table

A MyDB table consists of one or more data series that have been exported from an instance of ODM. Each record in a MyDB table represents a single observation of a specific variable at a specific site made using a specific method and having a specific quality control level. In its simplest form, MyDB contains all of the data values that make up one or more data series and their value level attributes (i.e., attributes that are specific to each data value and not the data series as a whole). Data series level attributes are stored in a companion metadata file, which is described below. Table A-1 lists the fields from ODM that are included in the most basic implementation of MyDB, called the Core MyDB, along with the field's data type and a description of the field. MyDB field names have been adopted from ODM for consistency. There are four general data types for MyDB fields. The four types are:

- Integer – a whole number, e.g. 234
- Decimal – a decimal number, e.g. 234.44

- Text – text, e.g. “EPA”
- DateTime – date, e.g. 4/14/1976 3:00 PM

Table A-1. Core MyDB fields.

MyDB Field Name	Data Type	Description	Can Be Null
SeriesID	Integer	Unique ID for the data series. This ID is the link into the metadata file where all of the additional information for the data series would be stored	No
ValueID	Integer	Unique ID for each data value in the data series. This is included to preserve the linkage from values in the MyDB table to the values in the ODM.	No
DataValue	Decimal	Numeric value of the observation/data value	No
ValueAccuracy	Decimal	Accuracy of the observation/data value	Yes
LocalDateTime	DateTime	The local date and time associated with the observation/data value	No
SiteCode	Text	Code used by organization that collects the data to identify the site	No
VariableCode	Text	Code used by organization that collects the data to identify the variable	No
OffsetValue	Decimal	The value of the offset associated with the observation/data value	Yes
OffsetTypeID	Integer	An ID that points at the OffsetType as stored in the metadata file. Since the offset is a value level attribute, the offset information must remain with the data values in the MyDB file. However, one OffsetType may qualify many data values in one or more data series within the MyDB file. MyDB will store only an ID associated with the OffsetType, and the OffsetTypes will be defined in the metadata file	Yes/No If an OffsetValue is given, there must be and OffsetTypeID
CensorCode	Text	An indication of whether the observations/data values are censored or not. Conforms to the CensorCode controlled vocabulary	No
QualifierID	Integer	An ID that points to a list of qualifier descriptions in the metadata file. One qualifier may describe many observations and so only the ID is stored in MyDB. The actual descriptions of the qualifiers are stored in the metadata file	Yes
SampleID	Integer	An ID that points to a list of samples and related laboratory methods in the metadata file. One sample may result in many observations, and so only the ID of the sample is stored in MyDB. All other descriptive information about the sample and laboratory methods is stored in the metadata file	Yes

Expanded MyDB Table

The Core MyDB table is compact and consists of the minimum amount of information that is required for the full specification of a data series. However, the Core MyDB table is not particularly friendly to human eyes because only numeric IDs are included in the table. These IDs suffice for computer applications that read a MyDB table and its associated metadata file. For example, a computer application can match the SeriesID in the MyDB table with the corresponding SeriesID in the companion metadata file to find the definitions for all of the

numeric IDs as well as any additional series level attributes. However, users that simply want to import the MyDB table to a spreadsheet will not want to dig through the metadata file to get all of the information that they want. For this reason, the Core MyDB table can be expanded to include additional attributes. Table A-2 lists additional attributes that can be added to the Core MyDB table based on whether a user requests them or not. An expanded MyDB table may contain any or all of the attributes in Table A-2 depending on the preference of the user. It is the responsibility of the application that creates the MyDB table to provide users with the ability to specify which attributes (in addition to the Core attributes) to include in the expanded MyDB table.

Table A-2. Attributes that can be added to the MyDB table to create an expanded version.

Field Name	Data Type	Description	Can Be Null
Time Attributes			
DateTimeUTC	DateTime	Universal UTC date and time at which the value was observed	No
UTCOffset	Integer	Offset in hours from UTC time of the corresponding LocalDateTime value	No
Site Attributes			
SiteName	Text	Full name of sampling location	No
Latitude	Decimal	Latitude in decimal degrees	Yes
Longitude	Decimal	Longitude in decimal degrees. East positive, West negative	Yes
SRSName	Text	Name of Spatial Reference System associated with the latitude and longitude coordinates	Yes
Variable Attributes			
VariableName	Text	Name of the variable that was measured, observed, modeled, etc.	No
VariableUnitsName	Text	Full name of the units associated with the variable	No
VariableUnitsAbbreviation	Text	Text abbreviation of the units associated with the variable	No
SampleMedium	Text	Text description of the medium in which the sample or measurement was made.	No
Qualifier Attributes			
QualifierCode	Text	Text code used by organization that collects the data to identify the data qualifying comment	Yes
QualifierDescription	Text	Full text of the data qualifying comment	No
Offset Attributes			
OffsetDescription	Text	Full text description of the offset type	Yes – if no offset exists
OffsetUnitsName	Text	Full name of the units associated with the offset	Yes – if no offset exists
Source Attributes			
Organization	Text	Name of Organization that collected the data. This should be the agency or organization that collected the data, even if it came out of a database consolidated from many sources such as STORET	Yes
SourceDescription	Text	Full text description of the source of the data	No

MyDB Companion Metadata

In interpreting data for analysis, it is important that the full context of the data (i.e., the metadata) be available. To ensure that the metadata is not lost, MyDB includes a companion metadata file,

formatted as XML, that contains all of the metadata from ODM that may or may not be included in the MyDB table, depending on whether the Core MyDB table is used or whether an expanded MyDB table is used. The information contained within the metadata file comprises all of the data series level attributes contained within ODM, whereas the MyDB table lists the actual data values and the data value level attributes. Between the MyDB table and the companion metadata file, the full ODM specification of a data series is preserved. The following shows the general outline of the MyDB metadata file, including the major sections. Each of the major sections has one or more attributes within them as described in the text of the sections that follow. If more than one data series is contained within a MyDB table, the companion metadata file will have multiple data series elements, each one containing the full metadata for a data series. The link between the data series in the MyDB table and the series metadata in the companion metadata file is the SeriesID, which is given in the <DataSeries> element.

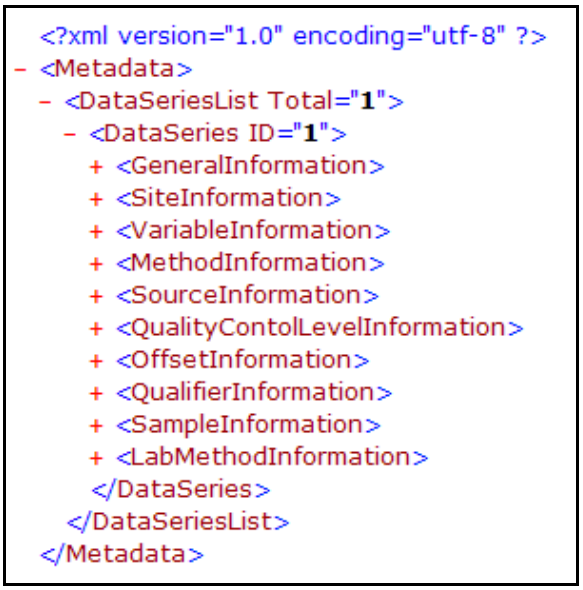


Figure A-1. Example MyDB metadata general outline.

General Data Series Information

The GeneralInformation section of the metadata file contains general metadata information about the data series contained in the MyDB table. This information is extracted from the ISOMetadata table in the ODM. It also contains the date on which the metadata was created. Figure A-2 is an example of the GeneralInformation section of the MyDB metadata file.

Site Information

The SiteInformation section of the metadata file contains all of the information describing the monitoring site associated with the data series contained in the MyDB table. This information is extracted from the Sites table in the ODM and includes spatial reference information from the SpatialReferences table of ODM. Figure A-3 is an example of the SiteInformation section of the MyDB metadata file:

```
<GeneralInformation>
    <TopicCategory>Water Quality, Inland Waters</TopicCategory>
    <Title>Specific conductance, unfiltered at Logan River Above State Dam, Near Logan,
      UT</Title>
    <Abstract>Specific conductance, unfiltered data retrieved from the USGS National Water
      Information System (NWIS) for site code:10109000, obtained through CUAHSI
      Hydrologic Information System. NWIS parameter code:00095, Units: microsiemens
      per centimeter at 25 degrees Celsius, 193 measurements with irregular time steps. A
      value of -9999 indicates no value. Site is located at 1426.8 m with reference to
      NGVD29 datum.</Abstract>
    <ProfileVersion>NULL</ProfileVersion>
    <MetadataLink>NULL</MetadataLink>
    <MetadataCreationDate>2/20/2007 5:00 PM</MetadataCreationDate>
</GeneralInformation>
```

Figure A-2. Example of MyDB metadata GeneralInformation section.

```
<SiteInformation>
    <SiteCode>NWIS:10109000</SiteCode>
    <SiteName>LOGAN RIVER ABOVE STATE DAM, NEAR LOGAN, UT</SiteName>
    <GeographicCoordinates>
        <Latitude>41.74326439</Latitude>
        <Longitude>-111.78272</Longitude>
        <SRSID>4269</SRSID>
        <SRSName>NAD83</SRSName>
        <IsGeographic>True</IsGeographic>
        <Notes>NULL</Notes>
    </GeographicCoordinates>
    <LocalCoordinates>
        <LocalX>NULL</LocalX>
        <LocalY>NULL</LocalY>
        <SRSID>NULL</SRSID>
        <SRSName>NULL</SRSName>
        <IsGeographic>NULL</IsGeographic>
        <Notes>NULL</Notes>
        <Elevation_m>1426.8</Elevation_m>
        <VerticalDatum>NGVD29</VerticalDatum>
    </LocalCoordinates>
    <PosAccuracy_m>NULL</PosAccuracy_m>
    <State>Utah</State>
    <County>Cache</County>
    <Comments>NULL</Comments>
</SiteInformation>
```

Figure A-3. Example MyDB metadata SiteInformation section.

Variable Information

The VariableInformation section of the metadata file contains all of the information describing the variable associated with the data series contained in the MyDB table. The information for this section is extracted from the Variables table in the ODM and includes units information from the Units table of ODM. Figure A-4 is an example of the VariableInformation section of the MyDB metadata file:

```
<VariableInformation>
    <VariableCode>NWIS:00095</VariableCode>
    <VariableName>Specific conductance, unfiltered</VariableName>
    <VariableUnits>
        <UnitsName>microsiemens per centimeter at 25 degrees Celsius</UnitsName>
        <UnitsType>Conductance</UnitsType>
        <UnitsAbbreviation>uS/cm</UnitsAbbreviation>
    </VariableUnits>
    <SampleMedium>Surface Water</SampleMedium>
    <ValueType>Field Observation</ValueType>
    <IsRegular>False</IsRegular>
    <TimeSupport>0</TimeSupport>
    <TimeSupportUnits>
        <UnitsName>hour</UnitsName>
        <UnitsType>Time</UnitsType>
        <UnitsAbbreviation>hr</UnitsAbbreviation>
    </TimeSupportUnits>
    <DataType>Instantaneous</DataType>
    <GeneralCategory>Water Quality</GeneralCategory>
    <NoDataValue>-9999</NoDataValue>
    <PeriodOfRecord>
        <BeginDateTime>9/13/1967 7:35:00 AM</BeginDateTime>
        <EndDateTime>10/2/1991 4:00:00 PM</EndDateTime>
        <BeginDateTimeUTC>9/13/1967 2:35:00 PM</BeginDateTimeUTC>
        <EndDateTimeUTC>10/2/1991 11:00:00 PM</EndDateTimeUTC>
        <ValueCount>193</ValueCount>
    </PeriodOfRecord>
</VariableInformation>
```

Figure A-4. Example MyDB metadata VariableInformation section.

Method Information

The MethodInformation section of the metadata file contains the description of the method used to generate the data series contained in the MyDB file. The information in this section is taken from the Methods table in the ODM. Figure A-5 is an example of the MethodInformation section of the MyDB metadata file:

```
<MethodInformation>
    <MethodDescription>Measured using a Hydrolab DataSonde 4 field specific conductance
      sensor</MethodDescription>
    <MethodLink>http://www.hydrolab.com</MethodLink>
</MethodInformation>
```

Figure A-5. Example MyDB metadata MethodInformation section.

Source Information

The SourceInformation section of the metadata file contains all of the information describing the source associated with the data series contained in the MyDB table. This information is extracted from the Sources table in the ODM. Figure A-6 is an example of the SourceInformation section of the MyDB metadata file:

```
<SourceInformation>
    <Organization>United States Geological Survey</Organization>
    <SourceDescription>Data retrieved from the USGS National Water Information System
      (NWIS)</SourceDescription>
    <SourceLink>http://waterdata.usgs.gov/nwis</SourceLink>
    <Contact>
        <ContactName>Water Webserver Team</ContactName>
        <Phone>1-888-275-8747</Phone>
        <Email>h2oteam@usgs.gov</Email>
        <Address>12201 Sunrise Valley Drive, MS 439</Address>
        <City>Reston</City>
        <State>VA</State>
        <ZipCode>20192</ZipCode>
    </Contact>
</SourceInformation>
```

Figure A-6. Example MyDB metadata SourceInformation section.

Quality Control Level Information

The QualityControlLevelInformation section of the metadata file contains all of the information describing the quality control level of the data series contained in the MyDB table. This information is extracted from the QualityControlLevels table in the ODM. Figure A-7 is an example of the QualityControlLevelInformation section of the MyDB metadata file.

```
<QualityContolLevelInformation>
    <QualityControlLevel>1</QualityControlLevel>
    <Definition>Quality controlled data</Definition>
    <Explanation>Quality controlled data have passed quality assurance procedures such as
      routine estimation of timing and sensor calibration or visual inspection and removal of
      obvious errors. An example is USGS published streamflow records following parsing
      through USGS quality control procedures.</Explanation>
</QualityContolLevelInformation>
```

Figure A-7. Example MyDB metadata QualityControlLevelInformation section.

Offset Information

The OffsetInformation section of the metadata file contains all of the information describing the offset or offsets associated with a data series contained in the MyDB table. Since data series may have multiple offsets, the OffsetInformation section may contain information about more than one offset type. The ID within the <Offset> elements corresponds to the OffsetTypeID field in the Core MyDB table. Figure A-8 contains an example of the OffsetInformation section of the MyDB metadata file in which two different offsets, both of which apply to the same data series, are described.

```
<OffsetInformation>
    <Offset ID="1">
        <OffsetDescription>Below water surface</OffsetDescription>
        <OffsetUnits>
            <UnitsName>Meters</UnitsName>
            <UnitsType>Length</UnitsType>
            <UnitsAbbreviation>m</UnitsAbbreviation>
        </OffsetUnits>
    </Offset>
    <Offset ID="2">
        <OffsetDescription>Above reservoir bottom</OffsetDescription>
        <OffsetUnits>
            <UnitsName>Meters</UnitsName>
            <UnitsType>Length</UnitsType>
            <UnitsAbbreviation>m</UnitsAbbreviation>
        </OffsetUnits>
    </Offset>
</OffsetInformation>
```

Figure A-8. Example MyDB metadata OffsetInformation section.

Qualifier Information

The QualifierInformation section of the metadata file contains all of the information describing any data qualifying comments associated with a data series contained in the MyDB table. Since data series may have multiple data qualifying comments, the QualifierInformation section may contain information about more than one data qualifying comment. The ID within the <Qualifier> elements corresponds to the QualifierID field in the Core MyDB table. Figure A-9 contains an example of the QualifierInformation section of the MyDB metadata file in which two different data qualifying comments, both of which apply to the same data series, are described.

```
<QualifierInformation>
    <Qualifier ID="1">
        <QualifierCode>e</QualifierCode>
        <QualifierDescription>Value has been edited or estimated by USGS personnel and is
          write protected</QualifierDescription>
    </Qualifier>
    <Qualifier ID="2">
        <QualifierCode>I</QualifierCode>
        <QualifierDescription>Value has been interpolated from previous and next
          value</QualifierDescription>
    </Qualifier>
</QualifierInformation>
```

Figure A-9. Example MyDB metadata QualifierInformation section.

Sample and Lab Method Information

The final two sections of the metadata file contain the sample and lab method information for a data series contained within the MyDB table. A data series may be the result of one or more samples, and each of the samples associated with a data series is listed and described in the SampleInformation section. The ID in the <Sample> element corresponds to the SampleID in the Core MyDB table. Figure A-10 shows an example of the SampleInformation section of the metadata file where two separate grab samples are described. Information for the SampleInformation section of the metadata file is extracted from the Samples table of the ODM.

Each of the samples in Figure A-10 was analyzed using a laboratory method. In this case both samples were analyzed using a laboratory method with an ID of 1. The ID of the laboratory method is contained within the <LabMethodID> element of the sample description. The LabMethodInformation section of the metadata file contains descriptions of any laboratory methods that are associated with the samples contained within the SampleInformation section. In Figure A-10, both samples were analyzed using LabMethodID = 1, and a description of the method appears in the LabMethodInformation section. Information for the LabMethodInformation section of the metadata file is extracted from the LabMethods table of the ODM. If multiple laboratory methods are referenced in the SampleInformation section of the metadata file, the LabMethodInformation section would contain descriptions of each, and each laboratory method would be identified by its ID.

```
<SampleInformation>
   <Sample ID="1">
      <SampleType>Grab Sample</SampleType>
      <LabSampleCode>ABCD12345</LabSampleCode>
      <LabMethodID>1</LabMethodID>
   </Sample>
   <Sample ID="2">
      <SampleType>Grab Sample</SampleType>
      <LabSampleCode>ABCD12346</LabSampleCode>
      <LabMethodID>1</LabMethodID>
   </Sample>
</SampleInformation>
<LabMethodInformation>
   <LabMethod ID="1">
      <LabName>Utah Water Research Laboratory Environmental Quality
        Laboratory</LabName>
      <LabOrganization>Utah Water Research Laboratory</LabOrganization>
      <LabMethodName>EPA 160.2</LabMethodName>
      <LabMethodDescription>Residue, Non-Filterable (Gravimetric, Dried at 103-105
        Degrees C)</LabMethodDescription>
      <LabMethodLink>http://infotrek.er.usgs.gov/pls/apex/f?p=119:38:57922112629004
        40::::P38_METHOD_ID:5212</LabMethodLink>
   </LabMethod>
</LabMethodInformation>
```

Figure A-10. Example MyDB metadata SampleInformation and LabMethodInformation sections.

ODM Compatibility

MyDB is described in this document such that it is consistent with and compatible with the CUAHSI HIS ODM Version 1.0 (Tarboton et al., 2007), which is being released as part of the CUAHSI HIS Version 1.0 package currently under development by the CUAHSI HIS Team.

Appendix C
Rules for Editing and Deriving Data Series Stored in ODM 1.0 Using the ODM Tools Application

Introduction

The ODM Tools application has functionality that allows users to modify, delete, change, or otherwise make edits to data series stored within ODM. In addition, ODM Tools provides functionality to create derived data series, or data series that are calculated from data series already stored in ODM (i.e., calculate a time series of discharge from a time series of stage, or calculate a time series of daily average temperature from a time series of hourly observations). The purpose of this Appendix is to clarify how data series editing and creation are managed by the ODM Tools application within the ODM Version 1.0 schema.

Rules for Editing and Deriving Data Series in ODM Version 1.0

The ODM Tools application uses the following rules so that versioning of and edits to data series can be managed within the ODM 1.0 schema:

1. *Data series with a QualityControlLevelID of 0 cannot be edited* – Level 0 data series represent raw data from sensors (i.e., stage measured by a water level recorder) or other products derived from raw data (i.e., discharge that is programmatically derived from stage before the stage values have been quality controlled). By definition, Level 0 data have not been quality controlled and may contain significant errors and spurious values. However, Level 0 data series represent the source from which all other derived data series are based, and as such should remain intact for archive purposes. Level 0 data series should not be used for analysis unless no other adequate options are available, and only if the user is aware that the data are raw. Level 0 data series can be removed entirely from the database using the delete functionality of ODM Tools, but only by removing the entire data series.
2. *Only one QualityControlLevelID 0 data series can exist for a Site, Variable, Method, and Source combination* – Only one raw data series for a Site, Variable, Method, and Source combination can exist within an ODM database. If multiple sensors are measuring the same variable at the same site, the method description would have to distinguish between the two (i.e., the method would be different).
3. *Only one QualityControlLevelID 1 data series can exist for each Site, Variable, Method, and Source combination* – Once a Level 0 (raw) data series has been loaded to the database, a Level 1 (quality controlled) data series can be "derived" from that Level 0 data series. ODM Tools does this by making a copy of the Level 0 data series, changing the QualityControlLevelID of the copy to 1. Once the Level 1 data series has been created, any necessary filtering or editing must be done so that the Level 1 data series is acceptable as quality controlled. In most cases, the majority of the values within a Level 0 data series and its corresponding Level 1 data series will remain the same. However,

where instruments malfunction or other conditions are present that affect the raw data values, Level 1 values may be deleted, adjusted, or otherwise edited so that the Level 1 data series is acceptable for use in analyses.

4. *Any edits to a data series are saved to that data series* – Level 0 data cannot be edited. With data series of Level 1 or higher, however, users will be allowed to edit and delete values. Each time an edit is made, the result overwrites the previous value within a data series. In other words, edits do not create new data series, they modify an existing one. This will be true even where edits are done within multiple editing sessions.
5. *Data series of QualityControlLevelID 2 or higher can only be created from data series of QualityControlLevelID 1 or higher* – Derived data series of Level 2 or higher can only be created from data series of Level 1 or higher. If a user wishes to create a derived data series from a Level 0 data series (such as raw discharge from raw, unchecked stage values) that derived data series will also be Level 0 because no quality assurance/quality control has been performed).