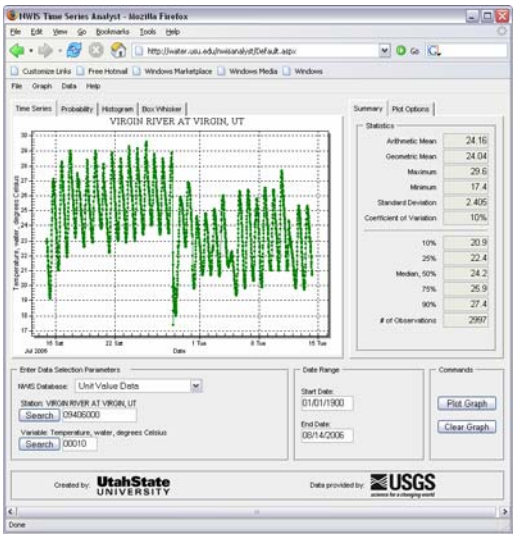
Time Series Analyst Tutorial



Jeff Horsburgh
Utah Water Research Laboratory
Utah State University
8200 Old Main Hill
Logan, UT 84322-8200
(435) 797-2946
jeff.horsburgh@usu.edu

8-14-2006

Introduction

The Time Series Analyst application, developed at Utah State University, was designed to provide users with plotting and export functionality for data at any United States Geological Survey (USGS) monitoring station in the United States. It implements the NWIS Web Services (http://water.sdsc.edu/wateroneflow/) designed and developed as part of the CUAHSI HIS project to provide access to the USGS data contained within the National Water Information System (NWIS).

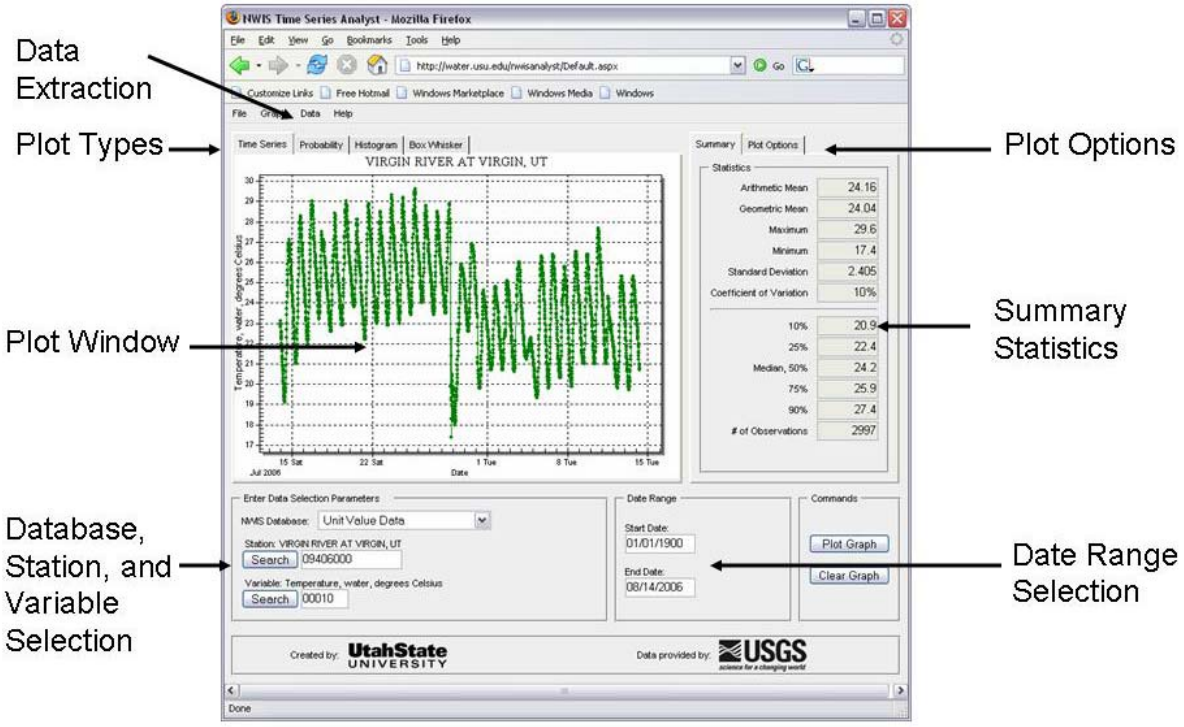
Opening the Time Series Analyst

The Time Series Analyst can be accessed by connecting your Internet browser to the following URL: http://water.usu.edu/nwisanalyst/

The Time Series Analyst is operated by selecting one of the NWIS databases, a USGS station, and a variable for which to retrieve data and then clicking on the "Plot Graph" button at the bottom right part of the Time Series Analyst Interface.

Parts of the Time Series Analyst Interface

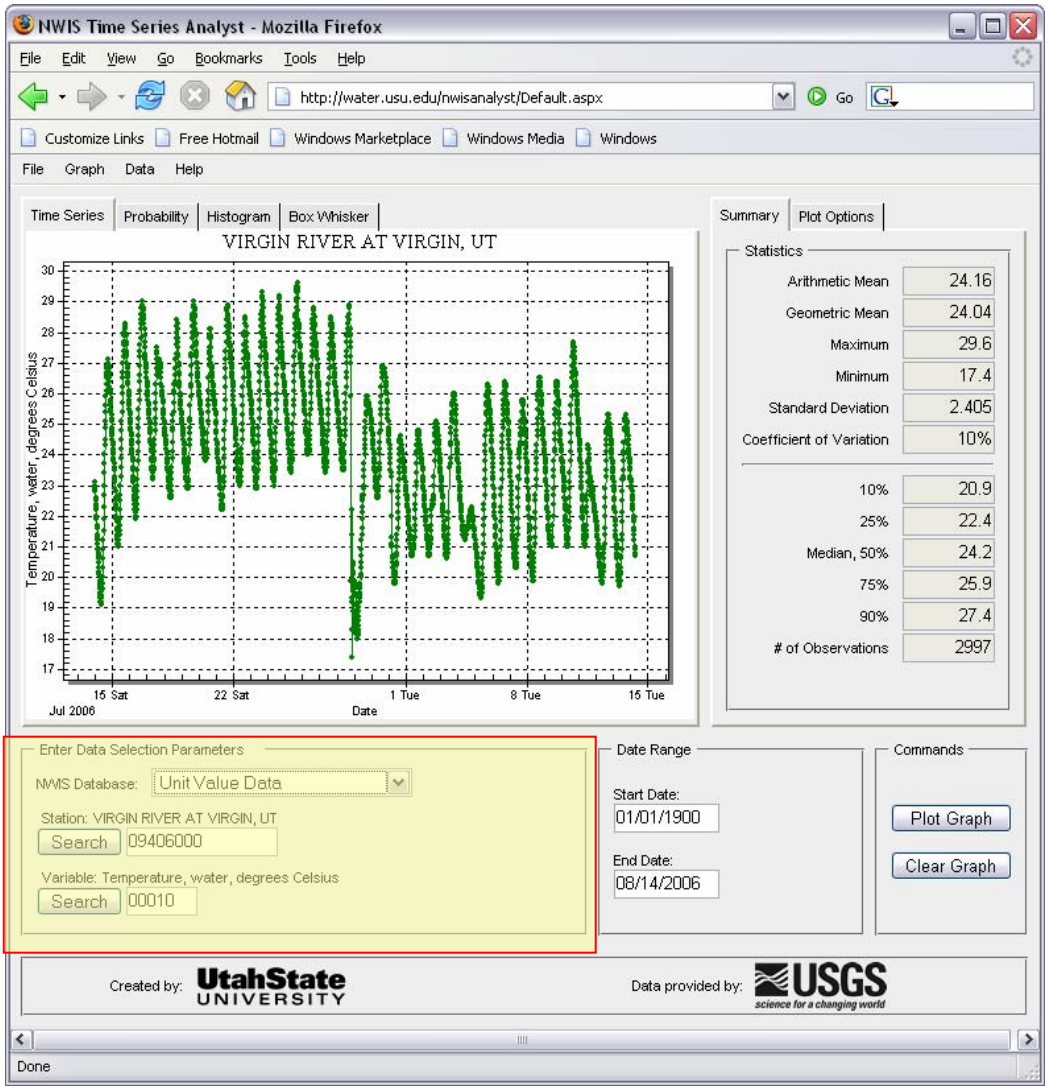
The figure below shows the different parts of the Time Series Analyst interface. The following sections describe each of these in more detail and describe how each is used.



Database, Station, and Variable Selection

Data is retrieved by the Time Series Analyst after a database, station, and parameter selection is made in the boxes at the lower left of the Time Series Analyst interface (see the figure below) and the "Plot Graph" button is clicked. The Time Series Analyst can retrieve daily value data (i.e., daily streamflow values), groundwater level data, unit value or realtime data, and instantaneous irregular data (i.e., water quality values) for any station in the USGS NWIS system (http://waterdata.usgs.gov/nwis). The type of values to be retrieved is set in the "NWIS Database" dropdown menu.

Stations are selected by typing the USGS station number into the Station text box. To retrieve daily streamflow data for a selected station, select "Daily Value Data" from the "NWIS Database" dropdown menu, type the station number into the "Station" text box, type "00060" into the "Variable" text box (00060 is the USGS Parameter code for discharge), and then click on the "Plot Graph" button. To retrieve data for a water quality parameter, select "Instantaneous Irregular Data" from the "NWIS Database" dropdown menu, type the station number into the "Station" text box, type the six digit USGS parameter code for the desired water quality variable into the "Variable" text box, and then click the "Plot Graph" button.



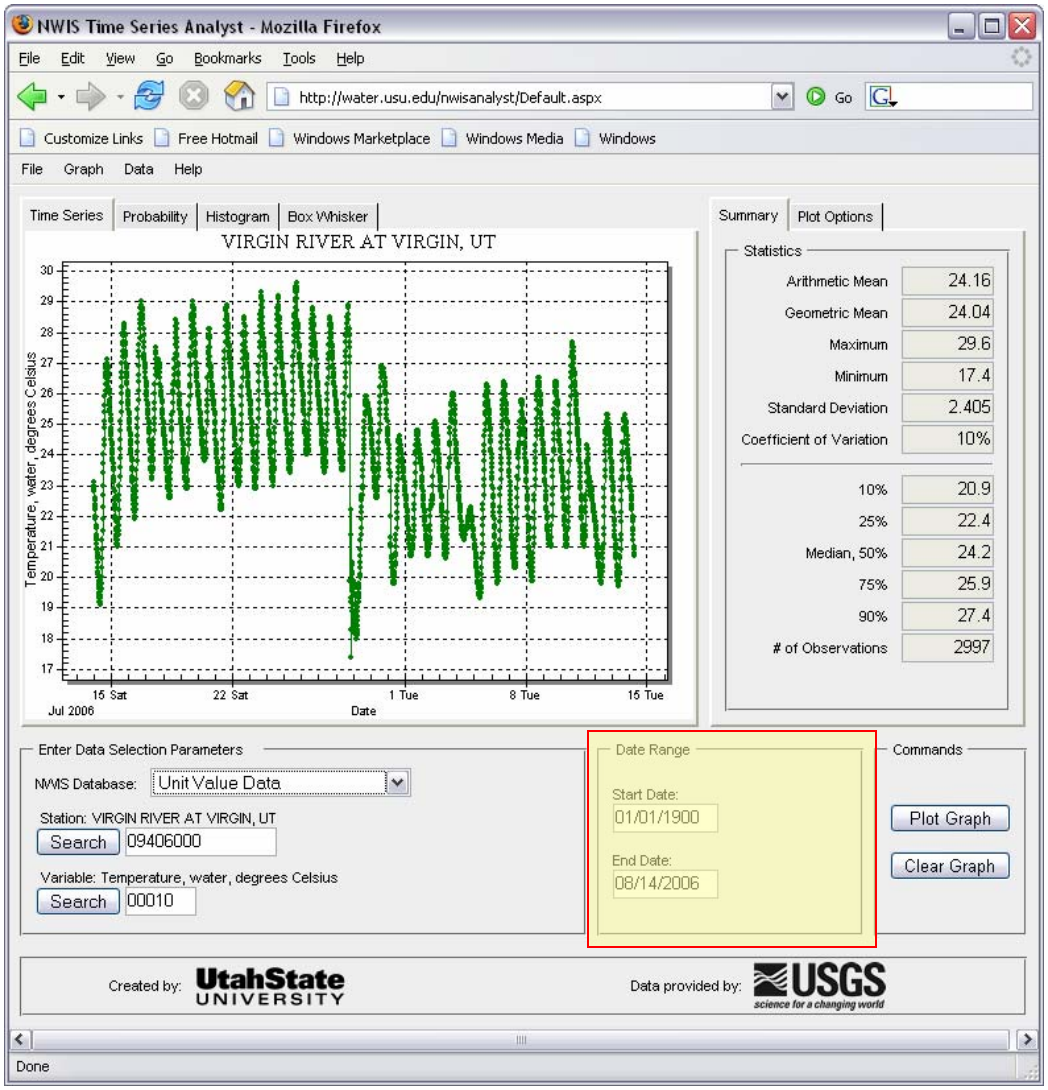
A complete list of the USGS NWIS parameter codes can be found at: http://nwis.waterdata.usgs.gov/usa/nwis/pmcodes. Common USGS parameter codes to try include:

00010 - Temperature, water, degrees Celsius
00060 - Discharge, cubic feet per second
00300 - Dissolved oxygen, water, unfiltered, milligrams per liter
00400 - pH, water, unfiltered, field, standard units

If no data are available for a database/station/variable combination, the Time Series Analyst will display a blank plot window and a "No Data Available" message will be displayed in the summary statistics tab.

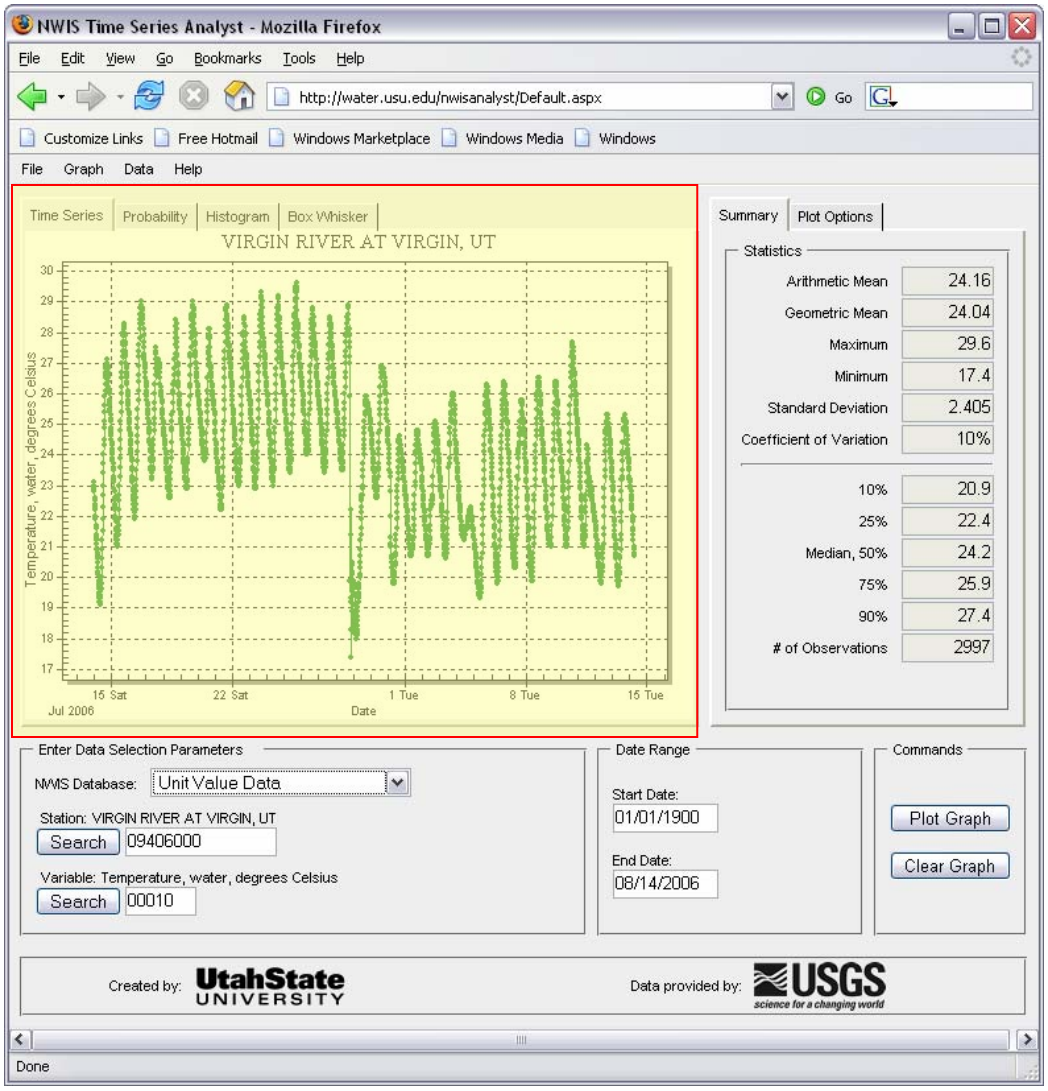
Date Range Selection

The date range of the data returned can be limited by typing a start date and an end date into the Date Range selection boxes. These dates should be input using the format MM/DD/YYYY where MM is the two digit month, DD is the two digit day of the month, and YYYY is the 4 digit year. If no date range is input by the user, the Time Series Analyst defaults to 01/01/1900 to the present date and returns all available data within that period. The following figure highlights the Date Range Selection area of the Time Series Analyst.



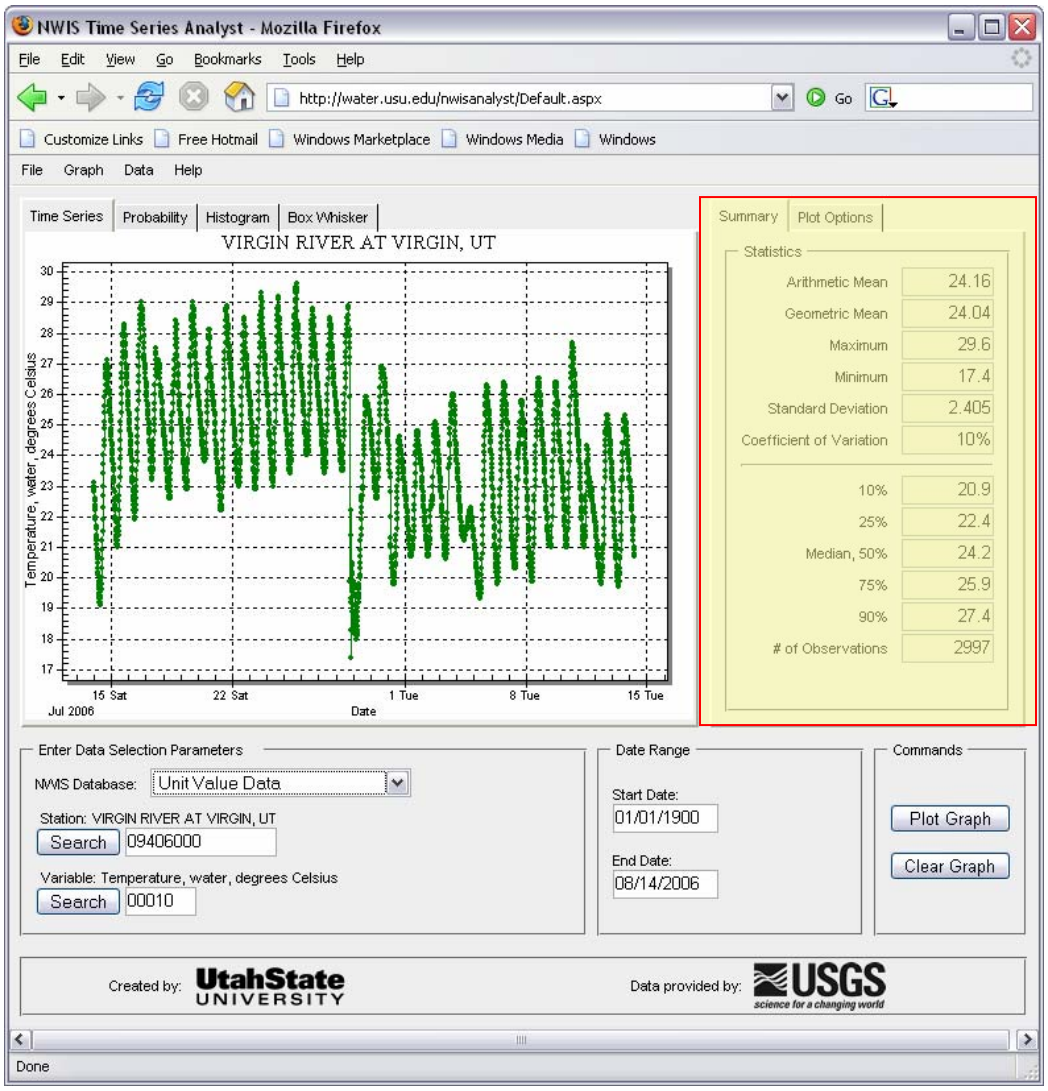
Plot Window and Plot Type Selection

The plot window is the area in which the plots are drawn. The Time Series Analyst currently includes several different plot types, including time series, probability, histogram, and monthly, seasonal, annual, and overall box and whisker plots. Different plot types can be selected by clicking on the tabs at the top of the plot window as shown in the following figure.



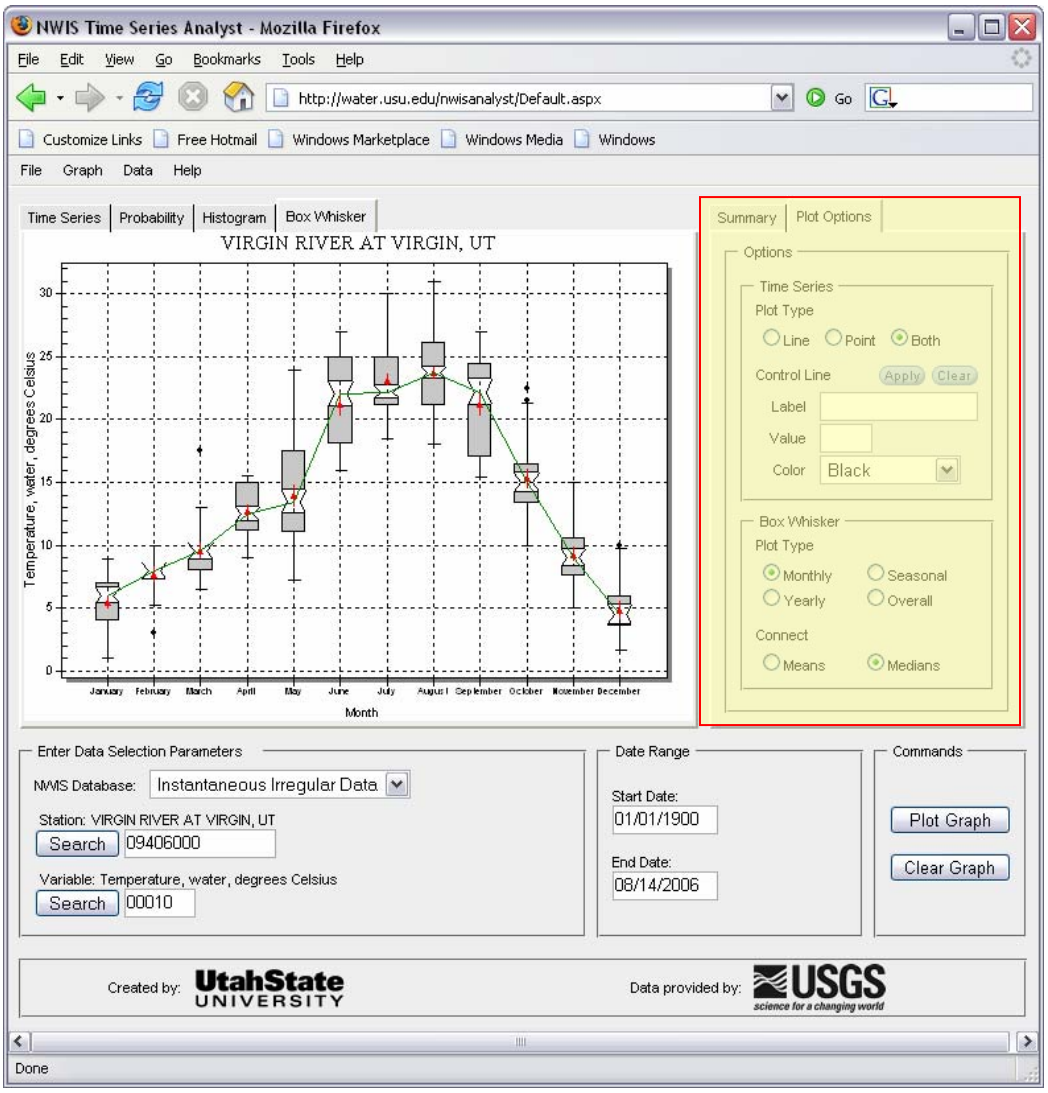
Summary Statistics

The Time Series Analyst calculates some simple descriptive statistics for the time series that is returned after a database, station, and parameter are selected. These statistics include the arithmetic mean, geometric mean, maximum value, minimum value, standard deviation, coefficient of variation, number of observations, and several percentile values. The summary statistics are shown in the tab at the right of the Time Series Analyst interface and are automatically calculated when the data are received. The following figure highlights the summary statistics tab.



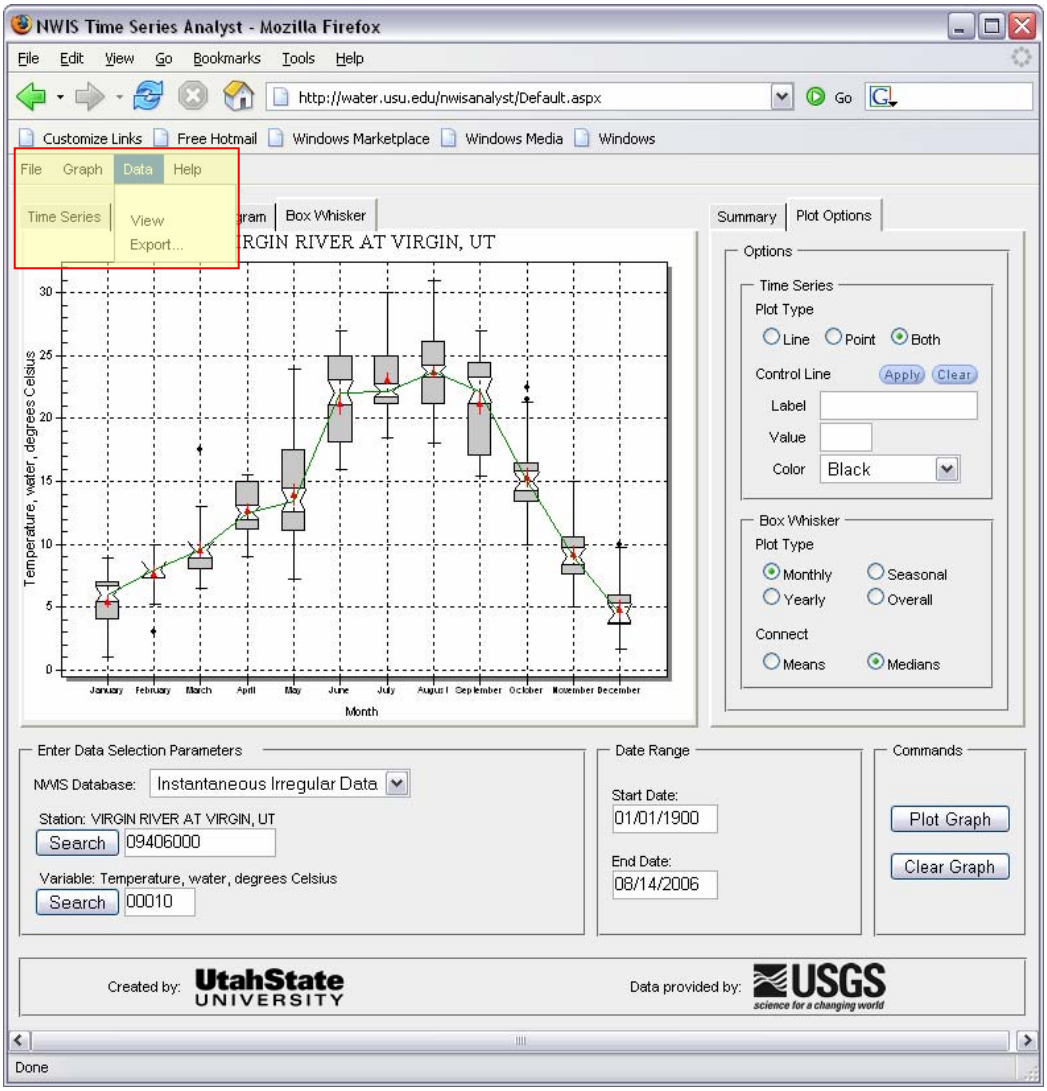
Plot Options

The Time Series Analyst has several options for customizing the plots that are produced. These options are available in the Plot Options tab at the right of the Time Series Analyst interface. Currently available plot options include changing the time series plots to show points, lines, or both, adding a control line (such as a water quality standard value) to the time series plot, changing the period of the box and whisker plot from monthly to seasonal, annual, or overall, and specifying whether to connect the mean or median values in the box and whisker plots. The following figure highlights the Plot Options tab.



Data and Plot Extraction and Printing

The data retrieved by the Time Series Analyst can be exported to an HTML file or to a Microsoft Excel file. This is accomplished by clicking on the "Data" pull-down menu at the top of the Time Series Analyst and choosing "View" to export to HTML or "Export" to export to an Excel file. A print version of the plot and descriptive statistics can also be generated as an HTML document by clicking on the "File" pull-down menu at the top of the Time Series Analyst and selecting "Print Version."

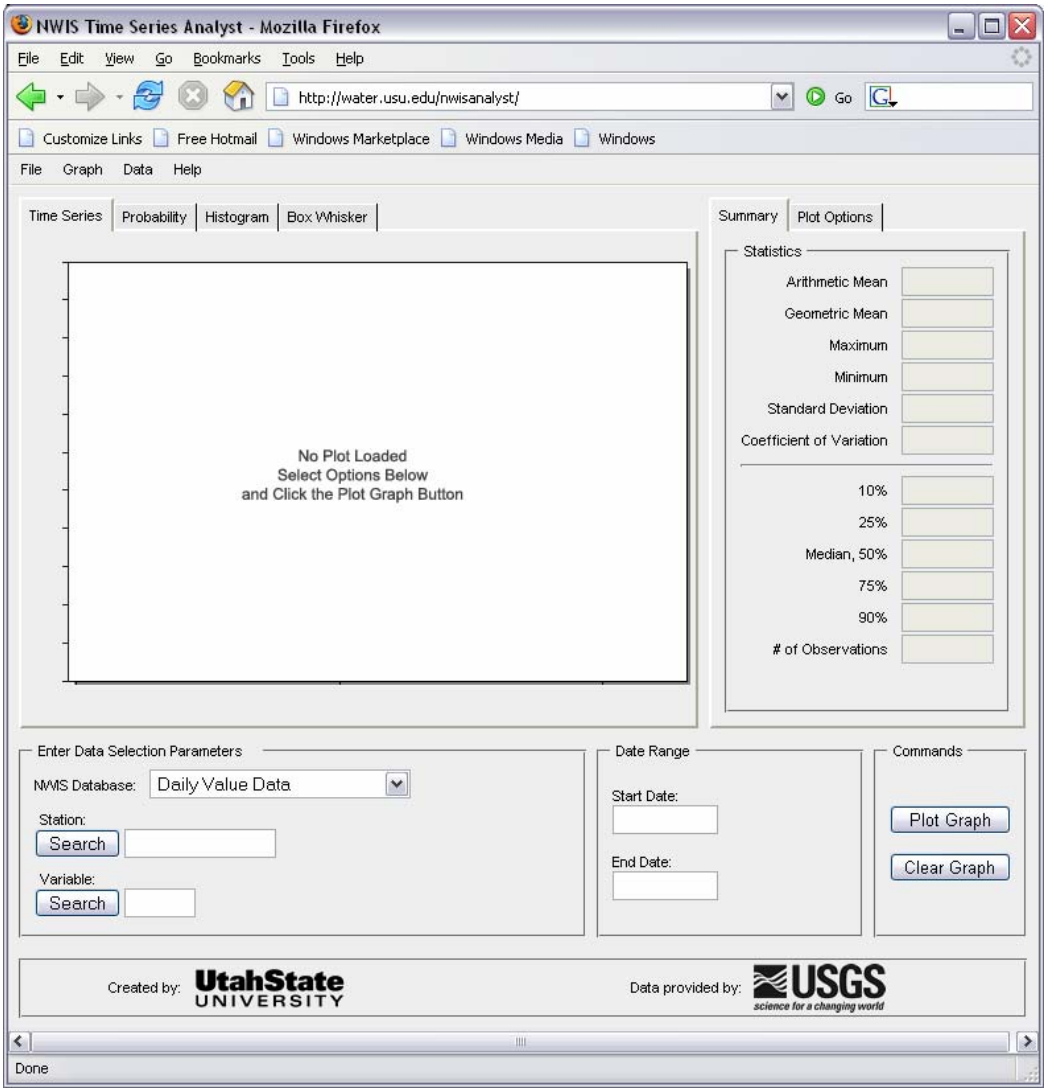


Step by Step Example

Open the Time Series Analyst by connecting your internet browser to the following link:

http://water.usu.edu/nwisanalyst/

Your browser should look similar to the following figure.



Retrieving and Plotting Daily Streamflow Data

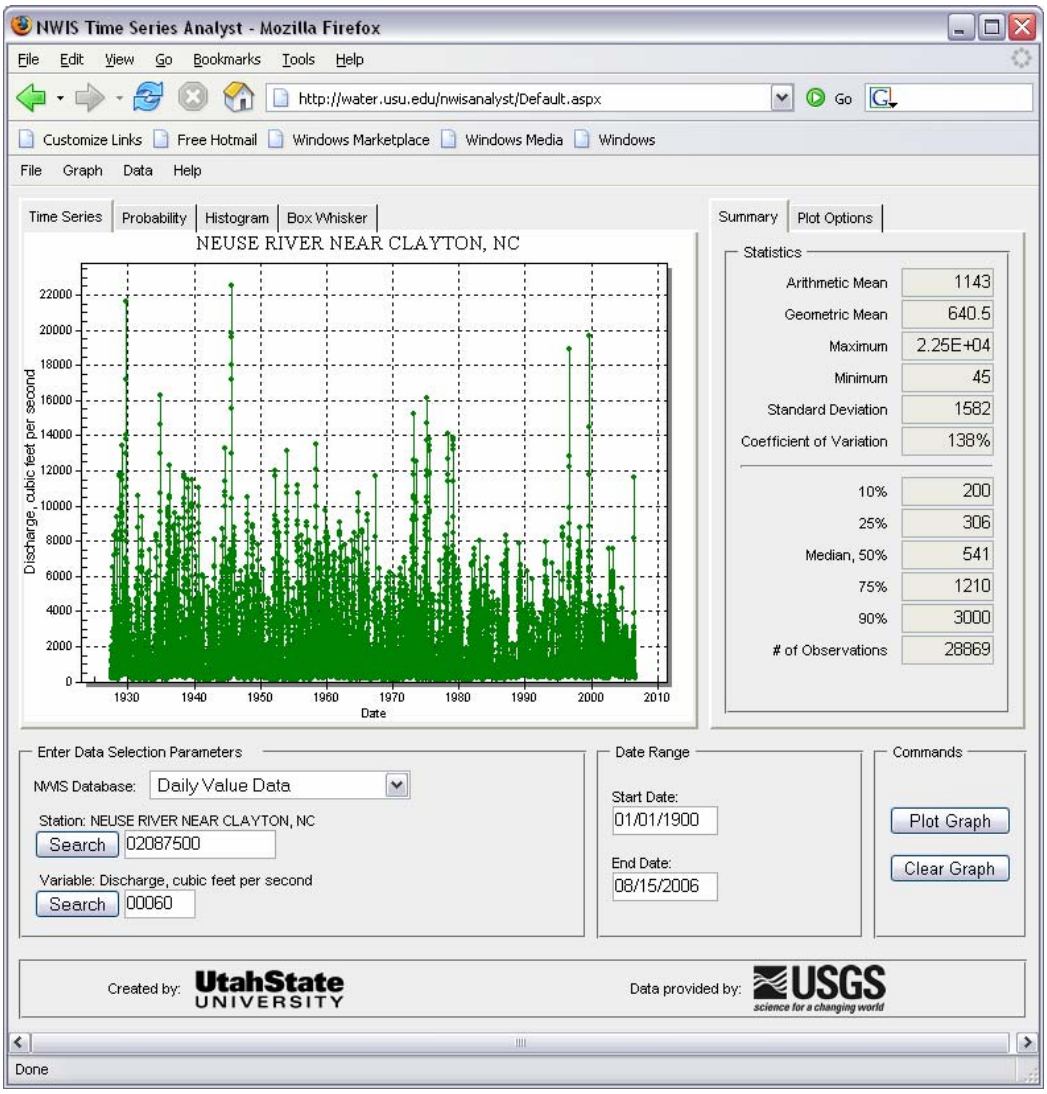
For this example, we will be looking at data for the Neuse River near Clayton, North Carolina. The USGS station identifier for this sampling location is 02087500.

NOTE: To search NWIS for stations and station identifiers click on the "Search" button next to the "Station" input text box.

We will begin by looking at the available daily streamflow record for this station. Make sure that "Daily Value Data" is selected in the "NWIS Database" dropdown menu, and then type the station identifier into the "Station" text box. Next, type the USGS parameter code for discharge (00060) into the "Variable" text box. We will retrieve the entire period of record for daily flows at this station, so we will leave the "Start Date" and "End Date" boxes empty.

NOTE: You can restrict the date range of data retrieved by the Time Series Analyst by typing a Start Date and an End Date into the Date Range boxes. Dates should be entered in the format MM/DD/YYYY where MM is the two digit month, DD is the two digit day, and YYYY is the four digit year.

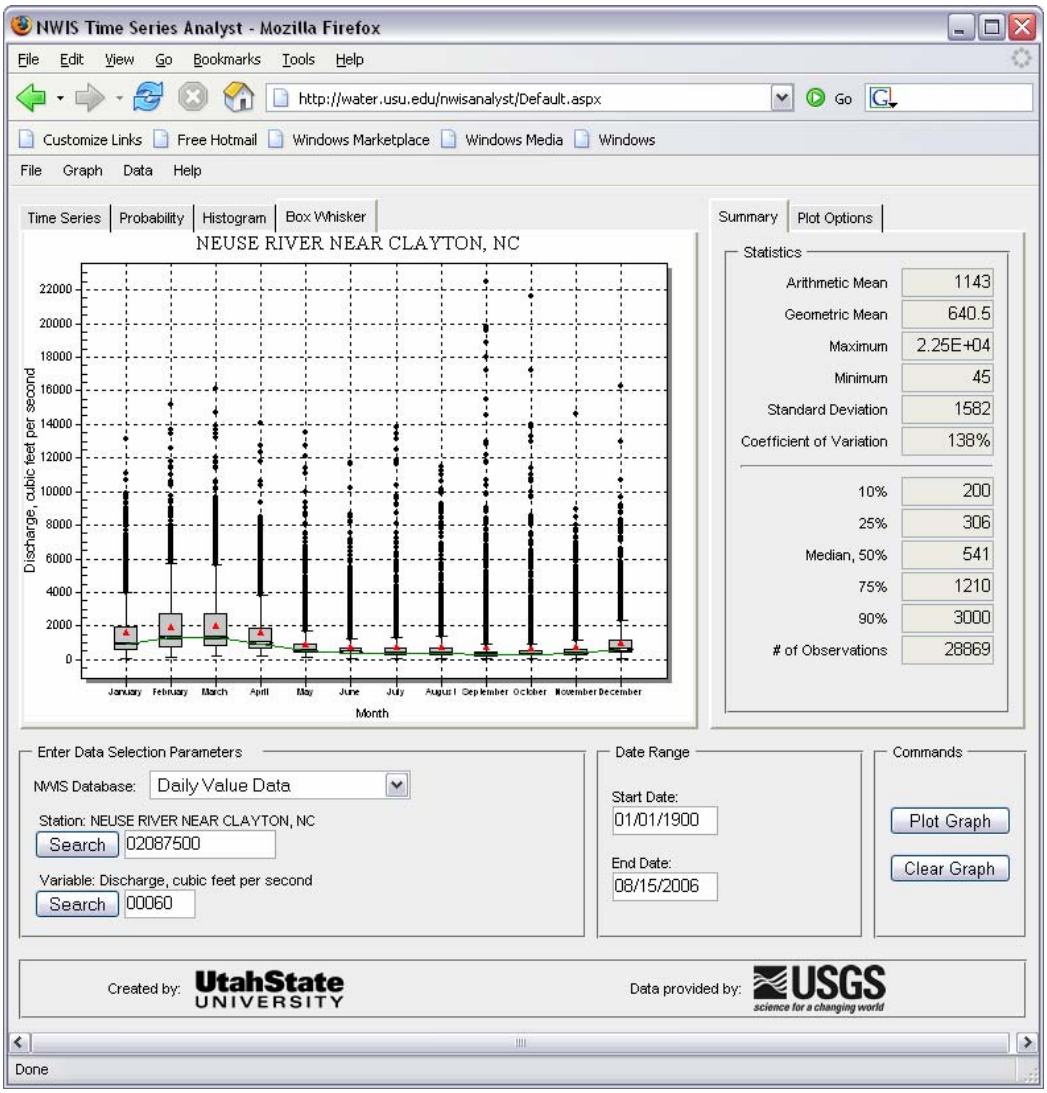
Once you have selected the database, station, and variable, click on the "Plot Graph" button at the lower right of the window. When you have clicked on the Plot Graph button, the Time Series Analyst will retrieve the data that you have requested. After a few moments, your window should look like the following.



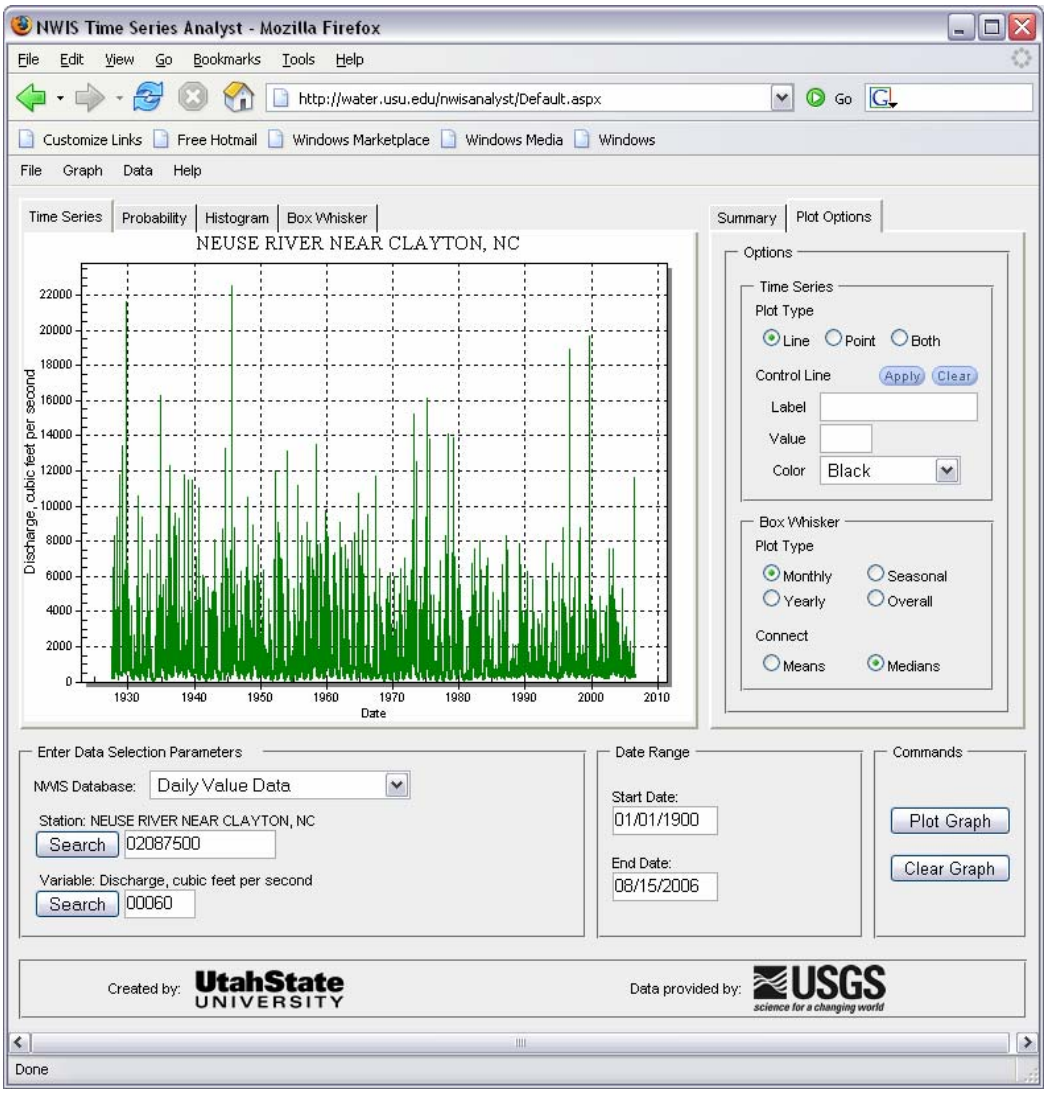
You will notice that the descriptive statistics in the "Summary" tab have automatically been populated based on the dataset that you have retrieved (see the right side of the window). These statistics will be dynamically generated each time you make a new data selection and click the "Plot Graph" button. The average flow at this station is 1,143 cfs, with a minimum of 45 and a maximum of 22,500 cfs.

Plot Types and Plot Options

You can select a new plot type by clicking on one of the tabs at the top of the plot window. If you click on the "Box Whisker" tab, you should see the following:

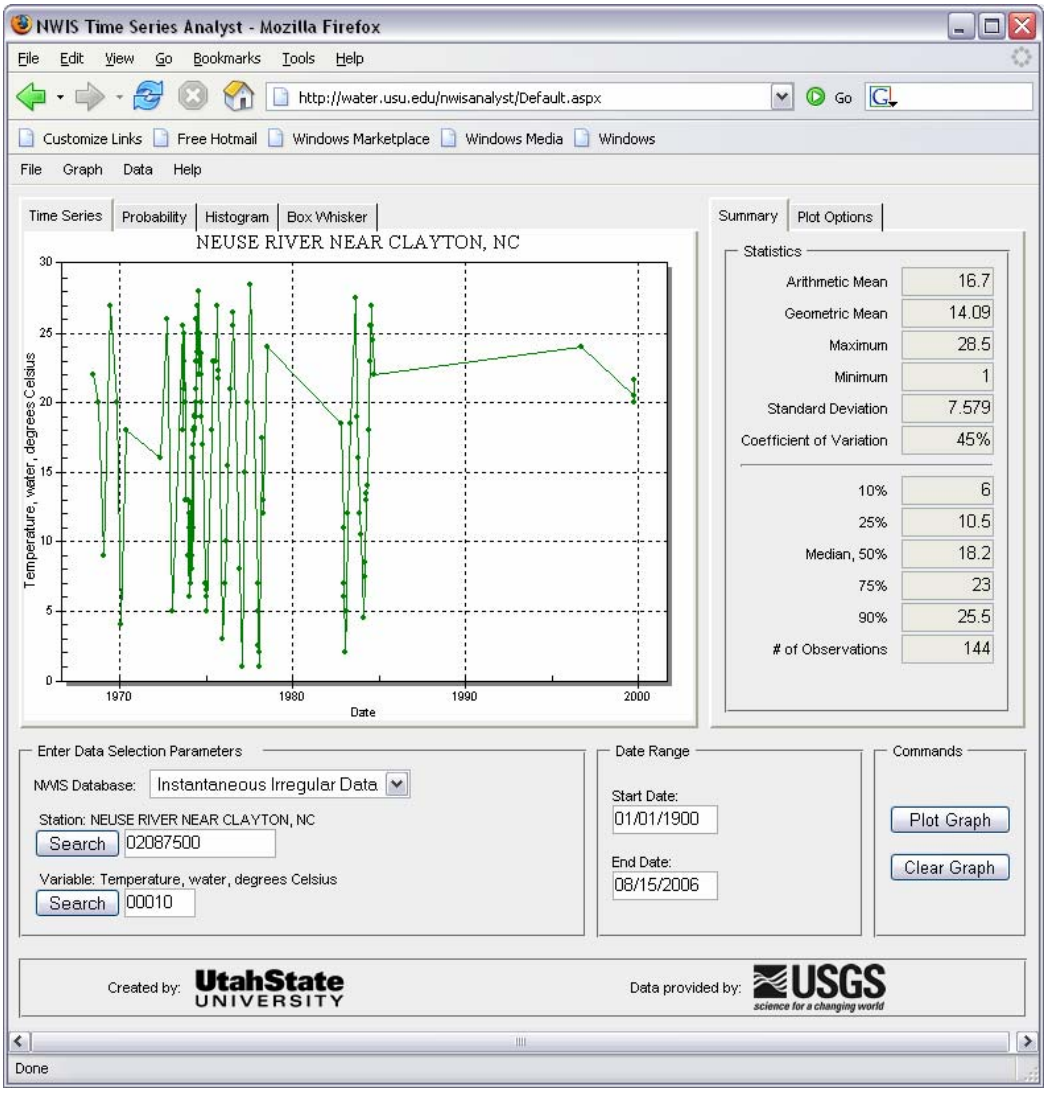


The "Plot Options" tab at the top right of the window allows you to customize to some degree the look and feel of the plots that are produced. For example, by clicking the line only feature in the Time Series plot options and then returning to the "Time Series" plot tab your window should look like the following:



Retrieving and Plotting Water Quality Data

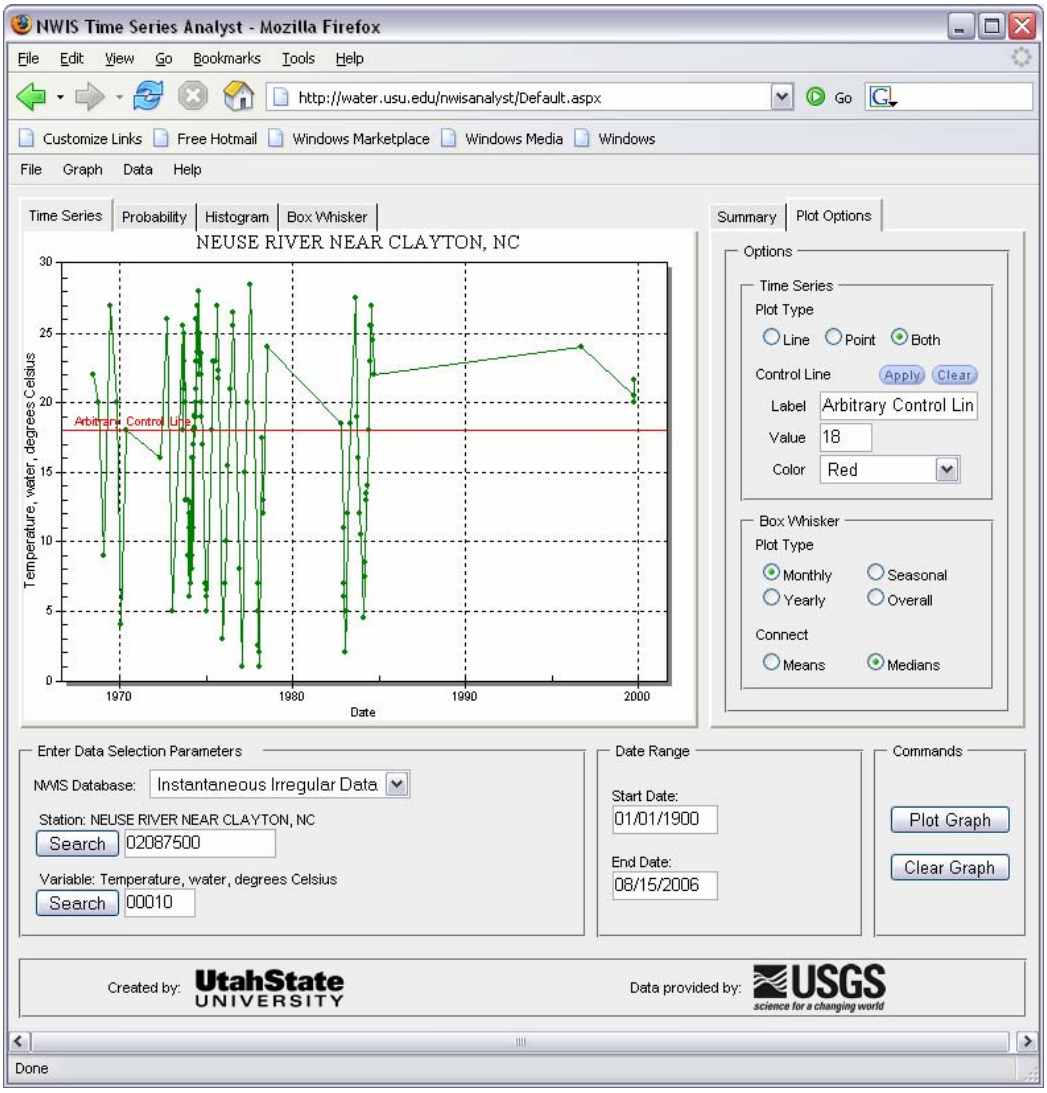
Now we will look to see if there are any water temperature data available at this location. Select "Instantaneous Irregular Data" from the "NWIS Database" dropdown menu and then type "00010" into the "Variable" text box. The USGS parameter code for water temperature in degrees Celsius is 00010. Click the Plot Graph button, and your window should look like the following:



NOTE: A complete list of the USGS NWIS water quality parameter codes can be found at: http://waterdata.usgs.gov/nwis/pmcodes/. You can access this search page by clicking on the "Search" button next to the "Variable" text box. Common parameter codes to try include:

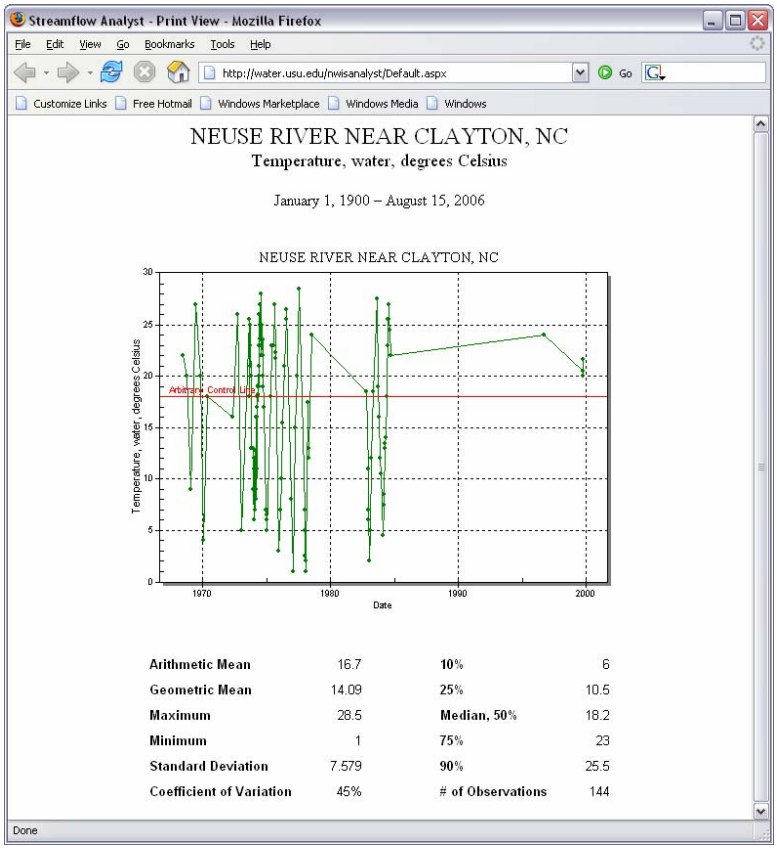
00010 - Temperature, water, degrees Celsius
00300 - Dissolved oxygen, water, unfiltered, milligrams per liter
00400 - pH, water, unfiltered, field, standard units

Using the Plot Options tab you can add control lines to the plot. This is useful if you wish to plot numeric criteria values on the plot with your data. Click on the "Plot Options" tab and in the "Label" box below the "Control Line" label give your control line a label by tying it into this box. Select a value for the control line by typing it into the "Value" box and select a color for the line by choosing one from the "Color" drop down. When you are finished, click the "Apply" button next to the Control Line label and your plot should look similar to the following. The "Clear" button will clear the control line from your plot.

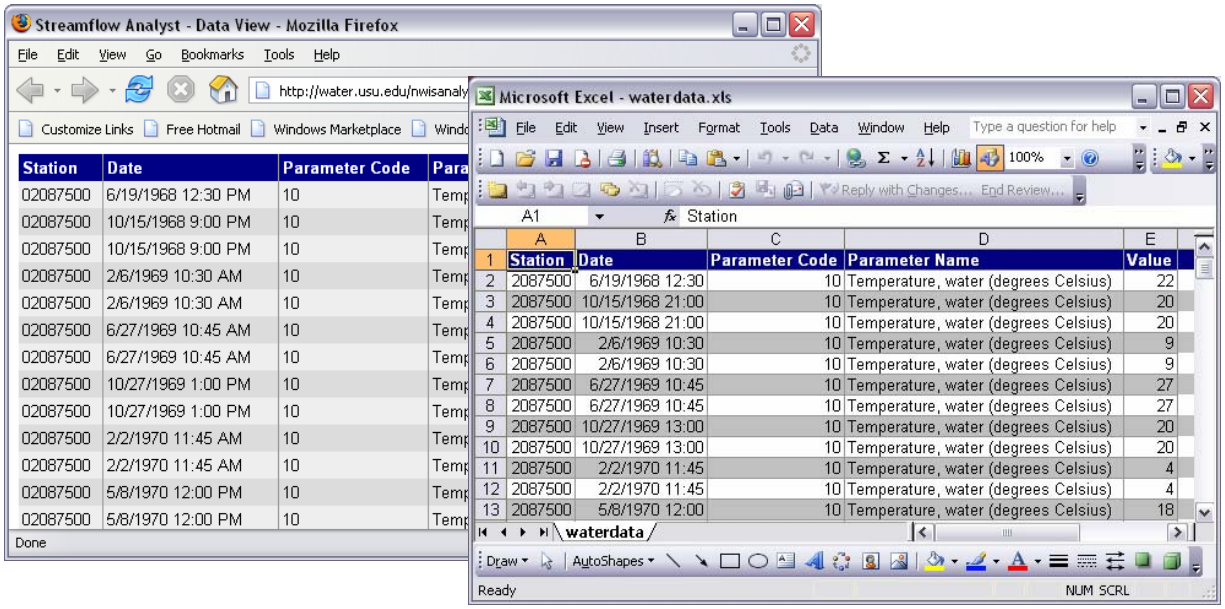


Exporting Plots and Data

If you would like to export the plot, you can either right click on the plot and select "Copy Image" to export the image to the clipboard, or you can choose "Print Version" from the "File" menu at the top of the window to create an HTML page with the plot and the descriptive statistics that can be printed. The following is an example of the Print Version:



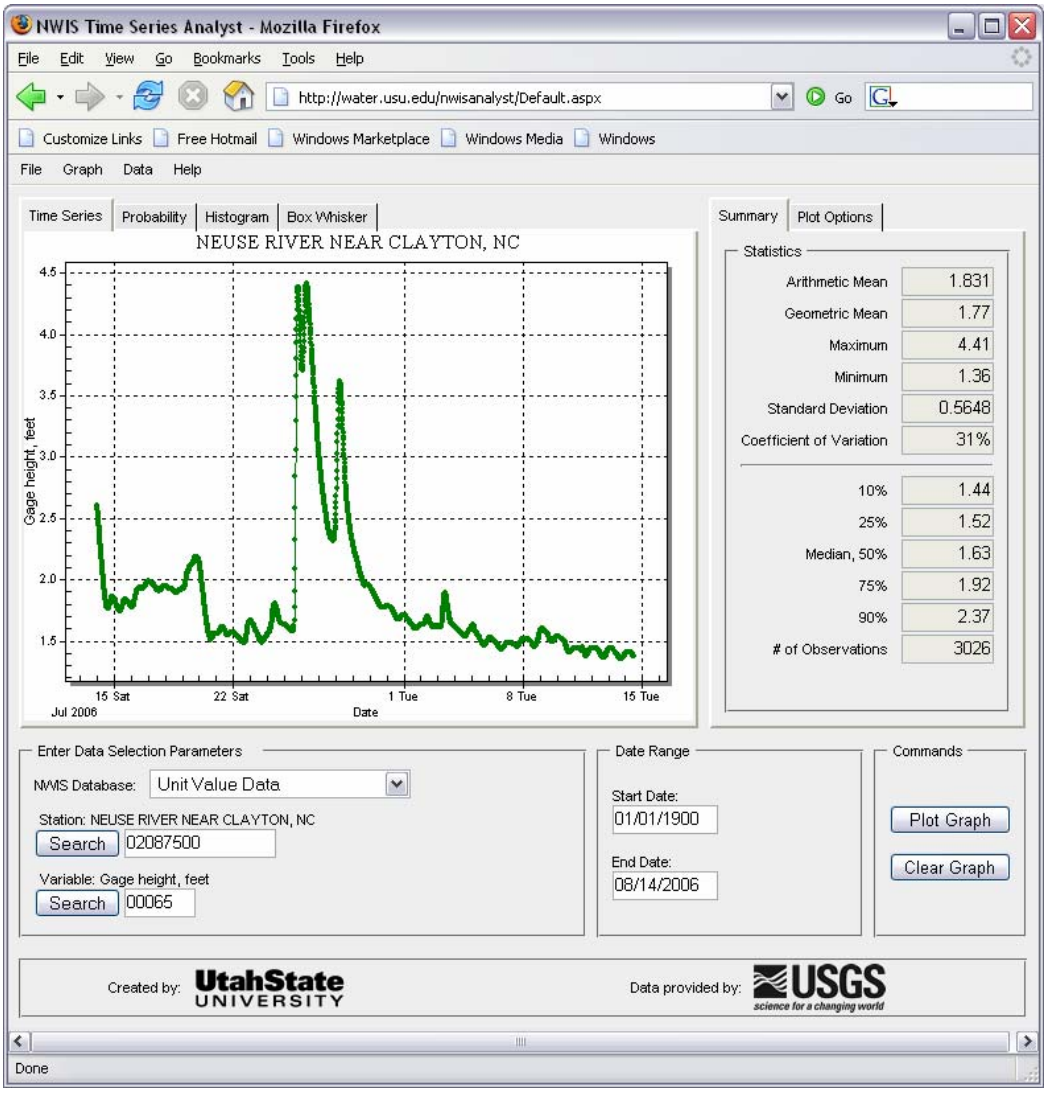
If you wish to view or export the data values, you can choose “View” from the “Data” menu at the top of the window to get an HTML formatted list of the data shown in the plot, or you can choose “Export” from the “Data” menu to export the data to a Microsoft Excel file. These files look similar to the following:



Retrieving and Plotting Realtime Data

The station that we are using for this example, 02087500 - Neuse River near Clayton, North Carolina, just happens to be a USGS realtime station. Using the NWIS Time Series Analyst we can access up to the last 31 days of real time data at this station.

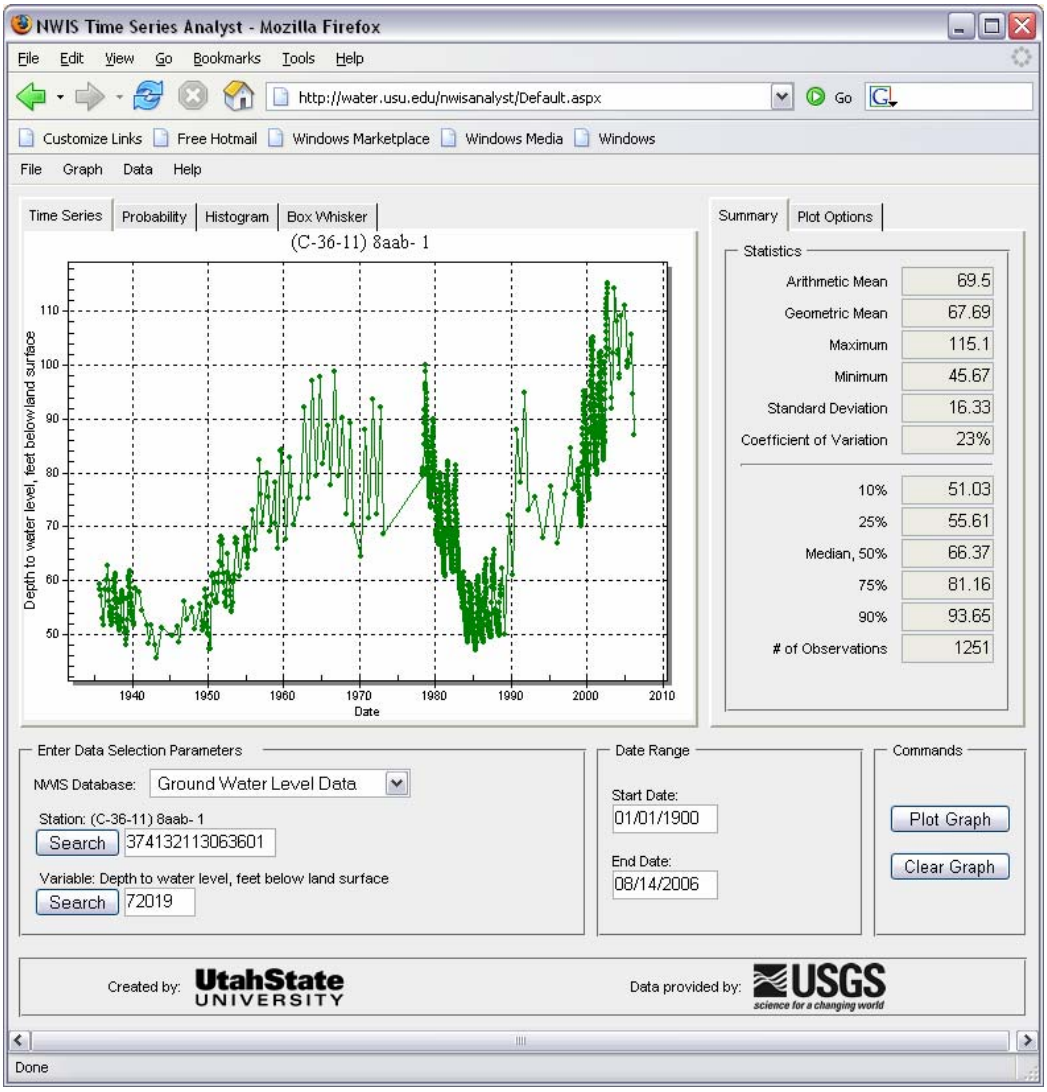
Select "Unit Value Data" from the "NWIS Database" dropdown menu, makes sure that "02087500" is still in the "Station" input text box, type the code "00065" into the "Variable" input text box, and then click the "Plot Graph" button. Your screen should look something like the following. Keep in mind that your request will retrieve the most recent data, so your results will not be exactly like the following figure.



In addition to gage height (00065), you can request real time data for discharge (00060) at this station. Many USGS real time stations have additional realtime variables such as water temperature (00010), turbidity (00076), dissolved oxygen (00300), and specific conductance (00096). A USGS station with all of these variables is 09406000 – Virgin River at Virgin, UT.

Retrieving and Plotting Groundwater Level Data

Historical groundwater levels can be retrieved for thousands of stations located throughout the country. Since the Neuse River station that we have been using for our example represents a river, it does not have groundwater level data available. For the purposes of this example, we will use an arbitrary station from NWIS – 374132113063601. Groundwater level data are accessed using a parameter code of 72019, which is the USGS parameter code for depth to water level in feet below land surface. Select "Ground Water Level Data" from the NWIS Database dropdown menu, type the station number into the "Station" input text box (374132113063601), type the variable code for water level into the "Variable" input text box (72019), and then click the "Plot Graph" button. Your result should look like the following:



NOTE: Many USGS wells collect water level data in realtime. These data can be accessed through the "Unit Value Data" rather than the "Ground Water Level Data" option of the "NWIS Database" dropdown menu.

About the Time Series Analyst

Technical Specifications

An Internet browser is the only client software required to operate the core functionality of the Time Series Analyst. Users do not have to download or install any special software. Time series data downloads are formatted as Microsoft Excel files, and so Excel is required to read the data files exported from the Time Series Analyst. The Time Series Analyst has been tested on Windows operating systems using Internet Explorer and Mozilla Firefox browsers.

The Time Series Analyst server side application was developed using Microsoft ASP .Net technology in the Microsoft Visual Studio .Net development environment. It is running on a Microsoft Windows 2003 server and requires Microsoft IIS and the Microsoft .Net Framework to be installed on the server. The data plots are generated using a 3rd party plotting control called ProEssentials that has been developed by Gigasoft (http://www.gigasoft.com). It implements the NWIS Web Services (http://water.sdsc.edu/wateroneflow) designed and developed as part of the CUAHSI HIS project to provide access to the USGS data contained within the National Water Information System (NWIS). For more information about the Time Series Analyst, see the description file.

Contact Information

The Time Series Analyst was developed by the Environmental Management Research Group, which is a research unit of the Utah Water Research Laboratory at Utah State University. For more information contact:

Jeff Horsburgh
Utah Water Research Laboratory
Utah State University
Logan, UT 84322-8200
(435) 797-2946
jeff.horsburgh@usu.edu