# Using the NWIS Snapshot Tool to Analyze Surface Water in the Los Angeles River Watershed

A User Experience

Prepared By: David Navar

**Introduction**

In 2010, the USGS Community of Data Integration (CDI) started a program to develop software to link the data accumulated in the National Water Information System (NWIS) with the Geographic Informatics of ArcGIS (Holl).  The result is the NWIS Snapshot tool which is an Add-in to the Arc Map Software.  I will learn the software and use it to find data to analyze the Surface water characteristics of the Los Angeles HUC 8 watershed in Southern California.  I will provide an overview of the process I used, and comment on my experience utilizing the data retrieved through the tool in ArcGIS.

**Overview of the Snapshot Tool**

The NWIS Snapshot tool is a powerful data retrieval method which allows an ArcGIS user to retrieve data from 4 NWIS Data Service systems, and populate a geodatabase in the ArcGIS software.  The NWIS Data service system is a robust conglomeration of collected data across all 50 States.  Approximately 1.5 million USGS sites populate North America with hundreds of millions of data points (Holl).  The Snapshot tool is a user interface which queries data from four USGS NWIS Data services using an Extensible Markup Language (XML) to access the data stores in the USGS website. These services are accessed through the tabs towards the top of the snapshot tool.  The four available services are defined as follows:

Sites Service - This service geographically displays the locations of USGS measuring sites in the limits of an ArcGIS base map window.  The service populates a layer in ArcGIS with site attributes as well.   Within the window, the user may select the desired sites through the ArcGIS software selection tool.  Once desired sites are selected the user can then utilize the other three services to find desired data.

Water Quality Web Services-   Water quality data is kept as instantaneous irregular data, this means there is a data value for a single instant in time.  For a given time the USGS site and time, USGS has many parameter values available, although the measurements are few and far between.

Daily values service -This service provides historical statistical daily measurements for specified parameters and time periods.  An example of a daily data point is an average value over a day's multiple measurements.  The data spans back to dates in the 1800's and is available in a variety of formats.

Unit Values service- This service is also called the Instantaneous Values Web Service and provides instantaneous data for USGS sites.  The most recent data for continuous time series

observations is available through this service.  For example if one wanted to know the discharge rate of a river at this instant, this service would be utilized. Typically the data is available in 15 minute increments and is available within the USGS database for a length of 120 days (USGS Water Service).

**Study Area**

Surface water for the Los Angeles River HUC 8 Watershed in Southern California is the object of study.  Mountain ranges cover the north of the watershed and a highly urbanized coastal plain resides in the south of the watershed.  I chose this watershed because I predicted that highly urban coastal plains in the south of the watershed would provide interesting data characteristics.
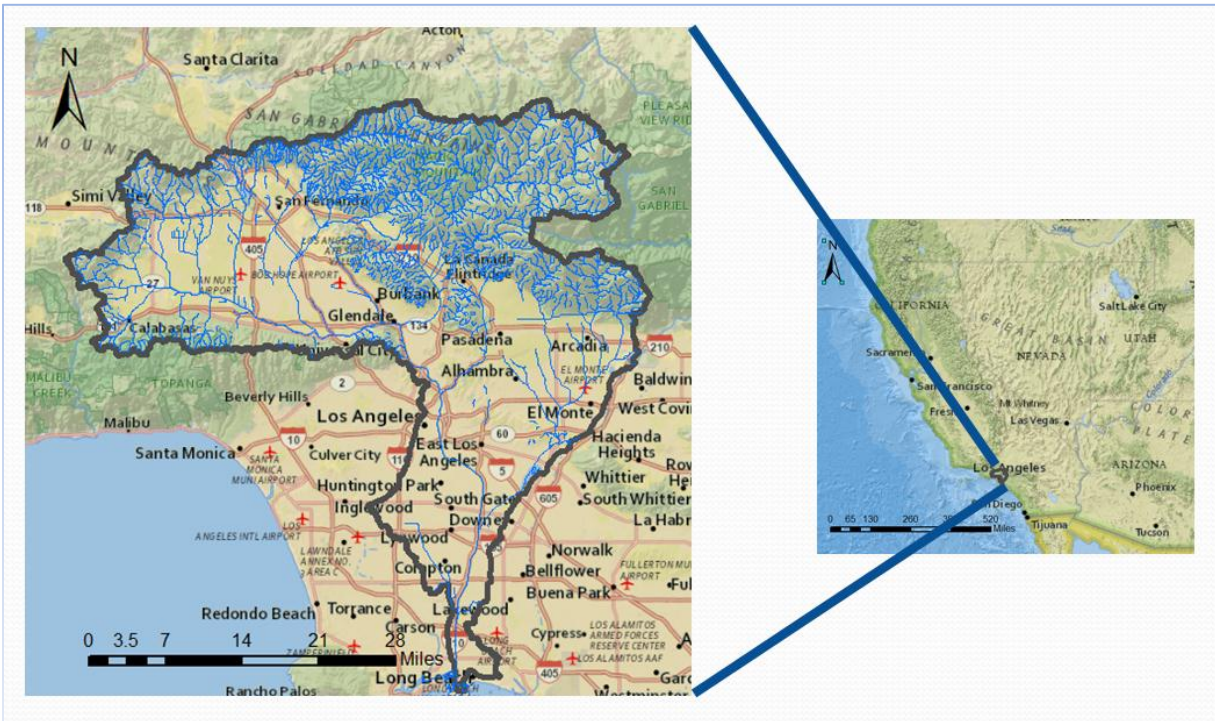


Figure 1: Location of the Los Angeles River Watershed

**Methods**

Each data service that the NWIS snapshot tool offers was utilized to analyze hydrologic characteristics of the Los Angeles River Watershed.  I did a water quality analysis for dissolved oxygen, a daily discharge value analysis and a unit value discharge analysis. Each of the three data service analyses consisted of the same steps with slight variations for the different services.  The steps are detailed below:

1. Site Locating

To locate available sites the first tab of the tool is utilized which is the USGS Sites service (Figure 2).  I input the type of sites I was looking for and the snapshot tool populated a NWIS Sites layer to display these sites.  Figure 2 shows one scenario I used for water quality analysis.
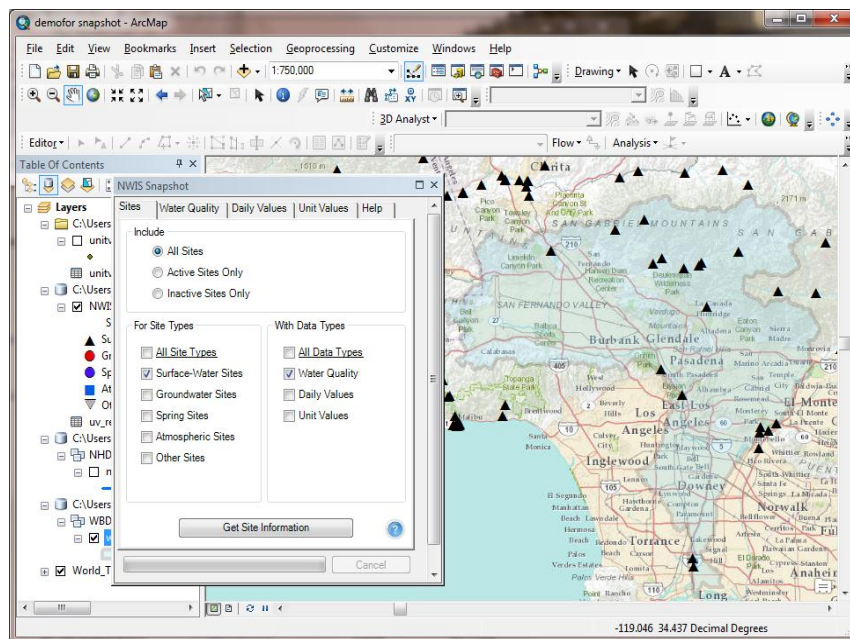


Figure 2: NWIS Snapshot Sites Service

2. Recognizing Available Data

Once the USGS sites can be viewed in the window, it is possible to find what data is available. To do this, the identify tool is selected and a USGS site is selected as well.  The identify window shows a link which opens a page which displays available data.  This step is done to know what data to expect from the Snapshot tool.
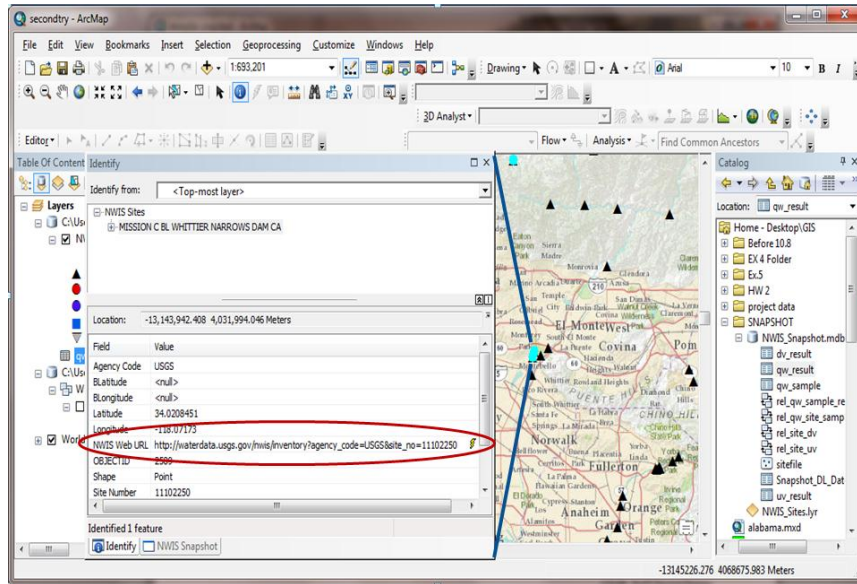
Figure 3: Information website access through the Identify tool



Figure 4: Website Displaying USGS site data Information CREDIT: USGS water Service

## 3. Site Selection

Once all available site locations are viewable, the desired sites may be selected using the select features tool in ArcGIS.  I placed a layer of the Huc8, Los Angeles River Watershed under the NWIS Sites layer to select all NWIS Sites within the outline of the watershed boundary.  Figure 4 shows what this process looked like in Arc GIS.
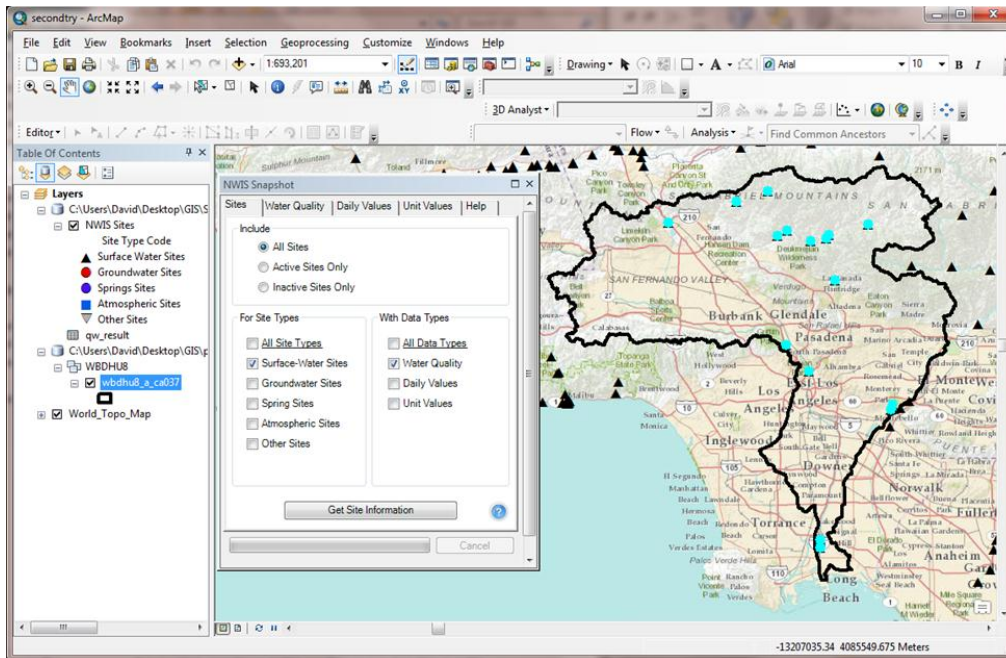


Figure 4: Selected sites within the Los Angeles River Watershed

## 4. Data Acquisition

Once I selected all the sites within the Los Angeles River Watershed, I chose the parameter.  For the water quality tab (Figure 6) I chose all available data for all available time.  I ended up with a huge amount of data points which I sorted through in Arc GIS using the Select by Attribute function.  I chose to use the parameter of total Dissolved Oxygen in mg/l which has the USGS parameter code 00300.

One alternative would have been to enter a parameter code.   I did not know the desired parameter number so I used an all-inclusive table to view which data points were available and select the attribute I wanted to project in GIS.  I chose a parameter this way because looking on the website for a desired parameter was difficult.    Alternatively, I could have also used the parameter menu expansion, Figure 5, which shows many types of available data parameters by name which would have been easier.
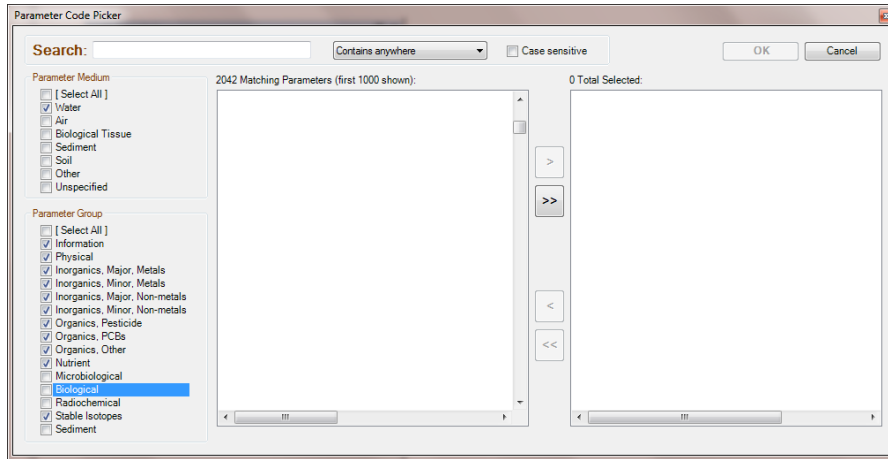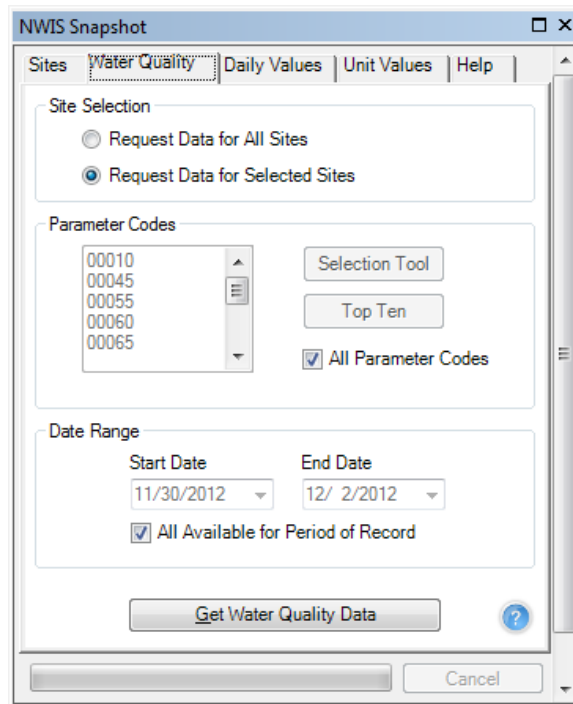
Figure 5: Parameter expansion menu



Figure 6: Water Quality Service Tab

Figure 7: Daily Values Service Tab



Figure 8: Unit (Instantaneous) Values Service Tab

For the daily values and unit values analyses I used the parameter of discharge in units of cubic feet per second (cfs) I then input the desired date range into the respective menus.  For the daily values service I used the date range of Jan 1, 1952 to June 1, 1952.  This data range was chosen because many USGS sites had data available for those dates.  For the unit value date range I indicated the 24 hour period prior to the date and time of December 1, 2012 at 12:00 AM.

5. Data Acquisition

Data output of the Snapshot tool is produced as a series of tables which appear in the ArcGIS Snapshot installation folder.  All possibly produced tables appear in the diagram of figure 9.  For every data transaction in the Snapshot tool, the Site Information table is populated as well as the Data Request Log.  For the water quality service, two tables are produced, namely qw_sample and qw_result.  For the daily values service, one data table is produced, this table is called dv_resut, and for the unit values service the uv_result table is produced. Dv_result and uv_result are designed to be used in conjunction with the Site Information table to comprise a complete dataset.
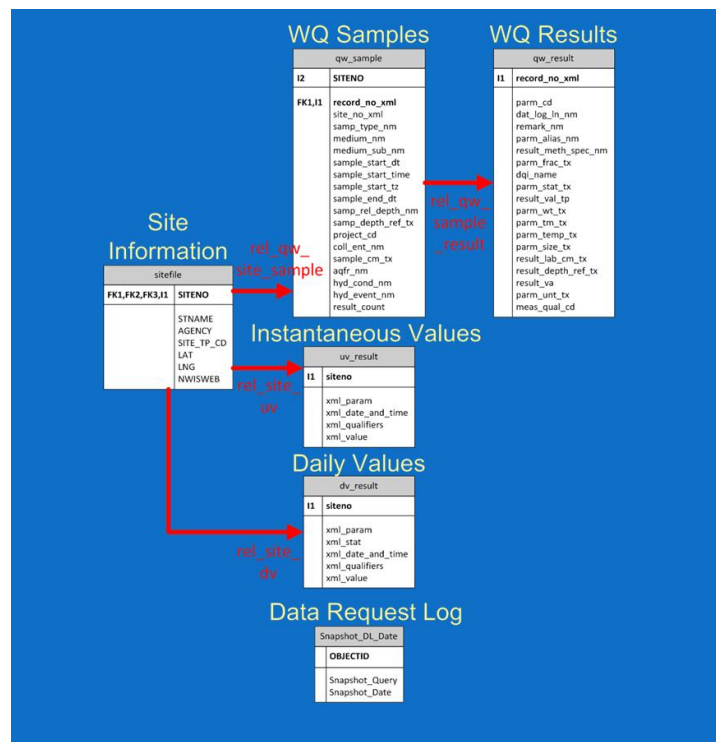


Figure 9: Schematic of data output tables from the Snapshot tool

6. Data Utilization

Utilizing the Data in ArcGIS was more of a challenge than expected. Two obstacles inhibited the direct utilization of the data in ArcGIS. These two obstacles were the separation of the data into separate tables and the format that the data came in. Overcoming these obstacles added a few extra steps to the process of using the data in ArcGIS.

All of the data that the Snapshot tool produces is in string format. Values in string format must be converted into the "double" format. Data must be converted in ArcGIS before a data table is joined or related. To do this, a new field was created in the original data table and the desired data field was copied and pasted in the new field. The new field data type was designated as "double". The fields I converted were that of longitude, latitude, and the value field. The other utilized fields, site number and long sample number, were acceptable as string data.

Correlating multiple data tables was annoying, but not difficult. Once the data was in the proper format, the table joining was simple. The water quality data format was unique in that it was necessary to correlate three tables. These tables are symbolized in figure 10. WQ Result was joined to the WQ sample table using the long sample ID number. The combined tables were then joined to Site information by the USGS site number.

The WQ Result table contained the actual values of interest and the values were only characterized by a long sample ID number. The WQ Sample table contained more information about the WQ Result table. Lastly, it was always necessary to unite the data result table with the Site Information table because this table contained the longitude and latitude information for the data points.
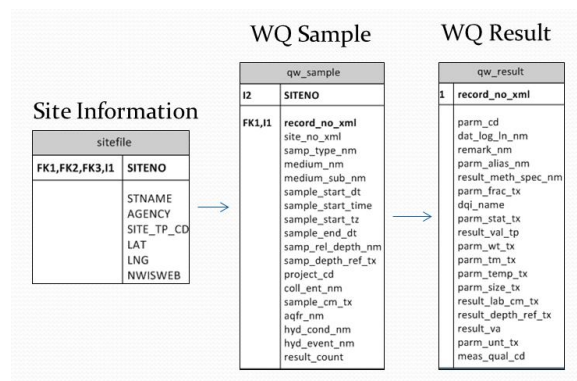


Figure 10: Output tables form the Water Quality Service

The daily values service and the unit values service produced data through two tables. The daily value service populated the Sites Information table as well as the dv result table. The unit value service populated the Sites Information table as well as the uv_result table. For each case, the necessary data in each table was converted to double data from string data, and the resulting tables were joined to the sites table through the site number. The resulting tables contained the data value, a USGS site number associated with the value, and longitude and latitude.
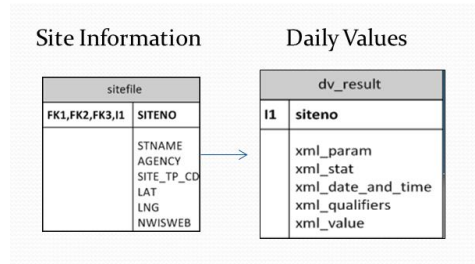


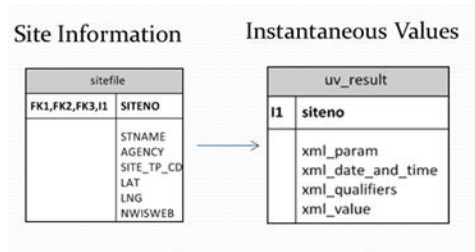Figure 11: Output tables from the Daily Values Service



Figure 12: Output Tables from the unit (Instantaneous) Values Service

Once the tables had been changed to accommodate for the correct data type and the necessary tables were merged, I exported the tables as a text file into Excel in order to convert the date and time data into a format that ArcGIS could read. The format came in the form of DD/MM/YYYY HH:MM, The necessary format for the date was the format of "YYYYMMDDhhmmss." "Y" is year "M" is month "D" is day, "h" is hour, "m" is minute, and "s" is second. The conversion was easy in excel as I only had to change the cell format for the column. After the date formatting is completed the table is imported to ArcGIS. The completed data table was uploaded to ArcGIS as a text file. Occasionally importing the file results in null columns in ArcGIS but after repeated attempts the table was displayed correctly. The exact reason for this problem is unknown, but the file was able to be loaded and used.

**Results**

The resulting map of the acquired water quality information is shown in figure 13. All of the data for dissolved oxygen concentration available through the Snapshot tool for the Los Angeles Watershed is geographically depicted in this map. Larger dots on the map represent

USGS sites which measured larger concentrations of Dissolved Oxygen, and small dots represent sites which measured lesser oxygen concentrations.  The mountainous region of the watershed produces lower dissolved oxygen levels in runoff water. The higher dissolved oxygen concentrations are found in the southern region of the watershed and by the coast.  I speculate that the urbanization of the southern region of the watershed leads to higher dissolved oxygen concentrations in the southern region.
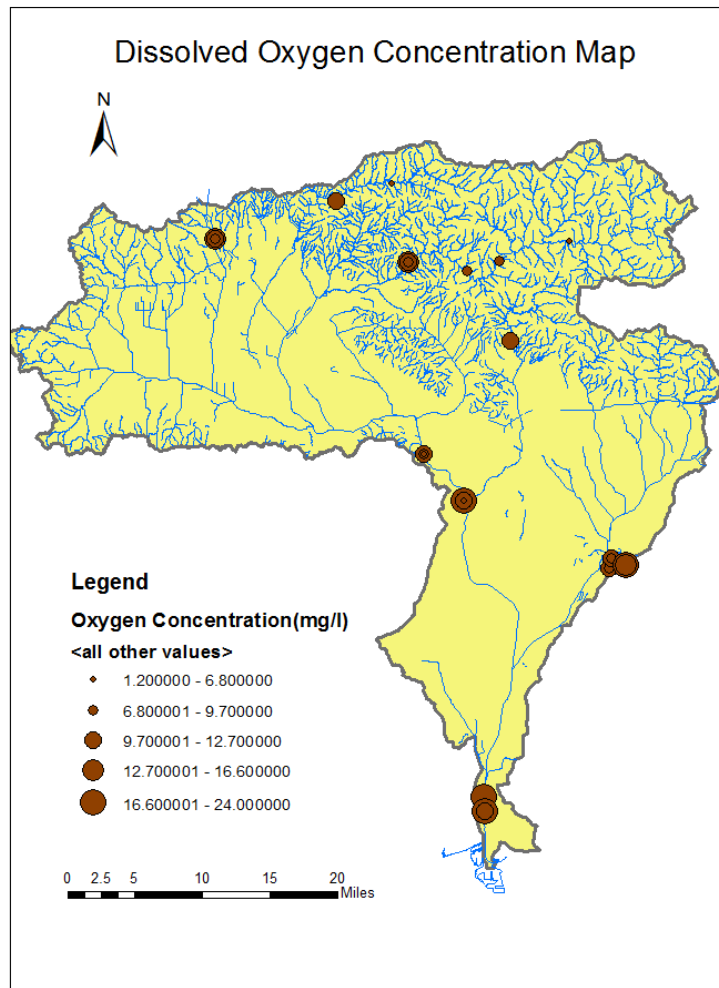


Figure 13: Dissolved Oxygen Levels in the Los Angeles River Watershed

The daily values and unit values data change in magnitude with time. Due to this fact, the data is easily viewed as a video.  Figure 14 shows clips from the daily values video. And figure 15 shows clips from the unit values video.  The videos show a larger dot for increased discharge and smaller dots for decreased discharge.  Patterns in the sizes of the dots are seen through the videos.  For example, in the daily values video, all the dots increase uniformly and decrease uniformly.  It is likely that precipitation cased this uniform increase. This uniform increase shows the validity of the data because precipitation affects the entire watershed and

not only a limited number of sites. The units values video shows a uniform increase and decrease in gage readings indicating precipitation as well.
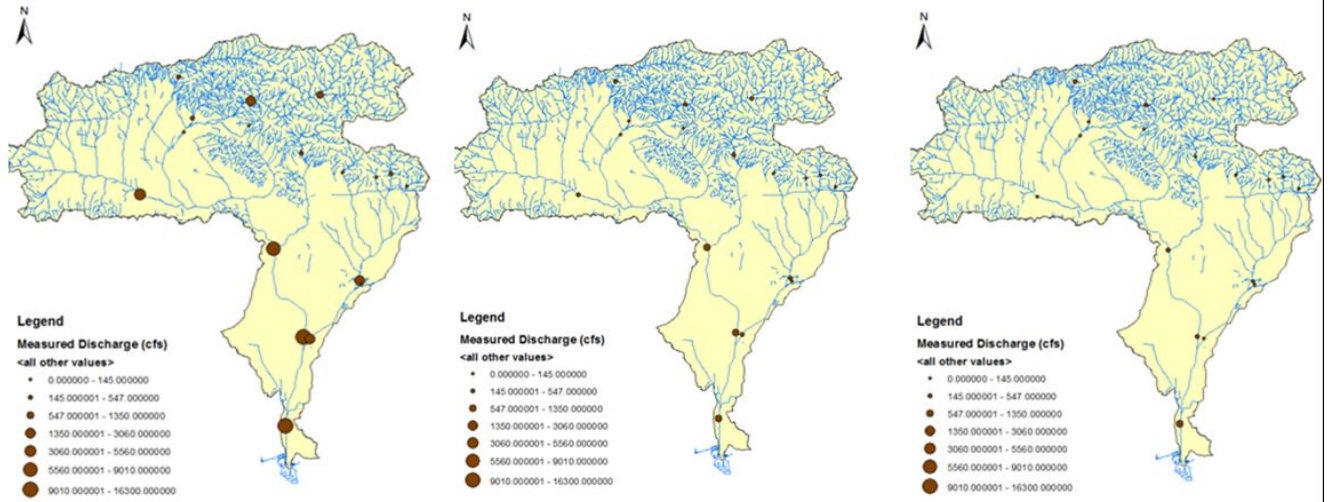


Figure 14: A 3 frame segment of the video for the daily value data.  Note the decreasing discharge from left to right.
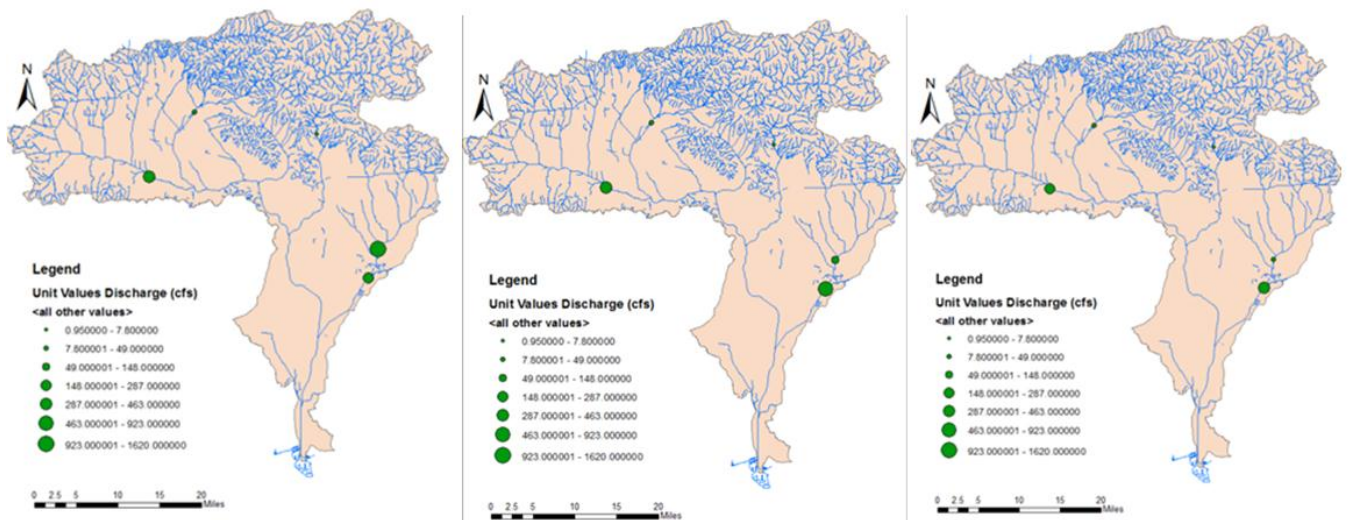


Figure 15: A 3 frame segment of the video for the unit value data.  Note the decreasing discharge from left to right.

**Conclusion**

The NWIS Snapshot tool is very adept for searching for and acquiring USGS Data.  The data acquired through the tool is not ArcGIS ready and requires a few manipulations to make it acceptable to ArcGIS.  The data format and the division of information amongst multiple table presented technical difficulties in utilizing the data. I mostly used Arc GIS to manipulate the data.  It would have been just as easy to export the data into Microsoft Excel or Microsoft Access, manipulate all of the data using that program, and re-upload it into ArcGIS.

Operation of the snapshot tool menus was rather pleasant.  It is a well-designed interface for acquiring data through GIS for those who understand how USGS data is labeled. Understanding the type and name of the data available through USGS allows the user to easily use the tool to find data.   If the user knows the parameter code for the data type in question the user can simply type it in the box.  If the parameter code is not known the user can utilize a larger menu for parameter selection or use the website to learn parameter codes. Familiarization with parameter codes only takes a short time if looking on the USGS website.

There are three recommendations I have for the data output format of the tool which would make the data very usable in ArcGIS.  One recommendation is to change the data output from string characters to double type data.  Double type data is more versatile and directly usable in ArcGIS.  Another suggestion is to produce data values and longitude and latitude in the same table.  This step would eliminate the need for correlation of multiple tables.  Lastly, the column of the date data could be more easily utilized if it were in a format usable to ArcGIS. The date format I converted the column to is rather untraditional outside of GIS, however the change would make the data seamlessly usable in ArcGIS.  The above suggestions eliminate the steps I undertook to use the data in GIS and would make data seamlessly applicable to ArcGIS applications.  Snapshot data most likely will have purposes outside of ArcGIS and the data format should account for that at well.

The NWIS Snapshot tool is a very powerful tool and radically simplifies user end complexity in data retrieval.  I think this tool is paving the way for future data acquisition tools. With some refining and development the snapshot tool has the capability to be an even more powerful tool for scientists as well as non-scientists.

**Data Sources**

United States Department of agriculture Natural Resource Conservation Service Geospatial
Data Gateway.12 Nov 2012. http://datagateway.nrcs.usda.gov/GDGOrder.aspx

USGS Water Services.25 Nov 2012. http://datagateway.nrcs.usda.gov/GDGOrder.aspx


**Literature Cited**

Holl, Sally. 2011. U.S. Geological Survey Community for Data Integration NWIS Web Services S
snapshot Tool for ArcGIS. 30 December 2011.

USGS Water Service, USGS Department of the Interior 25 June 2012, 1 Dec 2012. H
http://waterservices.usgs.gov/