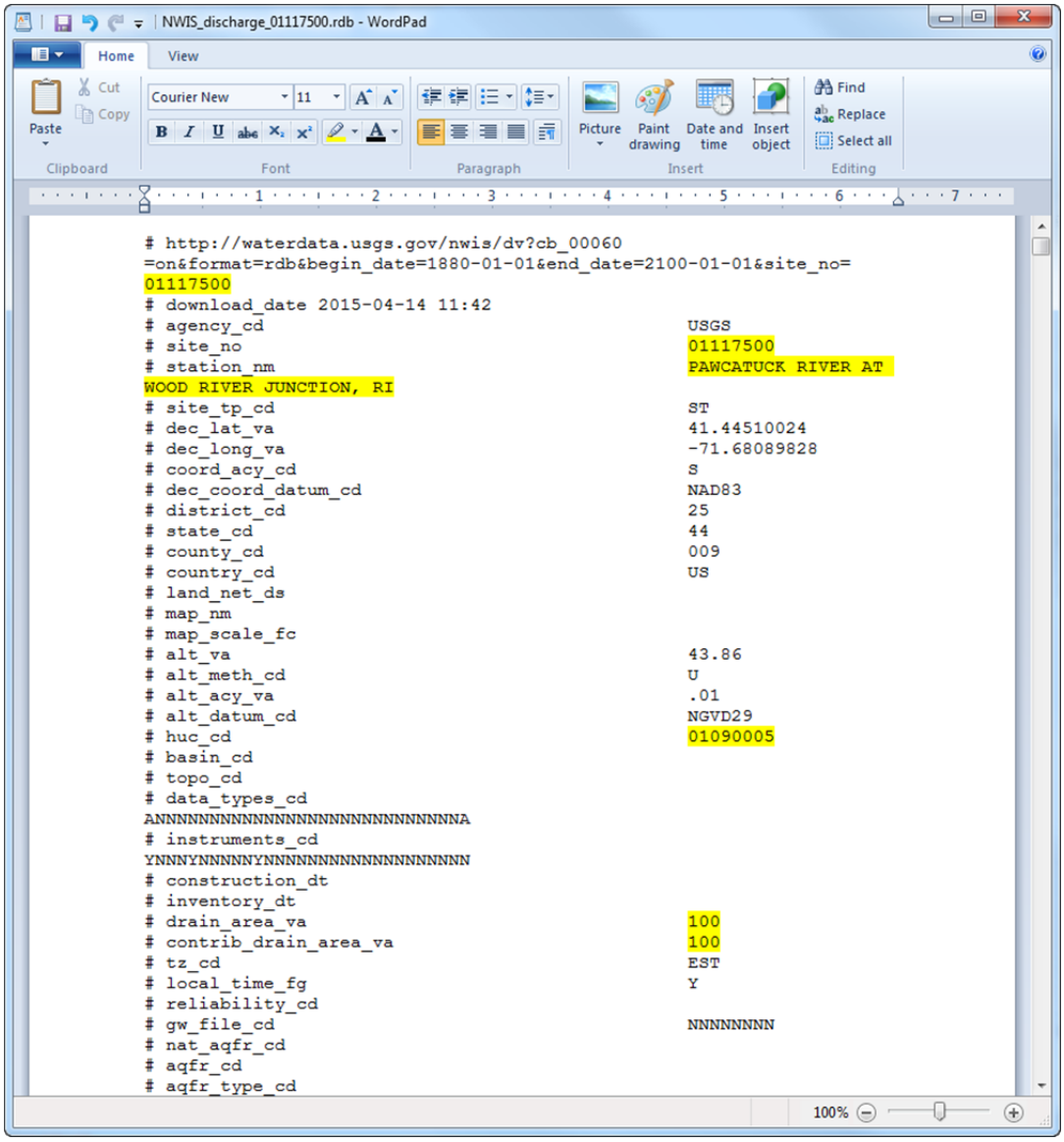
Groundwater Toolbox Tutorial

Reformatting streamflow data to the USGS RDB (relational database) format for importing into the Groundwater Toolbox

April 30, 2015

This document shows steps that can be taken to construct a USGS RDB streamflow data file from any streamflow data source, such as data collected at a streamgage that is not part of the USGS National Water Information System (NWIS). The RDB data file can then be imported into the Groundwater (GW) Toolbox for data analysis. The advantage of this approach lies in the completeness of the parameter values that are provided by the RDB file structure for most types of streamflow analyses that are provided in the GW Toolbox.

This tutorial starts with an existing USGS RDB streamflow data file template. The template can be from any existing USGS RDB streamflow data file that has already been saved to the user's computer, typically having the filename "NWIS_discharge_stationnumber.rdb." For example, the following screen capture shows an RDB file for the Pawcatuck River at Wood River Junction, RI, streamgage (USGS station number 01117500); the file name is "NWIS_discharge_01117500.rdb" (text that is highlighted in yellow in the screen capture is described below):



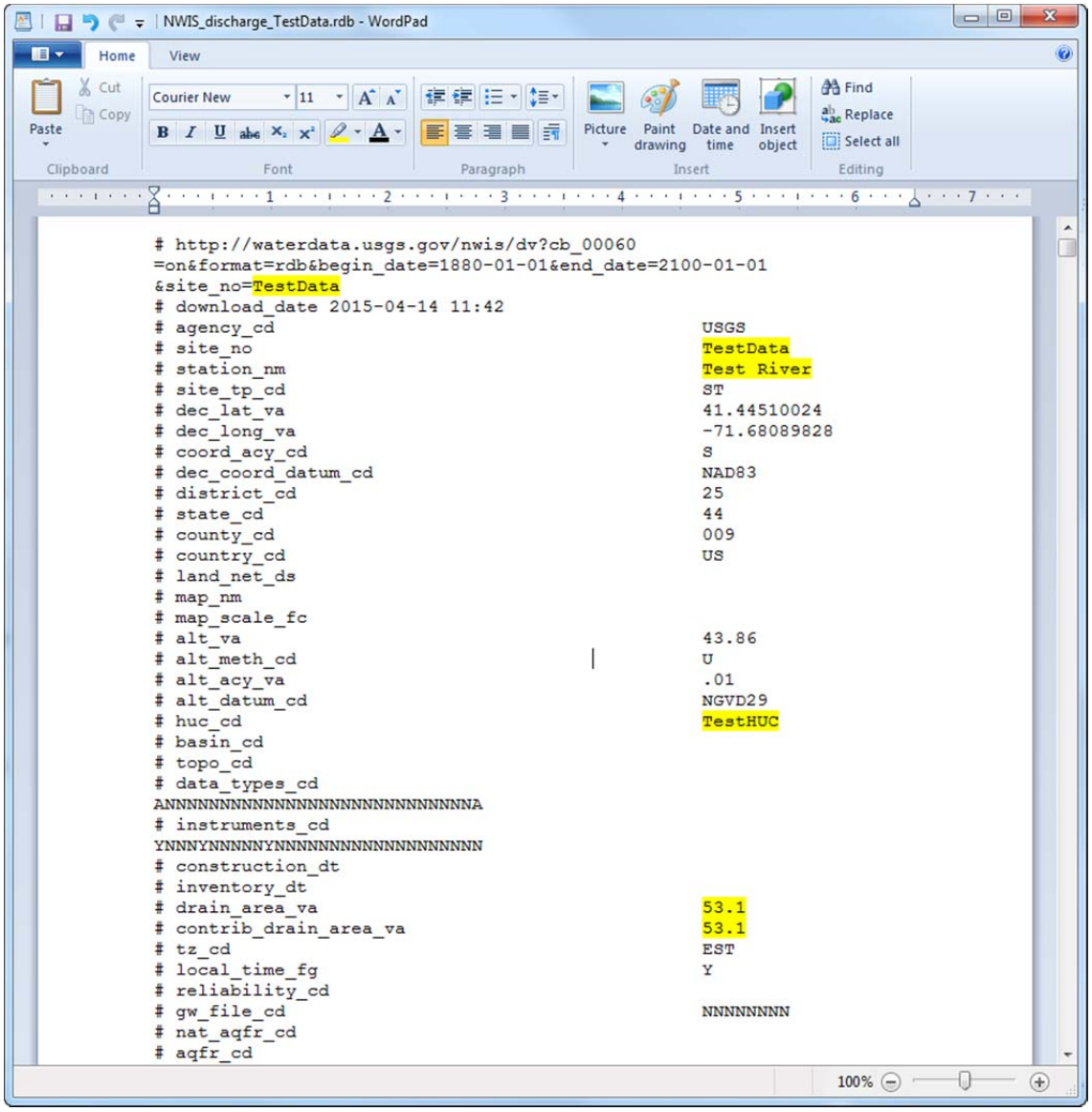
This RDB template file, which is included with the GW Toolbox distribution, contains a full set of parameters that are typically populated as part of the data-retrieval process from the NWIS Website. The template can be used to accommodate data provided by the user by modifying some of the parameters and then replacing the existing streamflow data for the Pawcatuck River with the user's data. As a start, the user might want to copy the template .rdb file to the new .rdb file that will be created (this will save the original template file for future use). In this example, the new file will be called "NWIS_discharge_TestData.rdb."

The key concept for creating the new RDB file is to maintain the header-line structure as it is provided in the template file so that the GW Toolbox can correctly read parameter and data values. Therefore, do not remove any header lines from the template or change the structure of the template; instead, simply replace certain key parameter values:

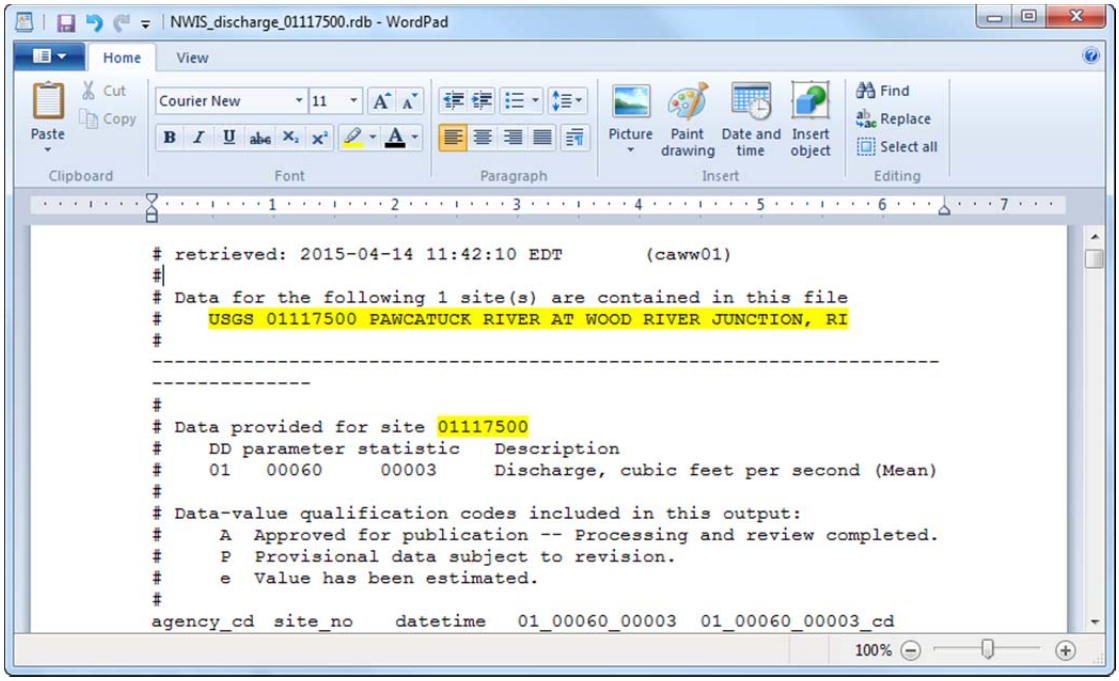
1. The first line of the header, showing the Web URL, should be kept intact, with the exception that the station number (in this case, “01117500”) can be replaced with the name of the user’s station (such as “TestData” as shown in the screen capture below).

2. The user may choose to change any parameter values in the top section of the header (above the --- WARNING --- line). These parameters might include the latitude, longitude, elevation, or drainage area of the station. Important note: Some of the functionality of the GW Toolbox, such as the hydrograph-separation methods, was designed under the assumption that streamflow has units of cubic feet per second and drainage area of the basin is reported in square miles; therefore, if the user’s data are in other units, such as metric, it is recommended that the data be converted to units that are used by the GW Toolbox prior to importing the data.

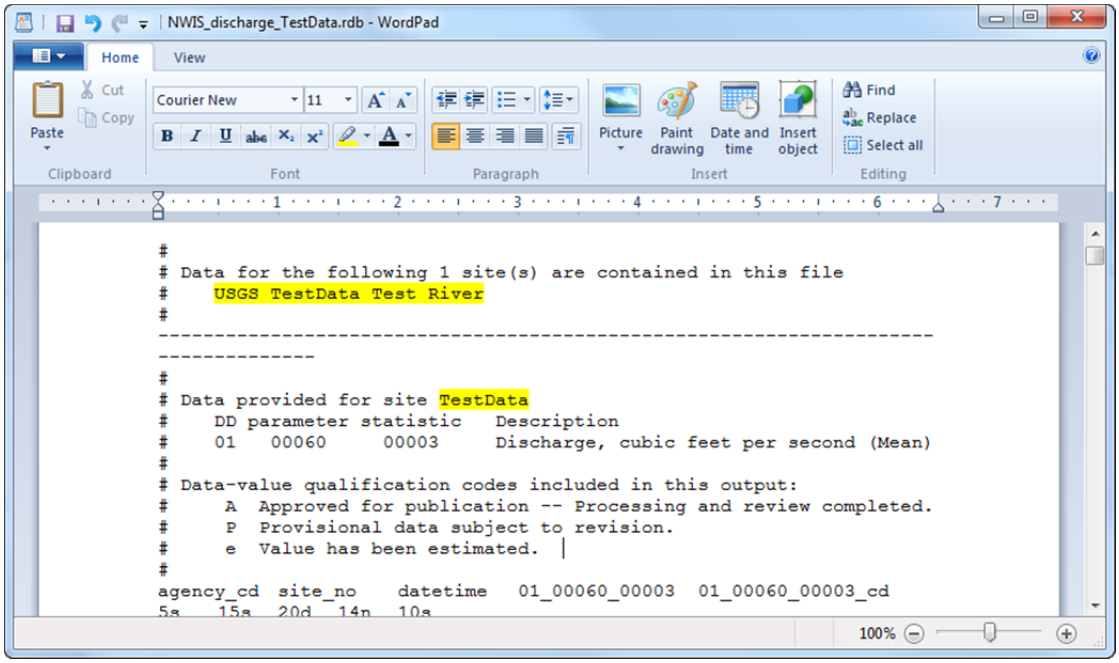
In the example below (“NWIS_discharge_TestData.rdb”) the parameters highlighted in yellow have been modified:



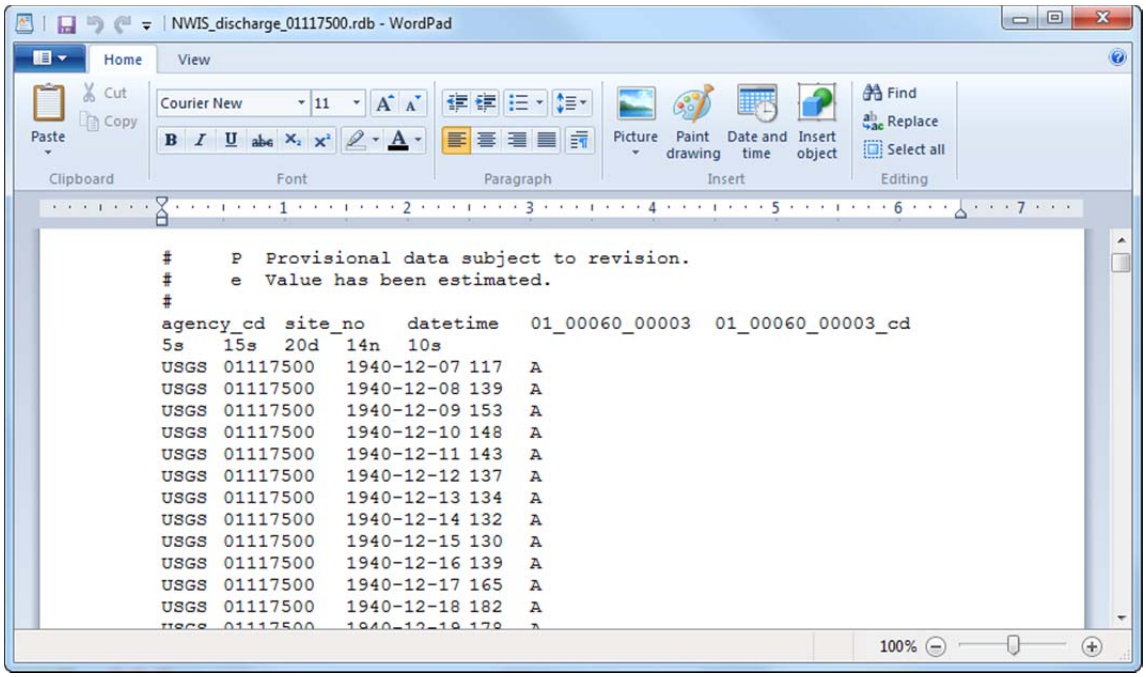
3. Only two changes are needed to the header section below the --- WARNING --- line, both of which relate to the site information. First, to the line that begins “Data for the following…” and second to the line that begins “Data provided for site…” The original data file appears as follows:



And the modified file as follows:

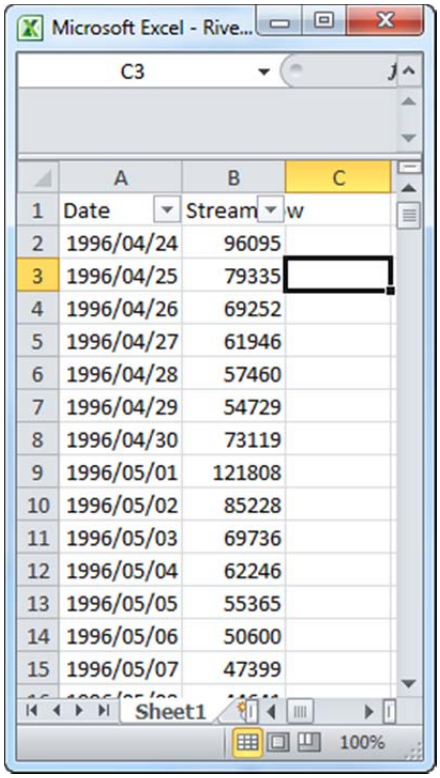


4. The next step is to delete the original data and replace them with the user's data. The original RDB file has the following format for the streamflow data:



For this tutorial, we will replace these data with data from an Excel file called "Riverflows2.xls."

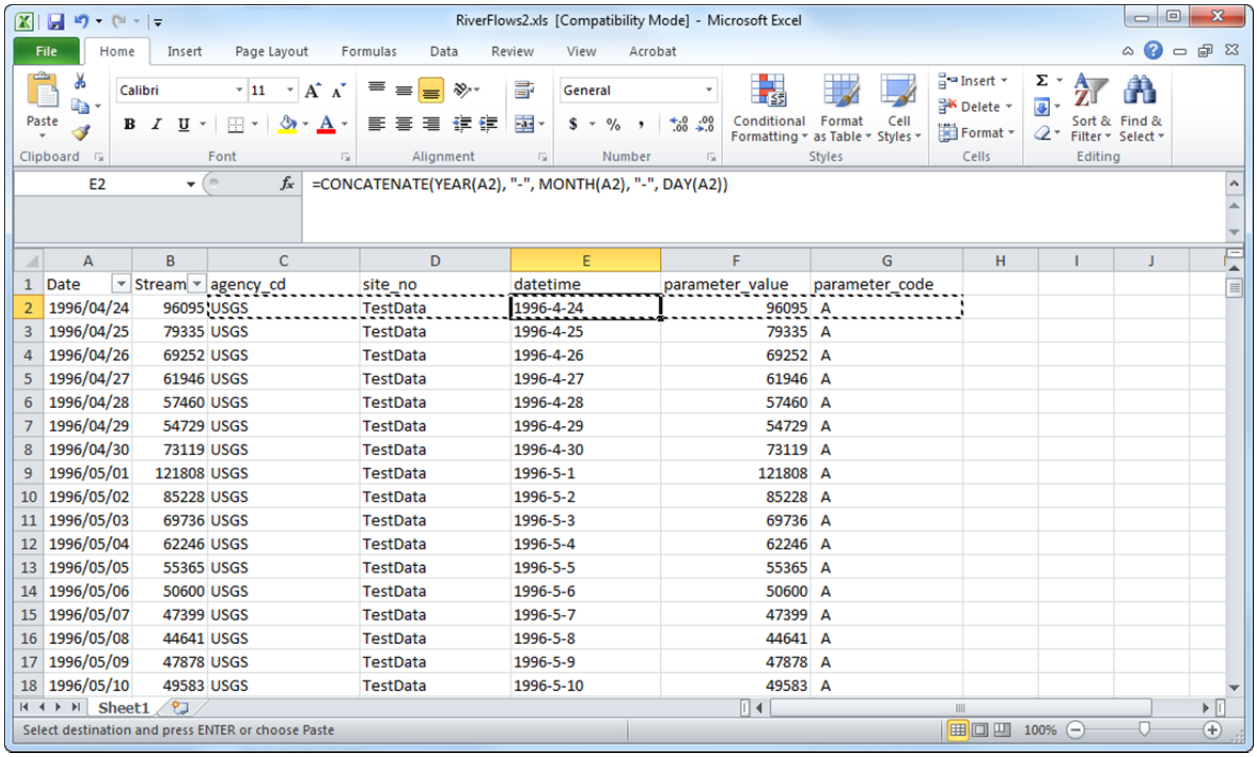
This file contains data in two columns, a date column and a streamflow-values column:



We want to modify the columns of the Excel file so that they have the same structure as the RDB file: agency_cd, site_no, datetime, parameter_value, and parameter_code. This is done by creating five new columns in the Excel file (see screen capture on next page), with 'USGS' in the first column, 'TestData' in the second, revised dates in the third, streamflow values in the fourth, and parameter_code ' A' in the fifth. Values in the new datetime column (column E) are created using the Excel Concatenate command, which is located in the '**Text**' options under the '**Formulas**' tab:

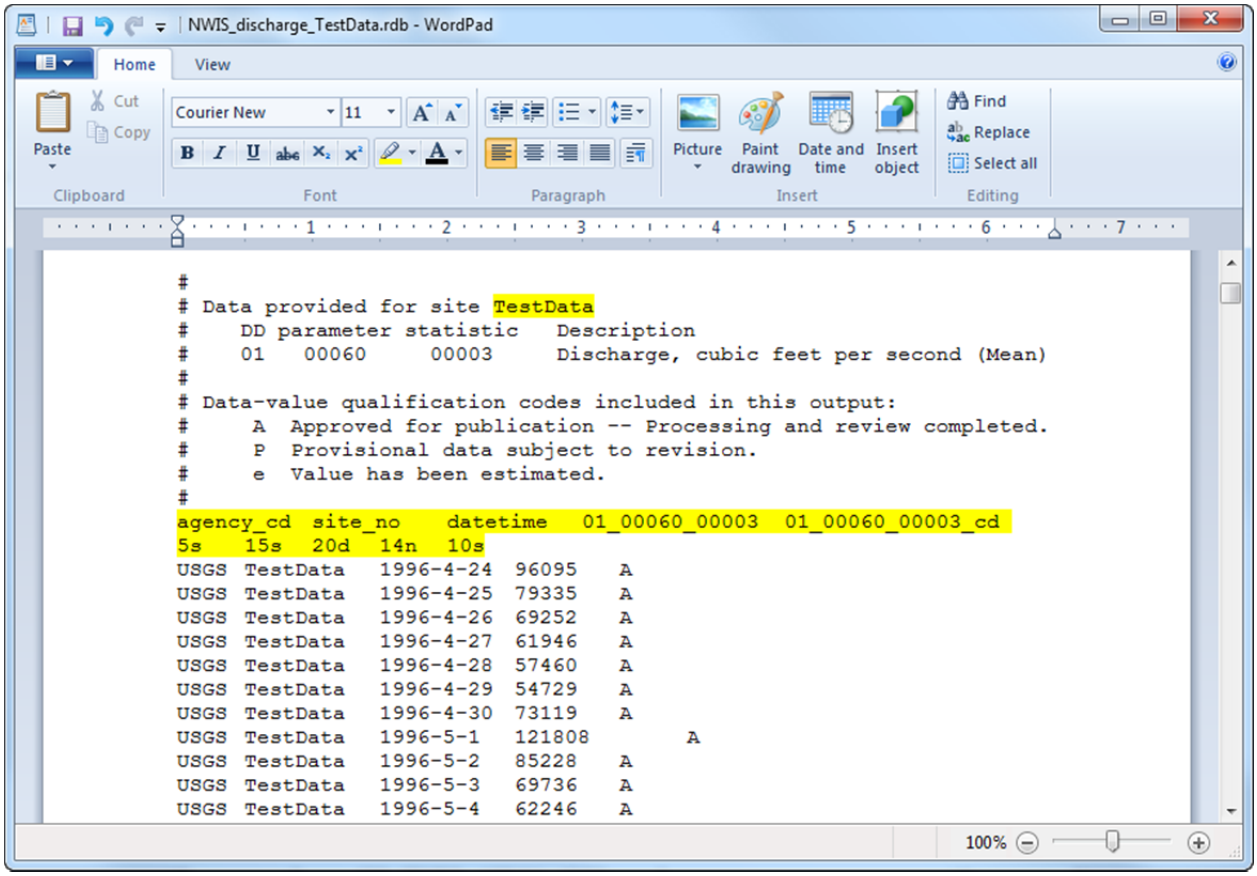
=CONCATENATE(YEAR(A2), "-", MONTH(A2), "-", DAY(A2))

The revised Excel file is then:



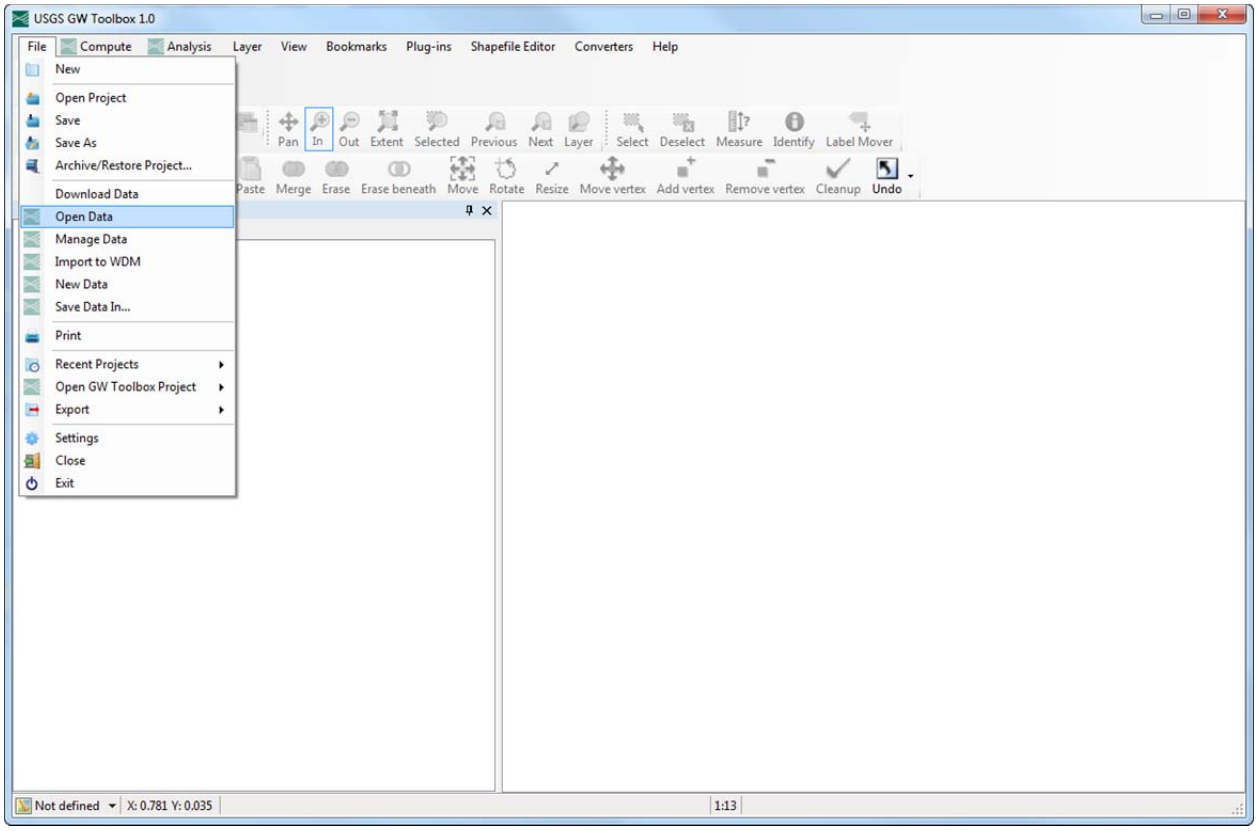
	A	B	C	D	E	F	G
1	Date	Stream	agency_cd	site_no	datetime	parameter_value	parameter_code
2	1996/04/24	96095	USGS	TestData	1996-4-24	96095	A
3	1996/04/25	79335	USGS	TestData	1996-4-25	79335	A
4	1996/04/26	69252	USGS	TestData	1996-4-26	69252	A
5	1996/04/27	61946	USGS	TestData	1996-4-27	61946	A
6	1996/04/28	57460	USGS	TestData	1996-4-28	57460	A
7	1996/04/29	54729	USGS	TestData	1996-4-29	54729	A
8	1996/04/30	73119	USGS	TestData	1996-4-30	73119	A
9	1996/05/01	121808	USGS	TestData	1996-5-1	121808	A
10	1996/05/02	85228	USGS	TestData	1996-5-2	85228	A
11	1996/05/03	69736	USGS	TestData	1996-5-3	69736	A
12	1996/05/04	62246	USGS	TestData	1996-5-4	62246	A
13	1996/05/05	55365	USGS	TestData	1996-5-5	55365	A
14	1996/05/06	50600	USGS	TestData	1996-5-6	50600	A
15	1996/05/07	47399	USGS	TestData	1996-5-7	47399	A
16	1996/05/08	44641	USGS	TestData	1996-5-8	44641	A
17	1996/05/09	47878	USGS	TestData	1996-5-9	47878	A
18	1996/05/10	49583	USGS	TestData	1996-5-10	49583	A

The five new columns of data (without the heading row) can now be copied and pasted into the new RDB file. Note that the data section of the RDB data file is tab-delimited, and users will want to ensure that the data that are pasted into the file also are tab delimited. Also, be sure to retain the two header lines in the data section, which are highlighted in yellow in the revised RDB file below:

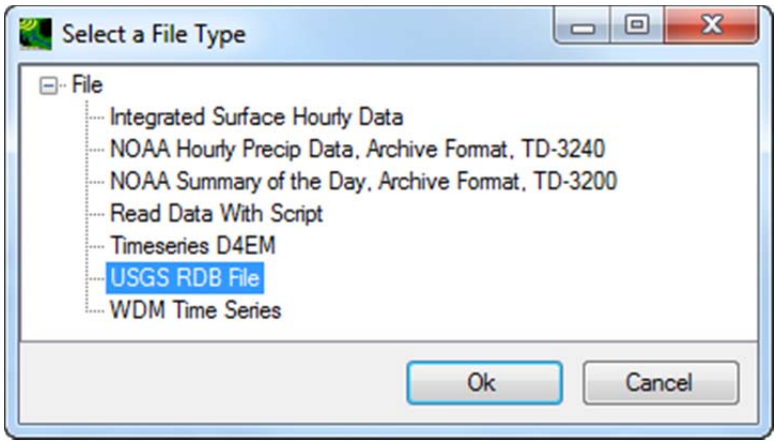


(Text continues on next page)

The revised RDB file is now ready to be imported into the GW Toolbox. Note that the data can be imported into the GW Toolbox without having a project area defined for the data. This is done by closing the "Welcome to the USGS GW Toolbox" dialog box that is shown after launching the GW Toolbox and then going directly to the "**File>Open Data**" menu option:

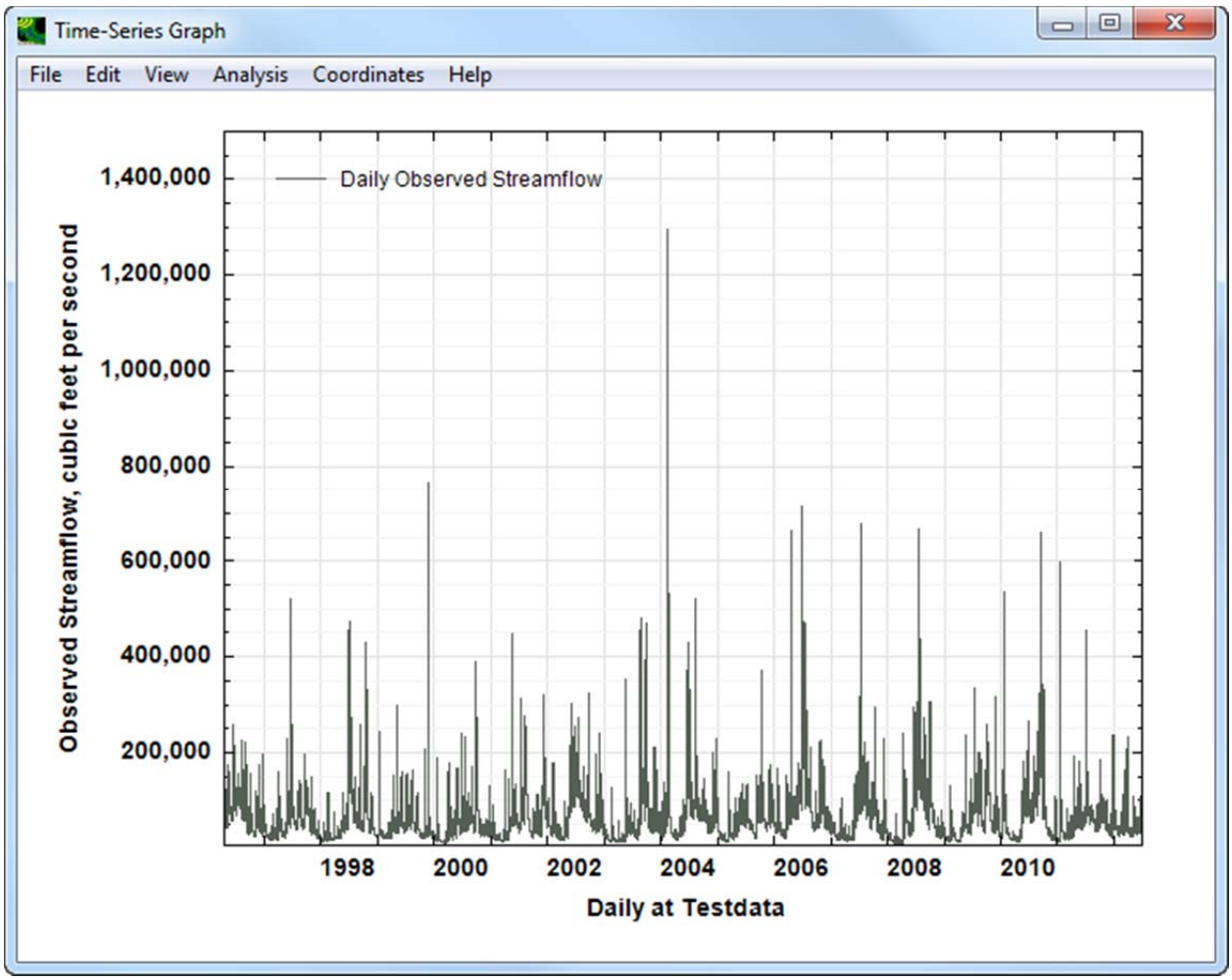


Then select "USGS RDB File" in the resulting dialog box:

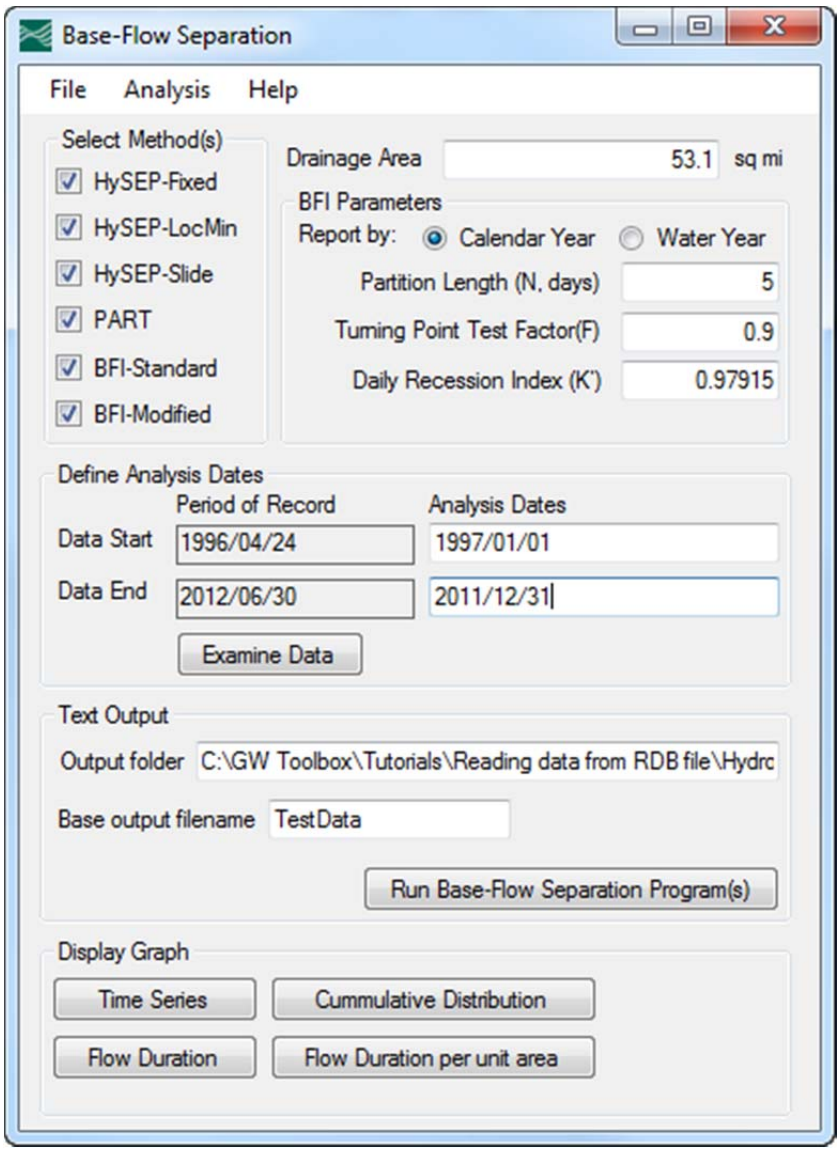


This option will allow the user to navigate to the new ".rdb" file.

Now that the data have been imported, the data can be used with GW Toolbox functionality, such as graphing:



or a Data Tree or List analysis. Notice that when the Base-Flow Separation analysis is selected, the Drainage Area for the basin (53.1 square miles), has been correctly populated in the dialog box:


Base-Flow Separation
File
Analysis
Help
Select Method(s)
HySEP-Fixed
HySEP-LocMin
HySEP-Slide
PART
BFI-Standard
BFI-Modified
Drainage Area
53.1
sq mi
BFI Parameters
Report by:
Calendar Year
Water Year
Partition Length (N, days)
5
Turning Point Test Factor(F)
0.9
Daily Recession Index (K')
0.97915
Define Analysis Dates
Period of Record
Analysis Dates
Data Start
1996/04/24
1997/01/01
Data End
2012/06/30
2011/12/31
Examine Data
Text Output
Output folder
C:\GW Toolbox\Tutorials\Reading data from RDB file\Hydrc
Base output filename
TestData
Run Base-Flow Separation Program(s)
Display Graph
Time Series
Cummulative Distribution
Flow Duration
Flow Duration per unit area