Data Issues

Hydrologists who use the PULSE program should be aware of data issues described below. The auxiliary program PREP will read streamflow data obtained from USGS web sites. The version of the program that reads such data have been available on the web site since 2002. Over the last few years there have been changes in the procedures for obtaining web data, and in July 2006 there was a change in data-file format which might affect program execution. Details are explained below.

OBTAINING STREAMFLOW DATA FILES

Here are instructions for obtaining streamflow data from USGS web sites. These procedures are different from the instructions in the User Guide for the PULSE Program (page 4).

A streamflow data file can be obtained from the following USGS web site: http://waterdata.usgs.gov/nwis/dv/?referred_module=sw. From this site, the user is first asked to "Choose Site Selection Criteria". For example, if the site of interest is Holiday Creek near Andersonville, Virginia, yet the site identification number is not known, the user might check the boxes for *State* and *Site Name* and then hit the "Submit" button. At the subsequent "Select sites that meet all of the following criteria" page, the user would then enter *Holiday* under site name, select the state of Virginia from the state list, check the "Streamflow, ft^3/s" box in the "Water Level/Flow Parameters" group of the Available parameters list, and then issue the request using the "Submit" button near the very bottom of the page. The subsequent page will be a list of all sites found that meet the specified criteria. Clicking on the site number will display a page including graph(s) of the previous 365 days of data. The final step to obtain a file of streamflow data is to check the "00060 Discharge (Mean)" item under the Available parameters heading, choose "Tab-separated" from the Output format list, enter the Begin and End dates of interest, and then hit GO. Once the data are displayed in your browser select "Save as" from the file pull-down menu of your browser and designate the directory where the data are to be written on the user's computer, and specify a file name. On some systems this might require copying and pasting into Notepad. File names should be in the format *xxxxx.txt*, where *xxxxx* is an abbreviated station name. In this case the file name might be designated as *Holiday.txt*. (Note: the length of the file name, including *.txt,* should be 12 characters or less.) Designate file type should be "Text File (.txt)".

Data-file format

In July 2006 there was a change in the format of streamflow data files that come from USGS web sites. Prior to this date, if there was a gap in record for a given streamflow-gaging station, and if the user retrieved data from a time period that includes the gap, the data file would not include lines for days of missing data. New data files will include lines for every day in that gap. In January 2007 a new version of the PREP program was placed on the PULSE web site. This new version of PREP can read streamflow data in the newer format or in the older format. (Note – The PULSE program does not read a streamflow datafile.)

A few more comments about data files obtained from a USGS web site. Beginning in 2006, some files may include text qualifiers about streamflow data. These qualifiers, which may be evident in recent data (2006+), can appear immediately after a streamflow quantity and in some cases the qualifier may be located in the place of the streamflow quantity. The program may not execute properly due to these text qualifiers. For this reason the program user may need to edit streamflow data files, scanning for these qualifiers, and removing them if they exist.

In some applications, the program user may choose to generate a data file from a source other than a USGS web site. In this case, a few items about file format may be helpful: (1) Most of the header information is not read by the program, and may not be necessary in a user-generated file. (2) The program starts reading data, two lines after it detects the string "agency" in the first five spaces of a line. (3) The columns of information in each data line are tab-separated in a data file from a USGS web site (these could be substituted with single blank spaces in a user-generated data file).

Al Rutledge, January 2007