

# FLOOD REGIME CHARACTERIZATION

Version 1.0 - August 2006

User's Manual



Written by Eric Booth

University of California  
Davis

Developed in Matlab 7



## **Introduction**

### **Companion Paper**

The research that led to the development of this computer program is discussed in the following paper:

*Booth, Eric, Jeffrey Mount, and Joshua Viers. Hydrologic Variability of the Cosumnes River Floodplain. San Francisco Estuary and Watershed Science. Vol. 4, Issue 2 [September 2006]. Article 3. <http://repositories.cdlib.org/jmie/sfews/>*

### **Background**

This manual describes how to use a program, developed in Matlab 7, called Flood Regime Characterization (FRC). This program is an attempt to transform the methods discussed in the companion paper into a portable computer program so that they can be used by researchers and managers on other watersheds in order to characterize a flood regime. The main inputs to the program are a daily streamflow record and a set of flood duration and magnitude thresholds that have some meaningful connection to the hydrologic, geomorphic, and/or ecological processes in the specific watershed. The main outputs are a list of floods, flood types, and water year types along with useful statistics for each list and estimates of the frequency of each flood type and water year type.

The entire foundation of the program is based on the concept that a flood regime is made up of four components: magnitude, duration, timing, and frequency. The flood types are created by separating out similar flood events based on their magnitude (i.e., peak flow) and duration and also, somewhat indirectly, their timing. Once the flood types are created, the frequency of each type can be found throughout the entire record. The water year types are then developed by clustering similar water years together based on the number of days each flood type occurs during each water year. The frequency of each water year type throughout the record is also calculated and presented.

### **System Requirements**

Program users must have Matlab 7 installed in order to use the FRC program. Previous versions of Matlab will NOT be able to run this program. Also, the Statistics Toolbox, which is not included with the standard Matlab 7 package, must also be installed because the hierarchical cluster analysis used to create the water year types calls functions that are only available in this specific toolbox. All of these products can be found on-line at <http://www.mathworks.com>

All files necessary to run the program add up to less than 200 KB. After the program has run, the addition of the output files should still bring the total amount of disk space needed to less than 2 MB. The calculations used in the program require very little computing power. This

program is command-line driven and does not use extensive graphics. Keep in mind that system requirements for the installation of Matlab 7 are far more extensive than for the FRC program.

## **Installation**

All Matlab M-files necessary for the program can be found at the following location: <http://watershed.ucdavis.edu/pages/programs.html>. The zip file 'FRC\_v1\_0.zip' contains 26 M-files, which compose the entire program. The file 'FRC.m' is the script file that needs to be called from the Matlab Command Window in order to run the program.

Once Matlab 7 is opened, change the current directory to the folder that contains the program files and call the 'FRC.m' script file by either typing 'FRC' in the Command Window or right-clicking on that file in the Current Directory window and selecting 'Run'.

## **Inputting Streamflow Data into FRC**

### **Downloading from the web**

Daily streamflow data can be obtained from various federal, state, and local natural resources agencies. However, the most extensive collection of daily streamflow data for the United States is available from the U.S. Geological Survey's National Water Information System. This database can be accessed on-line at the following location: [http://waterdata.usgs.gov/nwis/dv/?referred\\_module=sw](http://waterdata.usgs.gov/nwis/dv/?referred_module=sw). Daily streamflow data for Canada can also be obtained on-line from the Water Survey of Canada at the following location: <http://www.wsc.ec.gc.ca/hydat/H2O/>.

### **Format**

Once you have acquired a daily streamflow time-series, it must be formatted correctly in order for it to be read by the FRC program. First, the data should be arranged into four columns with a space or a tab in between each column.

- Column 1: month
- Column 2: day
- Column 3: year
- Column 4: streamflow

If your initial data set uses a one-value date format (e.g., 12/31/1999), a simple and straightforward way to transform the data into the correct format is to use the MONTH, DAY, and YEAR functions in Microsoft Excel.

Second, since the water year (October 1 – September 30) will be used exclusively in this analysis, the first data point must be on October 1 and the last data point must be on September 30 in order to correctly perform the water year type classification.

Third, there should be a value associated with every day in between your first and last data points. No day can be missing from the input file otherwise an error will occur.

Finally, the data set in this format must be saved as the text file ‘flow.txt’ in the folder that contains the program files. For an example of a properly-formatted ‘flow.txt’ file, see the Cosumnes example folder, which can be downloaded at <http://watershed.ucdavis.edu/pages/programs.html>.

### **Dealing with incomplete streamflow records**

Finding a sufficiently long streamflow record with no data gaps can be very challenging. If you do have gaps in your record, it is recommended that you use a correlation technique to extend your record by using precipitation data or other streamflow data from a nearby gage. A good discussion of these different techniques can be found in the following book:

Gupta, R.S. 2001. Hydrology and Hydraulic Systems.  
2<sup>nd</sup> Edition, Waveland Press, Inc. pp. 364-375

If you are unable to completely extend your streamflow record so that no gaps are present, the FRC program can still be used but several actions must be performed in order to properly input your data. As mentioned in the previous section, there can be no missing days from your input data file. Therefore, ‘filler’ values must be entered in place of the gaps in your streamflow record. It is recommended that you use some consistent value (e.g., -99). Once all of the gaps have been filled, you can then format and input the data into the FRC program.

The next action that must be performed is inputting the water years that you want to exclude from the FRC analysis (i.e., the water years with missing data). There are three instances in the program (calculating mean monthly streamflow, calculating empirical frequencies of each flood type, and creating water year types) where you will be asked if there are any water years that you would like to exclude from the following analysis. If you answer ‘yes’, you will have three options for inputting these excluded water years: 1) manually enter a list of the water years, 2) manually enter ranges of water years, or 3) input a file that stores the water years. This allows you to exclude different sets of water years for different parts of the FRC analysis. The following explains each of the three options in more detail.

#### *Manually enter a list of the water years*

If you select this option, all you have to do is enter in the water years that you would like to exclude. After you have entered the final water year, hit ‘Enter’ again without providing a water year.

### *Manually enter ranges of water years*

After you select this option, you will be asked to provide the number of ranges that will be entered. Next, provide the first and last water year of each range. This option was created to speed up the inputting process of long lists of excluded water years.

### *Input a file that stores the water years*

This is the quickest way to input a long list of excluded water years. After you select this option, make sure that you have created a text file called 'excl\_wys.txt', which stores all of the excluded water years in a single vertical column with one water year per row.

## **Running the FRC program**

### **Overview**

The main menu of the FRC program gives the following seven choices:

1. Plot the entire streamflow record
2. Calculate mean monthly streamflow
3. Separate flood events
4. Create flood types
5. Create water year types
6. Generate average hydrographs for each water type
7. Exit program

It is highly recommended that the user go through each of the choices sequentially. Certain variables created in steps 3, 4, and 5 will be used with other later options. Therefore, the third choice is the latest choice in the sequence that the user can start with if they are not interested in plotting out the streamflow record or finding the mean monthly streamflow values. An error will occur if any of the following occur:

- 4 is selected before 3
- 5 is selected before 3 and 4
- 6 is selected before 3, 4, and 5

After an initial choice is selected, the program will ask what type of units are being used (English or Metric) and remind the user to format the input file correctly. Once the user has gone through both of these displays, they will not have to repeat them again after selecting other options.

## **Plot the entire streamflow record**

This is a straightforward step that plots the streamflow record versus the date in a separate Matlab figure. The x-axis is the date shown in years and y-axis is the streamflow in the unit type specified by the user.

## **Calculate the mean monthly streamflow**

This step calculates the mean streamflow for each month of the year for the entire record and plots it in a separate Matlab figure. If you decide to exclude water years from this analysis (e.g., due to missing data), the program will not use data from the excluded water years to calculate the mean monthly streamflow values. For the Matlab figure, the x-axis corresponds with the month of the year (e.g., 1 = January) and the y-axis is the mean monthly streamflow in the unit type specified by the user. After the plot is shown, the user has the option of exporting this data. If the user selects 'yes', the file 'MMF.txt' is created and stores the mean monthly streamflow value for each month in the following format:

Column 1: month  
Column 2: mean monthly streamflow

A file header is present to denote the contents of each column.

## **Separate flood events**

This step separates all of the flood events from the streamflow record based on a user-defined flood magnitude threshold. A flood event is defined as any period when the streamflow equals or exceeds the flood magnitude threshold. This action of separating the flood events is performed after the user is prompted to enter the flood magnitude threshold. Once the action has been completed, a plot showing all separated flood events is shown in a separate Matlab figure. The x-axis of the plot is the duration of the flood event in days and the y-axis is the peak streamflow value (in the user-specified unit type) of the flood event. The color of each data point corresponds with the starting month of the flood event. Earlier in the water year corresponds with cooler colors while later in the water year corresponds with warmer colors as seen in the legend.

The user is then prompted as to whether or not they would like to export this data. If the user selects 'yes', the file 'floods.txt' is created and stores statistics associated with each flood event in the following format:

Column 1: water year  
Column 2: water year flood (the flood event number for that water year)  
Column 3: start date (in the format: mm/dd/yyyy)  
Column 4: start water year day (the day of the water year in that the flood started)

- Column 5: start water year fraction (the fraction of the water year in that the flood started)
- Column 6: start month (the month of the year in that the flood started)
- Column 7: start water year month (the month of the water year in that the flood started)
- Column 8: end date (in the format: mm/dd/yyyy)
- Column 9: duration (in days)
- Column 10: peak flow (in units of ft<sup>3</sup>/s (cfs) or m<sup>3</sup>/s (cms))
- Column 11: mean flow of the flood event (in units of ft<sup>3</sup>/s (cfs) or m<sup>3</sup>/s (cms))

A file header is present to denote the contents of each column.

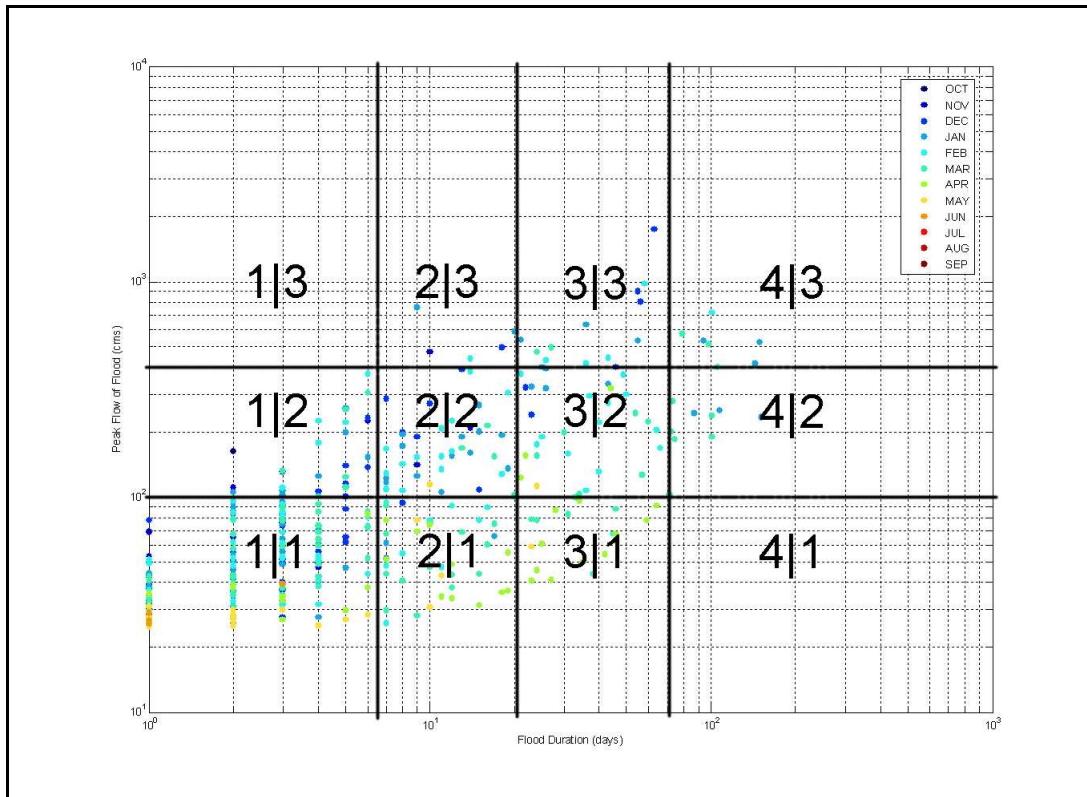
### **Create flood types**

Once the flood events have been separated, similar events can now be grouped into flood types. The grouping is performed by inputting several flood duration and magnitude (i.e., peak flow) thresholds. It is recommended that each threshold should reflect some physical or ecological threshold that is understood to be present in the study area. For example, on the Cosumnes River floodplain, it was found that floods that peak at or above 100 m<sup>3</sup>/s tend to transport and deposit sand from the channel onto the floodplain based on several years of observation. If the thresholds for your system are relatively unknown, it is recommended that you utilize the figure developed in the previous step by locating the thresholds that space out the flood event data at regular intervals. This technique of fitting the thresholds to the data is not preferred but in some cases unavoidable.

Flood duration thresholds are inputted first and the user is asked to provide the total number of duration thresholds that will be entered. The thresholds must be entered in increasing order. Next, the peak flow thresholds are inputted and the user is asked to provide the total number of peak flow thresholds that will be entered. Please keep in mind that you should not enter the flood magnitude threshold as one of the peak flow thresholds. That threshold has already been used by the program to separate out the flood events. An error will occur if a peak flow threshold that is equal to or less than the flood magnitude threshold is entered. As in entering the duration thresholds, peak flow thresholds should be entered in increasing order.

Once the final peak flow threshold has been entered, the program groups the flood events into flood types and outputs three text files that contain various statistics related to the flood events and flood types. The first file is 'floods.txt' and automatically replaces the previous version of that file if the export option was chosen in the previous step. The format for this new 'floods.txt' file is identical to the one described in the previous step with the exception that another column has been added that stores the flood type code:

- Column 12: flood type code



**Figure A1.** Example of how flood types are arranged on a plot of peak flow versus flood duration for the Cosumnes River.

The flood type code is composed of two numbers separated by a pipe symbol ( | ); the first is the flood duration type number and the second number is the peak flow type number. Figure A1 shows an example of how the flood types are arranged on a plot of peak flow versus flood duration for the Cosumnes River.

The next step for the program is to generate statistics for each flood type. If you decide to exclude water years from this part of the analysis, the only variables that will change are the empirical frequencies of each of the flood types. These frequencies will be calculated by using the length of the streamflow record minus the number of excluded water years. Please note that if any empirical frequency values are greater than one, this means that there was at least one flood event that occurred in an excluded water year. If this occurs, the empirical frequency value is not accurate. To ensure the accuracy of the calculation, make sure that the excluded water years contain no real data in your input file (e.g., the streamflow values for those water years are all -99).

The second file is 'flood\_types.txt', which stores statistics for each flood type in the following format:

- Column 1: flood type code
- Column 2: duration range (in days)
- Column 3: peak flow range (in units of ft<sup>3</sup>/s (cfs) or m<sup>3</sup>/s (cms))



Column 4: mean start water year day (the mean of the day of the water year in that each flood event of that type started)  
Column 5: mean start water year fraction (the mean of the fraction of the water year in that each flood event of that type started)  
Column 6: number of occurrences  
Column 7: empirical frequency of flood type occurring one or more times in a given year  
Column 8: empirical frequency of flood type occurring two or more times in a given year  
Column 9: empirical frequency of flood type occurring three or more times in a given year  
Column 10: empirical frequency of flood type occurring four or more times in a given year  
Column 11: empirical frequency of flood type occurring five or more times in a given year

A file header is present to denote the contents of each column. Also note that flood types that are not occupied by flood events are also presented in this file.

The third file is 'ft\_wy.txt', which stores the number of flood days for each water year arranged by flood type with the following format:

Column 1: water year  
Remaining Columns: number of flood days associated with the water year in the first column and the flood type given in the header above

As in the 'flood\_types.txt' file, flood types that are not occupied by flood events are also included in this file.

### **Create water year types**

Once the flood types have been created and the number of flood days of each flood type for each water year is calculated, the program can be used to create water year types. The FRC program uses a statistical tool called hierarchical cluster analysis, which is discussed in more detail in the companion paper.

First, you are asked whether or not you want to exclude any water years from the analysis. If you decide to exclude some water years, these water years are eliminated from the hierarchical cluster analysis routine. This option can also be used to investigate the influence of certain water years on the creation of water year types.

After you decide whether or not to exclude any water years from the analysis, you are asked to enter the number of water year types that you would like to create. This program is designed to handle a wide variety of watershed applications and thus the appropriate number of water year types will change on a case-by-case basis. It is recommended that finding an appropriate number of water year types should be an iterative process. You should use your knowledge of differences between years that have been observed on your site to assess whether

or not more or fewer water year types are needed to adequately differentiate unique water years. This iterative process is aided by the next step that generates average hydrographs for each water year type. If you notice that two water year types are not adequately distinguishable, then you may want to try and reduce the number of water year types. Conversely, if you notice that two water years that you know have very different ecosystem responses are clustered into the same water year type, then you may want to increase the number of water year types.

Once you decide on a number of water year types, the program performs the hierarchical cluster analysis using the water years as a set of observations and the number of flood days for each occupied flood type as the variables. After the analysis is complete, the program outputs two text files that store statistics related to each water year and water year type. The first file is 'wy\_stats.txt', which stores several statistics for each water year in the following format:

- Column 1: water year
- Column 2: water year type
- Column 3: water year volume (total streamflow volume in units of million-acre-feet (maf) or cubic kilometers (km<sup>3</sup>))
- Column 4: center of mass (streamflow center of mass in units of water year day)
- Column 5: peak flow (peak daily flow in units of ft<sup>3</sup>/s (cfs) or m<sup>3</sup>/s (cms))
- Column 6: number of flood days
- Column 7: number of flood events

The second file is 'wy\_types.txt', which stores several statistics for each water year type in the following format:

- Column 1: water year type
- Column 2: number of occurrences
- Column 3: empirical frequency of the water year type occurring per water year
- Column 4: average water year volume (in units of million-acre-feet (maf) or cubic kilometers (km<sup>3</sup>))
- Column 5: average center of mass (in units of water year day)
- Column 6: average peak flow (in units of ft<sup>3</sup>/s (cfs) or m<sup>3</sup>/s (cms))
- Column 7: average number of flood days
- Column 8: average number of flood events

Notice that the water year type numbers are arranged in increasing average water year volume order (e.g., water year type 1 will have the lowest average water year volume out of the entire set).

### **Create average hydrographs for each water year type**

After the user selects this option, the program determines the mean streamflow value for each day of the water year for each water year type. The standard deviation for each water year day is also calculated. These values are then used to plot the average hydrograph for each water year type. A separate Matlab figure is created for each water year type. The x-axis is the water

year day and the y-axis is the streamflow (in the unit type specified by the user). The blue line is the mean streamflow value while the two dotted black lines are the mean plus and minus one standard deviation. The vertical red line indicates the average streamflow center of mass for the water year type.

Use these plots to visually observe the differences between each water year type. As mentioned in the previous section, these plots can also aid in determining how many water year types are necessary to adequately differentiate water years.

The user is then prompted as to whether or not they would like to export this data. If the user selects 'yes', a file called 'wytN\_hydro.txt', where N is the water year type number, is created for each water year type. Each file has the following format:

Column 1: water year day  
Column 2: mean streamflow (in units of ft<sup>3</sup>/s (cfs) or m<sup>3</sup>/s (cms))  
Column 3: mean streamflow plus one standard deviation  
Column 4: mean streamflow minus one standard deviation.

A file header is present to denote the contents of each column.

## **Other Potential Errors**

When the program is attempting to create a new text file, you may encounter the following error:

```
??? Error using ==> fprintf
Invalid file identifier -1.
```

This is most likely caused by having the text file, to which the program is trying to write, already open in another program (e.g., in Microsoft Excel). Simply close out of that other program and the error should not occur.

## **Viewing and Working with FRC output files**

The best way to view and work with the FRC output files is to use a preferred spreadsheet software program such as Microsoft Excel. Each of the output files can be opened in Microsoft Excel using the Text Import Wizard. With the exception of the 'flood\_types.txt' file, you can import the text files as space-delimited data or fixed-width data. For the file 'flood\_types.txt', it is recommended that you treat the data as fixed-width and make sure to change the data format of columns 2 and 3 to 'Text' instead of 'General' (this can be done in Step 3 of 3 in the Text Import Wizard).

## **Converting Water Year Day to a Date**

Two files are also included in the zip file 'FRC\_v1\_0.zip' that store tables of water year day versus the corresponding date of the year. The first file 'reg\_year.txt' stores the relationship for a non-leap year and the second file 'leap\_year.txt' stores the relationship for a leap year.

## **Cosumnes River example**

The same website from that the program files were downloaded (<http://watershed.ucdavis.edu/pages/programs.html>) also has a link to example output files from an analysis of the Cosumnes River. These files are contained in the zip file 'Cosumnes\_example.zip'. One of the files in this package is 'command\_window.txt', which stores the entire output to the command window during the analysis of the Cosumnes River. You can use this file to determine what thresholds were used and how many water year types were desired.

## **User Support**

All questions, issues, and comments should be directed to:

Eric Booth, [egbooth@gmail.com](mailto:egbooth@gmail.com)