RWater Module 3

Understanding Rainfall-Streamflow Relationship from Observed Gage Station Data

Adnan Rajib and Venkatesh Merwade

Lyles School of Civil Engineering, Purdue University

Learning Goals

From the hypothetical examples shown in the previous module, students already have the concepts of rainfall hyetographs and streamflow hydrographs, along with relevant terminologies such as Peak Discharge and Lag Time. The United States Geological Survey (USGS) has thousands of gage stations all over the USA, monitoring the streamflow. After completing this module, students will be able to:

i. download the USGS daily streamflow data for particular locations directly through RWater
ii. plot rainfall hyetograph and streamflow hydrograph by writing simple programming lines in RWater interface, based on the actual data at any USGS location
iii. visualize the effect of rainfall intensity over lag time and peak discharge in an interactive way using the graphs created by their own

---

- Load the script for this module from your working directory.
- Select certain segments of the script and run in steps as shown below.
- **Make changes only in the portion with "XXXX" or as directed**
- Relevant explanations associated with each step are also provided here. These explanatory notes are only for building user's perception over the code.

Downloading the USGS Streamflow Data

The objective for this portion of the script is to enable direct downloading of USGS daily streamflow data from any valid location and for any valid date range. The downloaded data can be used for subsequent plotting and analysis without any sort of pre-processing. In this module, we are going to use the data from USGS 04180000 which is a gage station for the watershed called Cedar Creek near Cedarville in Indiana, USA. Click on http://goo.gl/Y5lYZZ and you can see the watershed as well as the gage station in a customized Google map.

```
### STEP 1
### Removing previously used scripts from RWater
### Removing all previously generated datasets and plots
cat("\014")
rm(list = ls())
dev.off()

### STEP 2
### Loading a specific package
library(waterData)


### STEP 3
mysite<-c("XXXX")
mysiteinfo<-siteInfo(mysite)
```

> A package is a collection of R functions to serve specific analysis purpose. Package names are case sensitive (**waterData**: **D** is capital letter)

> Getting information about a USGS site. In this example, replace **XXXX** with **04180000**. This is the USGS gage station ID for Cedar Creek watershed in Indiana
>
> **siteInfo** is the function name which pulls the information from USGS website. Click on 'mysiteinfo' in the RWater workspace (upper right corner)

```
### STEP 4
### Downloading streamflow data directly from USGS
### Save the downloaded data with a name 'cedarflow'

cedarflow<-importDVs("XXXX",
        code="00060", stat="00003",
        sdate="YYYY-MM-DD",
        edate="YYYY-MM-DD")
```

In this example, replace **XXXX** with **04180000**.

importDVs is the function name which pulls the data from USGS website (case sensitive: **D** and **V** are capital letters)

You are saving the downloaded data with a name **cedarflow**.

Click on 'cedarflow' (upper right). You will see that streamflow values are listed in the right column (column header name is **val**)

"00060" and "00003" indicates 'Streamflow' and 'Mean Daily Data' respectively

Provide a range of date within **sdate** and **edate** in YYYY-MM-DD format. In this example, use 2009-01-01 and 2010-12-31 respectively

Loading a Given Rainfall Dataset from RWater Working Directory

The most comprehensive source of daily precipitation data is on the website of the National Climatic Data Center (NCDC). To aid the easy usage of this module, we have prepared the rainfall dataset for Cedar Creek watershed, which you can store in your RWater working directory. To know how to set up a working directory, please refer to previous modules.

```
### STEP 5
### Caution: The following file path will be different in your case
cedarrain=read.csv("/home/mygeohub/
                   put your user name here/
                   put the name of your working directory/
                   cedarrain.csv",header = TRUE,
                   stringsAsFactors = FALSE)
```

Load the given rainfall dataset from your working directory. File name is **cedarrain**

Use your own RWater **username** and the **name of your current working directory**

stringsAsFactors is case sensitive (**A** and **F** are capital letters). See Appendix for the meaning of the function

Click on 'cedarrain' in the workspace (upper right). You can now see the rainfall data set (two columns: one is date and the other one is the rainfall value in mm)

```
### To enable the subset function, 'date' has to be a character
class(cedarrain$Date)
```

Plotting Rainfall Hyetograph and Streamflow Hydrograph

```
### STEP 6
par(mfrow=c(1,1))
par(mar=c(5, 4, 4, 8) + 0.1)
barplot(cedarrain$Rainfall.mm.,cedarrain$Date,
       space = c(0,1),width = 0.5,
       ylim=rev(c(0,XXXX)),
       xlab="Days in year 2009-2010",
       ylab="Rainfall(mm)",
       main="Cedar Creek near Cedarville, IN",
       axes=TRUE,las=1,xaxt="n",
       col="light blue", border="light blue")
```

Option of creating a multi-paneled plotting window (see Appendix for details). Here, we will have only one figure in the Plots window

Plotting the rainfall values in the 'cedarrain' file as a bar diagram (**barplot** function)

Put the maximum limit of the Y axis in **ylim**. Replace **XXXX** with **200** for this example

rev indicates that values are to be plotted in a reverse Y axis

```
par(new=T)
plot(cedarflow$val,type="l",pch=21,
     col="red", lty=12,lwd=1.5,
     yaxt="n", ylim=c(0,XXXX),
     xlab="",ylab="",axes=T)
axis(side=4)
mtext("Streamflow(cfs)",side=4, cex.lab=1,
       las=3,line=3, col="black")
legend(300,5500,"Streamflow", col = "red",
        lwd=1.5, lty=12, bty="n")
legend(300,4500,"Rainfall", col = "light blue",
        lwd=7, lty=1, bty="n")
border=c("black")
```

par(new=T) will plot the new graph over the previous graph

cedarflow$val commands RWater to plot only the streamflow values from the **val** column in the **cedarflow** file

type indicates what type of graph should be drawn. Here 'l' (small letter L) means a line plot

Put the maximum limit of the Y axis. For this example, use 6500 in **ylim**

See Appendix for all other syntax (example: xlab, lty, lwd, side etc.)

You have just created a graph showing real-time rainfall and streamflow data for the Cedar Creek watershed in Indiana! To read the graph, pick any day from X axis and trace a vertical line up to where it intersects the plotted hyetograph/hydrograph. Reading horizontally to the left and right, you can determine the rainfall and discharge of the stream for that date.

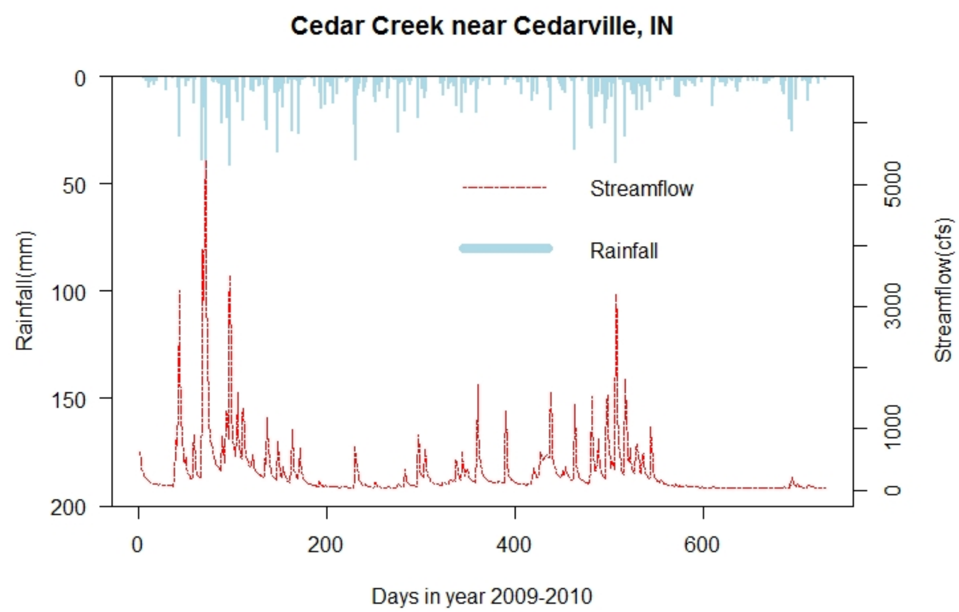

Figure 1

The graph which you have just plotted depicts vertical jumps and drops. What this means for the actual stream is that a stream exhibits increased flows immediately following the onset of a rainfall event and then return to pre-rain condition shortly after the end of rainfall.

Understanding Rainfall-Streamflow Relationship

In this step, we will split up the whole datasets being used in the previous step for a smaller duration of time. This will help to critically view the effect of rainfall intensity and temporal distribution over the streamflow response in an actual location such as Cedar Creek. This will also validate the theoretical concepts which we have discussed in Module 2, through a real-time situation.

```
### STEP 7

library(xts)
cedarx<-subset(cedarflow,
               dates>="2009-05-25"
               &dates<="2009-06-04")
cedarrainx<-subset(cedarrain,
                   Date>="2009-05-25"
                   &Date<="2009-06-04")
```

Clipping a date range directly from the original streamflow and rainfall datasets

We need an R package called **xts** for the sub-setting task

The new subset datasets are named as '**cedarx**' and '**cedarrainx**'

```
### STEP 8
### Plotting a segment of the whole downloaded time series

par(mar=c(5, 4, 4, 8) + 0.1)
barplot(cedarrainx$Rainfall.mm.,cedarrainx$Date,
        space = c(0,1),width = 0.5,
        ylim=(c(0,XXXX)),
        xlab="Days (May 25 - June 4, 2009)",
        ylab="Rainfall(mm)",
        main="Cedar Creek near Cedarville, IN",
        axes=TRUE,las=1,xaxt="n",
        col="light blue", border="dark blue")
par(new=T)
plot(cedarx$val,type="l",pch=21,
     col="red", lty=12,lwd=1.5,
     yaxt="n", ylim=c(0,XXXX),
     xlab="",ylab="",axes=T)
axis(side=4)
mtext("Streamflow(cfs)",side=4, cex.lab=1,
      las=3,line=3, col="black")
legend(5,900,"Streamflow", col = "red",
       lwd=1.5, lty=12, bty="n")
legend(5,800,"Rainfall", col = "light blue",
       lwd=7, lty=1, bty="n")
border=c("black")
```

Use appropriate name of the data set from which **rainfall data** needs to be plotted (here, this should be the new subset **cedarrainx**)

Put the maximum limit of the Y axis. For this example, use 60 in **ylim**

Use appropriate name of the data set from which **streamflow data** needs to be plotted (here, this should be the new subset **cedarx**)

Put the maximum limit of the Y axis. For this example, use 1000 in **ylim**

All functions have the same meaning as in STEP 5

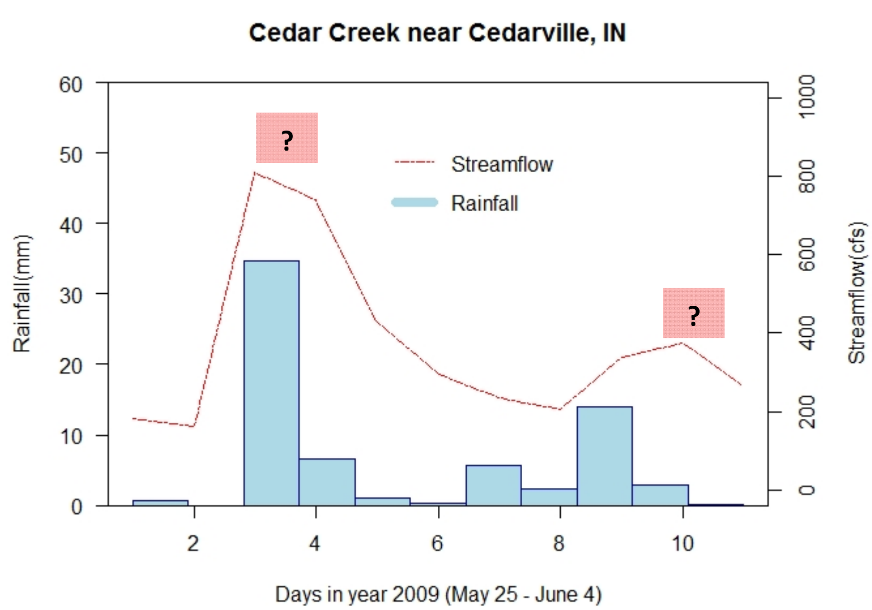

Figure 2

Quiz

1. From Figure 2 in the tutorial, we can identify two consecutive rainfall events of different intensity. Now, read the graph carefully and replace W, X, Y, Z in the following table from given options:

Rainfall Type	Maximum Rainfall (mm)	Peak Discharge (cfs)	Slope of Rising Limb (Steep/ Mild)	Lag Time (days)
High intensity	35	**X**	Steep	less than 1 day (almost instantaneous)
Low intensity	**W**	390	**Y**	**Z**

Take help from Module 2, if required.

a. W: 800, X: 35, Y: Steep, Z: less than 1 day
b. W: 15, X: 800, Y: Mild, Z: more than 1 day
c. W: 25, X: 800, Y: Mild, Z: more than 1 day

2. In order to download daily streamflow data directly through RWater, all we need to know is the _______
   a. County name
   b. USGS ID of the gage station
   c. Latitude and Longitude
   d. Watershed size