Laurel P. Goodell
Princeton University

Analyzing “Hometown Streams” using USGS NWIS data

For this lab, you will analyze discharge records for a stream or river of interest to you. You’ll download data from a gauging station, make graphs, and compile a report which will have the following sections:

A. Location of your station
B. Course of your river or stream
C. Annual pattern of precipitation for your river or stream
D. Annual peak discharge for your river or stream
E. Discharge recurrence interval for your river or stream
F. Interpretive essay
G. References

FOR THE PRE-LAB ASSIGNMENT, YOU MUST COMPLETE ITEMS A AND B. This involves searching an on-line database, so don’t leave it until the last minute. Bring your work on your computer, or email yourself electronic copies you can access in lab. You’ll work on the other sections during your lab session (but feel free to work on them beforehand as well). During lab, you’ll also see examples of how to interpret discharge data.

BACKGROUND INFORMATION ON YOUR RIVER Once you have selected your river (Part A), arm yourself with background information. This can come from personal experience and/or background research, and after your lab session you will likely need to do further research. For instance, you might:

- Interview relatives or friends.
- Research local newspapers for items of interest about your river (e.g. notable storms or floods).
- From the website for your gauging station, check out the items available in the drop-down menu under "data available for your site" -- especially check out "Offsite EPA Surf your watershed."
- “Google” your stream plus the state name plus the words "water resources." This may bring you to "official” sites. If that doesn't help, try searching for info on a larger stream that your stream runs into.
- “Fly to” your stream using Google Earth. You might also try “flying” along it from source to mouth.

Here are examples of background issues to consider:

- What is your personal experience with the stream or river?
- What are the flood/drought history of the area and the chances of future floods/droughts? Is there a flood/drought or wet season/dry season every year?
- Are there any restrictions to development along the river?
- Are there artificial controls (e.g. dams, levees, channelization, sea walls, jetties) on the rivers or coastlines in this area?
- What are the purposes of the controls (e.g. flood control, power, transportation, erosion control)?
- Have the controls been successful, expensive, or controversial?

After your lab session, as your analysis and write-up progress, you may want to do additional research.

OTHER CONSIDERATIONS FOR YOUR REPORT: Organize things so they are clear and easy to follow. Use a descriptive title. Number your figures and graphs logically, and refer to them in your text. Integrate your technical analysis in a significant way - your lab instructor will show examples of how to do this. Discuss your results and analysis with lab staff during class, or make an appointment to do so.

NOTE: Shaded in grey in this handout are worked-out examples for the Princeton Stony Brook station, as models for work on your own station.

Laurel P. Goodell
Princeton University

A. SELECTING A STATION AND SHOWING THE LOCATION OF YOUR STATION

Find a river/stream near your home, or some other area of interest. (Vacation spot? Summer camp? Grandma's house?) If you are not from the U.S., you could try searching for records available for non-U.S. rivers, or try something near Princeton or some other U.S. river/stream of interest (The Colorado River in the Grand Canyon? The Mississippi River in New Orleans? The Delaware or Hudson Rivers?) To search:

1. Go to the National Water Information System Mapper at
   http://wdr.water.usgs.gov/nwisgmap/index.html
2. In the upper left hand corner, make sure the "Include" drop-down menu specifies"Active sites." Then, below the "Welcome," make the box for Surface Water sites is checked, and none of the others.
3. Start browsing for a site by either:
   - Using the zoom buttons, direction arrows, and/or "Zoom Box" function to focus on an area, or
   - Entering a state, zip code, or address (yes, even your home address), and then clicking "GO."
4. As you zoom, you will see triangles marking gauging stations. If you zoom in enough, hovering your pointer over a station turns the cursor into a pointer-finger (if it doesn't, you haven't zoomed in enough). Clicking on a triangle with the pointer finger will bring up the station number and name, and then clicking on "Access Data" brings you to the website for that station.
5. Once you are at a station's website, use the drop-down menu entitled "Available data for this site," and choose "Surface-water: Peak streamflow" to check that the site has sufficient data for your analysis. For a station to be appropriate, it must:
   - Have **at least 30 *consecutive* years** of "Peak streamflow" data – the longer, the better.
   - Include recent years, at least to 2010.

   Here are two examples, the first a station that is not appropriate, the second a station that is:

   http://water.usgs.gov/nwis/peaks/?site_no=11174000&agency_cd=USGS
   This river in California lists data from 1910-2010. But the graph only shows data for years 1913-1930, 1956, 1961-1965 and 2000-2010. So there are not 30 consecutive years of data.

   http://water.usgs.gov/nwis/peaks/?site_no=11180500&agency_cd=USGS
   This river in California lists peak stream flow data starting in 1917. These data are only consecutive from 1959-2010, but this is long enough (51 yrs).
6. Once you have found a suitable station, under "Available data for this site" choose the site map. Figure out a way to capture this image, or find some other maps or images, to clearly indicate the location of your station. (You can also do this with Google Earth by zooming into a field of view and then using File/Save/Save image.) Annotate your maps/images to clearly indicate the site location.

7. As soon as you have found a station, email the following information to laurel@princeton.edu:
   - The station name
   - The station number
   - The latitude and longitude of the station

B. DESCRIBING THE COURSE OF YOUR RIVER/STREAM

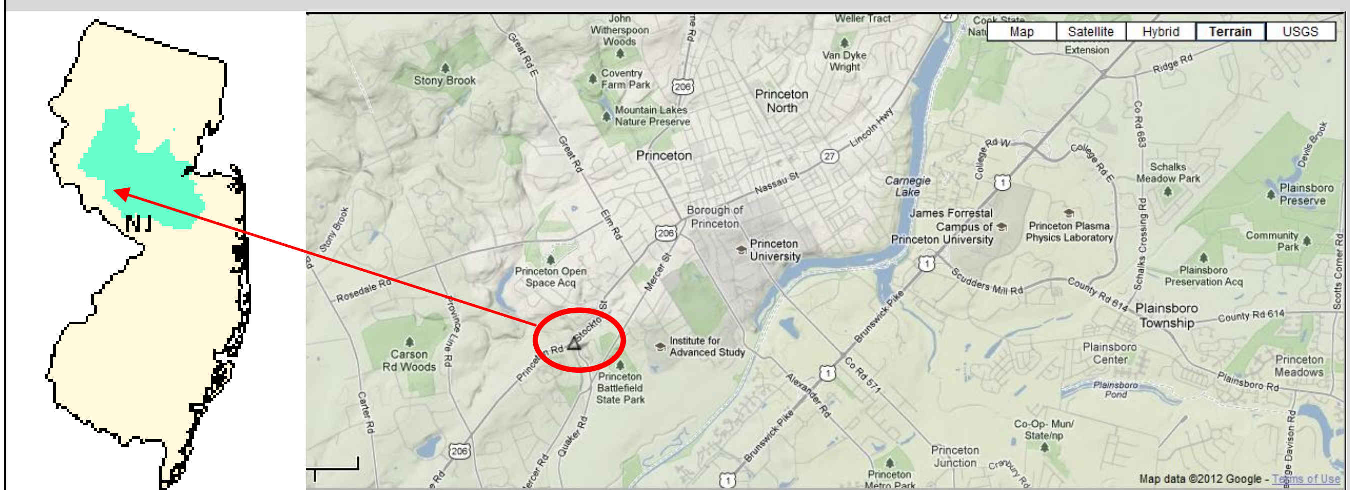
Describe the course of you river both upstream and downstream, i.e. from its headwaters to the ocean AND note where the gauging station is relative to the entire course of the river. Use Google Earth, or consult an atlas or map to answer this completely.

Laurel P. Goodell
Princeton University

A. Site information for USGS gauging station 01401000

Stony Brook at Princeton NJ

Mercer County, New Jersey
Hydrologic Unit Code 02030105
Latitude 40°19'59", Longitude 74°40'55" NAD83
Drainage area 44.5 square miles
Gage datum 62.23 feet above sea level NGVD29



B. Course of Stony Brook

Stony Brook has its headwaters in Hopewell Township northwest of Princeton. It then flows southeasterly into Lake Carnegie, which is formed by the damming of Stony Brook and the Millstone River. The stream leaves the dam as the Millstone River, which flows northeastward into the Raritan River, which then empties into Newark Bay and the Atlantic Ocean. The Raritan River drainage basin is shown in the figure at left.

Interesting note: There is no direct connection between the Delaware-Raritan Canal and the regional drainage, even though the canal is adjacent to Stony Brook and Lake Carnegie (and other streams) in places. An aqueduct carries the canal over the Millstone River where the river empties into Lake Carnegie.

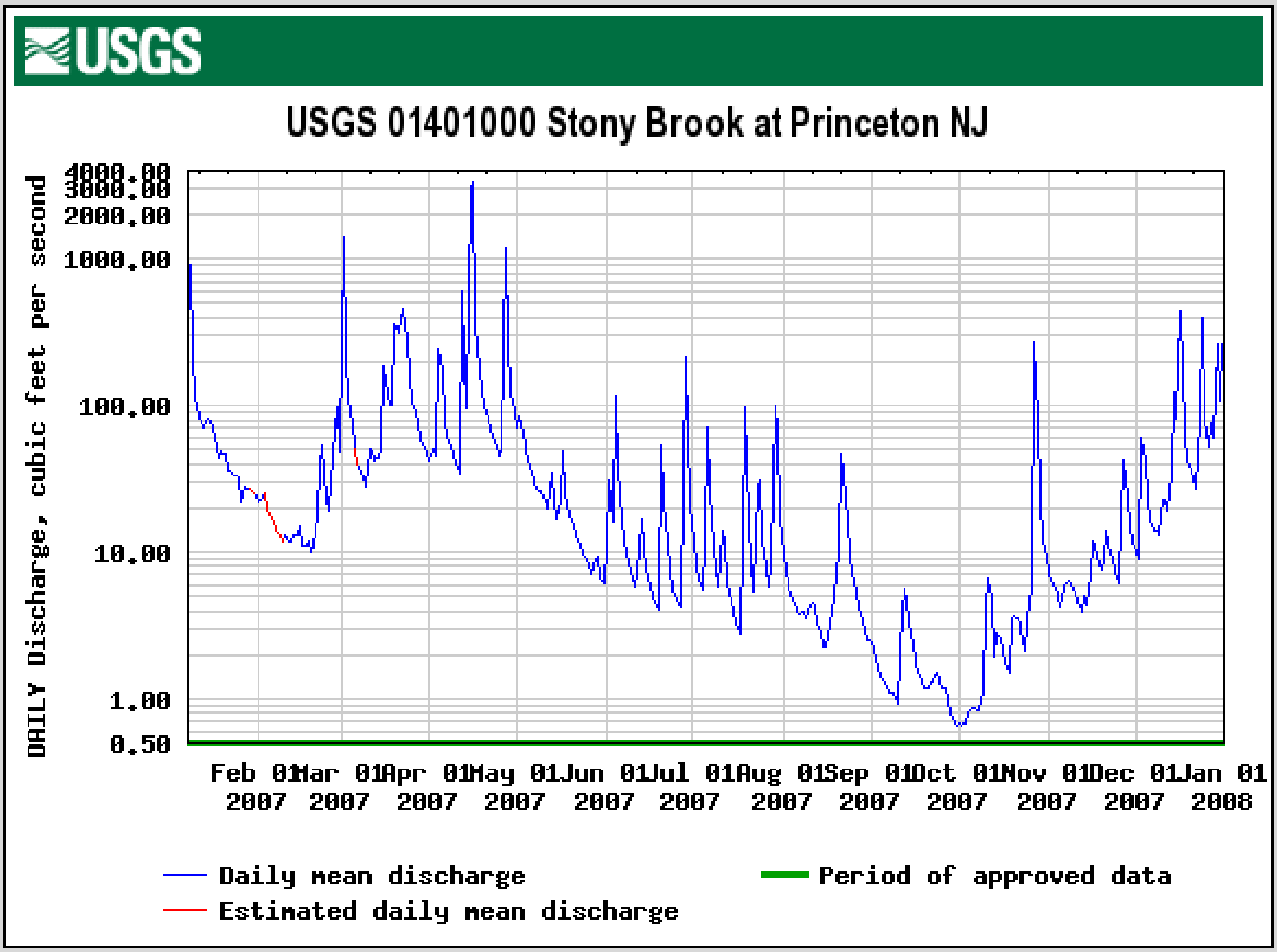
The gauging station is 4.5 miles upstream of where Stony Brook flows into the southwest end of Lake Carnegie.

Laurel P. Goodell
Princeton University

C. ANALYSIS OF THE ANNUAL PATTERN OF DISCHARGE FOR YOUR STREAM

1. From the website for your station, go to "Available data for this site" and choose "Time-series: Daily data." You'll get a webpage with a year's worth of annual discharge values. You can view other years by entering other "begin" and "end" dates. Look at a several years to get an idea of the typical annual pattern. Consider looking at early years and compare them to more recent years.
2. Save/print out the graphs for one or more years (choose "Create presentation-quality graph" before you save/print). Include at least one "typical" year, but you might also include years with significant droughts or floods.
3. Discuss how discharge varies over a typical year. You might discuss wet and dry times of the year, frequency and duration of storm events at various times of the year, etc. Has the annual pattern changed over time? Put comments or labels on the graph itself as appropriate.

C. for Stony Brook: Daily discharge data from Jan 1- Dec 31, 2007



Stony Brook flows all year long, with base discharge of about 5-100 cubic feet per second (cfs) from September through May. Several rainstorms a month cause episodic spikes in discharge of several hundred cfs; these last a day or two before dropping back down to base level. 2-4 events a year cause discharges of several thousand cfs. Summer and early fall are typically drier, with discharge dropping below 1 cfs.

Laurel P. Goodell
Princeton University

D. **<u>GRAPH OF ANNUAL PEAK DISCHARGE FOR THE LENGTH OF RECORD</u>**

1. Under "Available data for this site," click on "Surface water: Peak Streamflow" to display a graph of the largest event vs. year. Instead of printing out this graph, you will make your own.
2. To retrieve the data, click on "Tab-separated file." Ignore the explanatory material, and scroll down to Select/Copy and Paste the data into a blank Excel spreadsheet
3. Now in Excel, use Data/Text to Columns to split your data into columns so that the calendar year and the peak annual discharges are in separate columns. Delete all columns EXCEPT for the columns for the year and for the peak annual discharge. Edit the data as follows:
4. Some stations have gaps in the early years. Delete rows for early years to get a consecutive record.
5. Some years may list two peak annual flows and some may list none. This is because these data are collected on the basis of a "water year," not necessarily a calendar year. Edit the spreadsheet so that there is 1 point for each water year. *(If you discuss a particular peak discharge event in your essay, remember that you have done this; be sure to go back to the website to confirm the actual date.)*
6. Cut, insert, copy and paste columns until you get discharge in a column to the right of the column listing calendar year.

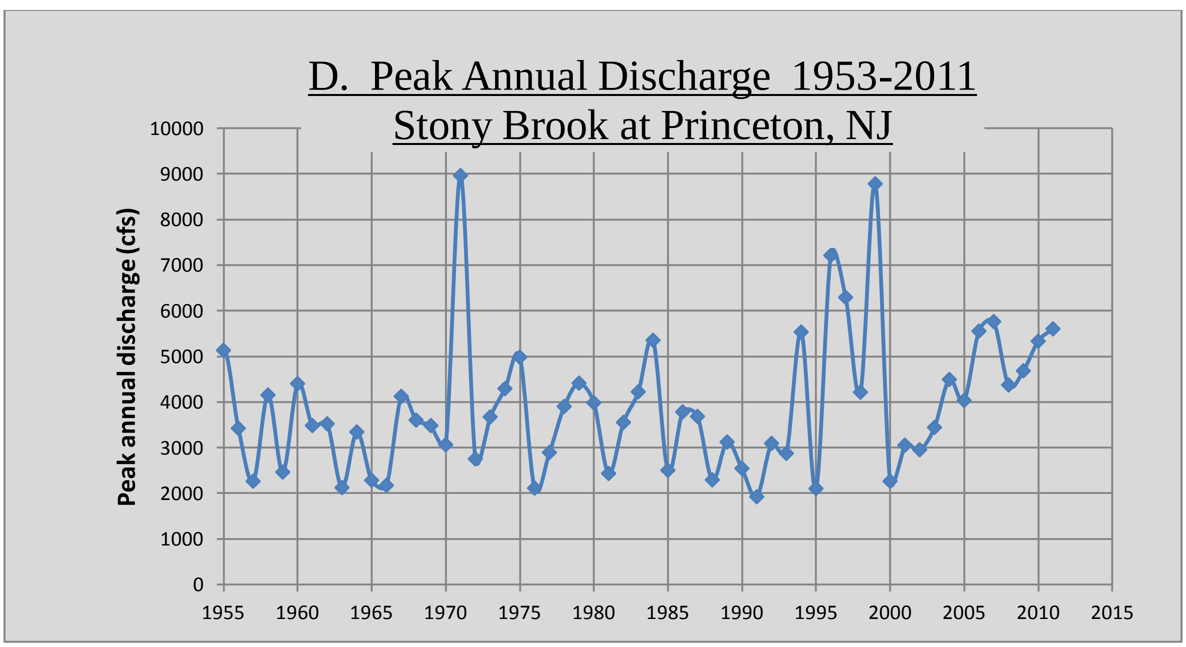
Raw data:

USGS	01401000	1953-12-14	3000	8.66
USGS	01401000	1955-08-13	5130	11.90
		↓		
USGS	01401000	2011-08-28	5600	10.60

Becomes:

1954	3000
1955	5130
↓	↓
2011	5600

7. Make an x-y scatter graph of peak annual discharge (y-axis) vs. year (x-axis), with straight lines connecting the points. Use linear scales for both the x and y axes. Format the graph appropriately.



8. When your graph is complete, copy your data onto another worksheet, preserving this version one with the graph. Make your next graph on the new worksheet.

Laurel P. Goodell
Princeton University

E. GRAPH OF DISCHARGE RECURRENCE INTERVALS, WITH ANNOTATIONS

1. On the new copy of your worksheet, delete the data for calendar year so that you are left with only a single column with the annual peak discharge data. Use Data/Sort to sort this column in order of **de**creasing discharge, from highest (top) to lowest (bottom).

2. Make a new column and assign a "rank" to the ordered discharges: "1" for the year with highest discharge, "2" for the second, and on down the list to the lowest discharge. (You can enter the numbers manually, or enter 1 then use Edit/Fill/Series, or enter and then select both 1 and 2 and then move your curser to the left bottom corner of the cell with the 2 and drag down to the bottom of the column.)

3. Determine and make a column for the recurrence interval for each discharge data point:

   a) Define N as the number of consecutive years for which you have a peak annual discharge for your stream). N for your stream = _______.

   b) The Recurrence Interval (RI) in years for a discharge event is defined as (N+1)/rank. For Stony Brook, for example, there are 58 years of consecutive data so N+1=58 and:

      the RI for the largest event = 58/1 = 59 yrs,
      the RI for the second largest event = 58/2 = 29.5 yrs,
      the RI for the third largest event = 58/3 = 19.667 yrs
      and so on.

rank	RI	discharge
1	59.000	8960
2	29.500	8780
3	19.667	7210
4	14.750	6290
5	11.800	5760
↓	↓	↓
54	1.093	2170
55	1.073	2120
56	1.054	2110
57	1.035	2100
58	1.017	1920

   To do this for your data, insert a column to the left of the column of discharge data. Select a cell in this new column and enter the RI formula by typing the following:

   "=" "(your N+1)" "/" " the cell address of the rank" and press Enter.

   Excel will then enter the RI into the cell. Once you have entered the formula for one rank, use Edit/Copy and then Edit/Paste to copy the formula from this cell to the rest of the RI column (you can also click and drag).

4. Make an x-y scatter graph of discharge (y-axis) vs. recurrence interval (x-axis). Format it using the Stony Brook example as a model: use logarithmic scales for both axes (so it is a log-log plot), adjust the ranges of the axes appropriately to show the data as clearly as possible, use appropriate titles, etc.

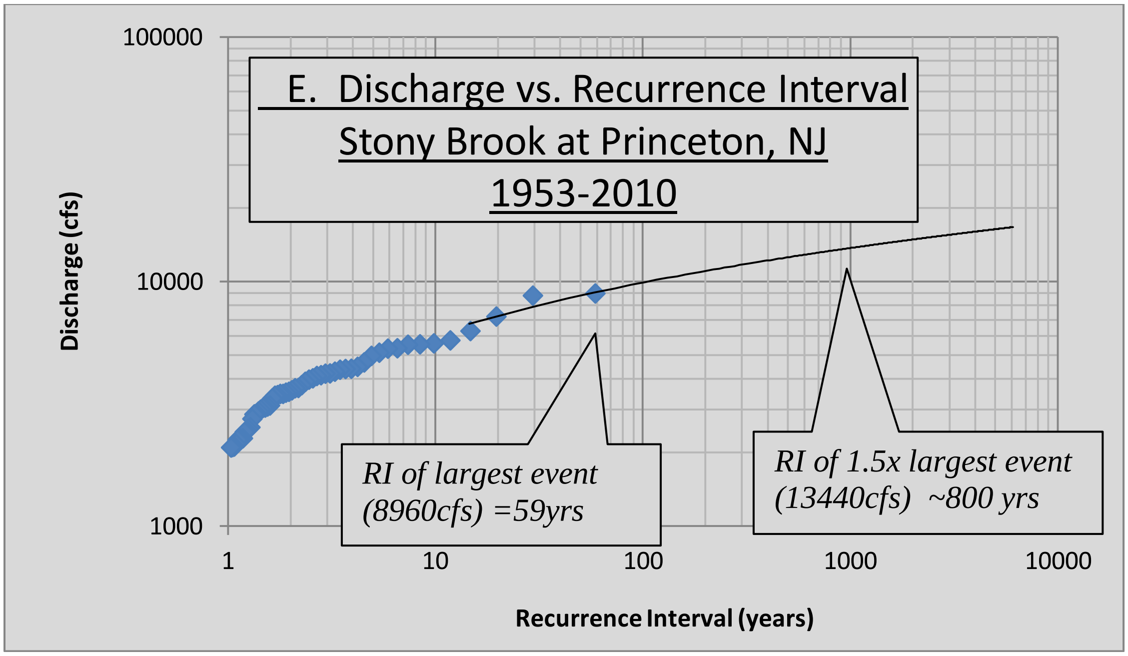
   This is a recurrence interval curve, discussed in your text book. This kind of analysis is similar to that used to forecast frequency and magnitude of earthquakes.

   Annotate and adjust your graph as follows:

   a) Indicate and label the RI of the largest peak.

   b) Forecast and indicate the RI for a discharge event 1.5 times that of the largest annual peak. To allow extrapolation, you will likely have to adjust the maximum values of your axes (do this by powers of 10 since it is a log-log plot.) It is probably best to do extrapolate to longer recurrence intervals by hand, but you could try "Add Trendline," and see if any of the choices there - linear, exponential, logarithmic, etc.- give an appropriate fit.

Laurel P. Goodell
Princeton University



F. <u>SEVERAL-PAGE INTERPRETIVE ESSAY</u>

Consider one or more the topics listed below or other topics suggested by your data, background research and personal experience. You must refer to and integrate your technical analysis in a significant way. This is to be an interesting analysis, not a comprehensive analysis; choose one or two interesting aspects to discuss. The suggestions below are to get you thinking, and are *not* to be a list of questions for you to answer. You are also encouraged to come up with your own topics, and discuss issues with lab staff.

- Has the pattern of discharge changed over time? Why?
- Discuss notable high water events (e.g. floods, hurricanes) and/or low-water events (e.g. droughts). What are the probabilities of discharge events of various sizes occurring over the next 10 years, 50 years, 100 years, etc?
- Is someone in the class studying a station upstream or downstream of you? If so, compare information and analyses. Check Blackboard as the week progresses, to see who's doing what stream.
- Is the flood plain of your river developed/urbanized? Has the extent of development changed over time and can you see the effect of this in the data?
- How are developments along your stream prepared for flooding?
- Are there artificial controls (e.g. dams, levees, channelization) on the rivers in the area? If so, what are their purposes (e.g. flood control, transportation, erosion control)? Have the controls been successful or controversial? Why or why not?
- Would you suggest restricting development on the 10-year flood plain? 50-year flood plain? 100-year flood plain? What should policy be for rebuilding in areas damaged by flooding? If development is restricted, what can the land be used for?
- What should insurance policy be in areas vulnerable to flooding? Is this is your home area, are you living in a high risk area? Do you have flood insurance? Should you get it?
- How would your analysis change if you were to eliminate the largest discharge event? That is, how sensitive is your analysis to this one event?

Laurel P. Goodell
Princeton University

G. <u>CITATION AND LISITING OF REFERECES</u> –

The paragraphs below illustrate the citation style you should use.

Do not use footnotes; instead use the style of citation demonstrated on this page and summarized nicely at Cornell University's library website (Cornell University Library, 2011), which is the only reference on the list below that actually exists. It is not necessary to reference a statement that is common knowledge, or in the case of scientific knowledge, can be found in numerous (generally greater than three) publications.

It is, however, necessary to cite new and specific information in support of your arguments, such as the number of earthquakes observed in 1996 in San Francisco (Shakey et al., 1997) or the distance of movement observed during the last quake on the San Andreas fault (Slink and Jones, 1998). References are usually cited by author's name immediately after the information attributed to the author, or directly in the text, followed by the date in parentheses. Slink and Jones (1998) may also have commented upon the relative distance traveled by fence posts in the central valley in several preceding earthquakes, but their information was disputed by subsequent authors (Smith and Johnson, 1999; Abel et al., 2001).

For internet resources, follow the style for the Cornell University Library (2011) website given below.

Direct quotations are generally not used; rather the information is restated in the writer's own words and attributed appropriately, according to Ward (2000). The first person active tense is not usually used in scientific writing. This helps avoid disagreements between persons and keeps discussions focused on the data and arguments themselves, rather than the people discussing them.

However, if some of your information is from personal communication (oral or written) is anecdotal, you might directly quote that person as follows: "There was never any flooding in town until that shopping mall was built over the county fairgrounds," (Mary K. Smith, the author's grandmother and local resident, personal communication, April 8, 2011.) Do not include personal communications in your list of references.

At the end of your paper, include a list of references in alphabetical order by author's last name; similar to the example below.

REFERENCES

Abel, P. and M. Jones, 1998, Lateral slip distances associated with the San Andreas fault. Internat. Jour. Slippery Slopes, 143, 34- 44.

Cornell University Library. "Citation Management," Cornell University, <http://www.library.cornell.edu/resrch/citmanage/mla#web_page>, accessed 7 April 2011.

Shakey, J., M. Rolling and P. Flip, 1997, Frequency of earthquakes in San Mateo and San Francisco Counties, Jour. San Franciscan Seism., 66, 123-135.

Slink, Q., R. Peterson and T. Williams, 2001, Abel and Jones don't know what they are talking about, Jour. of Disputed Results, 54, 10-22.

Smith, W. and E. Johnson, 1999, Re-evaluation of slippage at the San Andreas, *in* Reviews of Regional Earthquake Activity, T. Andrews and Y. Knot, editors, Bucking Publishers, Los Angeles, 25-47.

Ward, B., 2000, Many tedious examples of proper citation and referencing in scientific papers, Academic Publishers, 267p.