

Volcanic Gases Tutorial for CIDER 2019

Tobias Fischer and Kristen Rahilly

This tutorial aims to instruct students in practical collection of CO₂ flux data from an area on the Berkeley campus and come up with very rough estimate of the CO₂ flux for the chosen area (part 1). The second part of the tutorial uses gas flux data collected from Horseshoe Lake from Long Valley Caldera to practice statistical treatment of flux measurements. In the third part you will share your collected flux data with that of the other groups and analyze all results. You will need a laptop with excel and MatLAB or share with somebody.

1. CO₂ flux measurements: *Students need to work in groups of three*

What you need: CO₂ flux meter (provided), means to record data (notebook and pencil).

What you need to do: select an area of your choice on the grassy area of the UC Berkeley campus and collect a few (~ 10) CO₂ flux data, record the fluxes and type of soil/vegetation. You should record the type of soil/vegetation you measured on. This should take you no more than 15 minutes.

Try to estimate the total flux of your area – more about this below.

2. Horseshoe Lake Data Set

This data set is from Cardellini et al., 2003 JGR. PDF is provided. The data is in the excel file (HorseshoeLakeFlux). You will use this data set and the Sinclair statistical method to evaluate the flux data set to define the background and above background fluxes. The Sinclair, 1974 paper is provided for your reference.

Part 1: Use excel to construct the Sinclair plot. The detailed instructions are below.

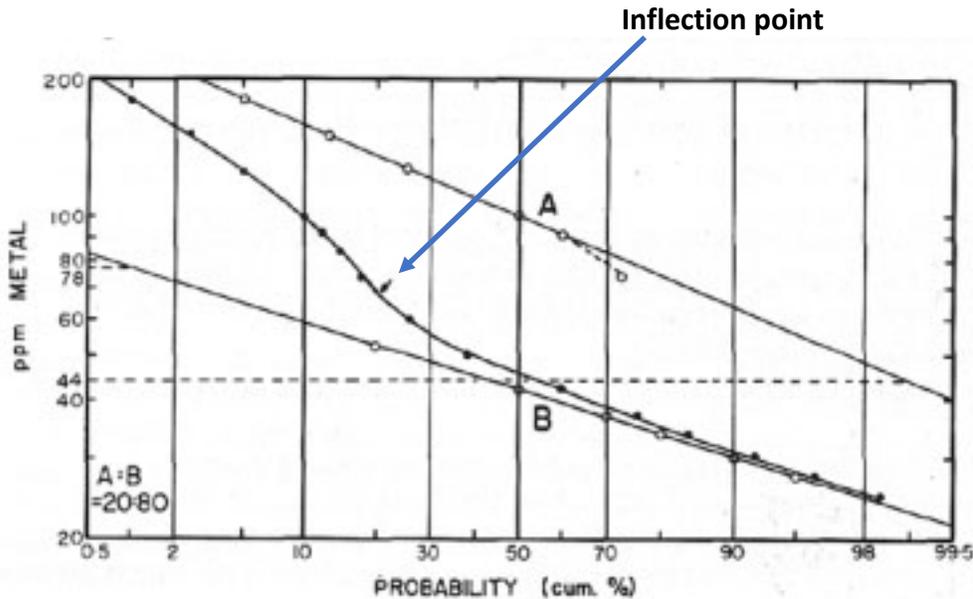
The steps for getting cumulative percentage in Excel are:

- 1.) Order data from smallest to largest
- 2.) Copy data into adjoining column. Then remove duplicates (under "Data" tab). Label this column "Bin"
- 3.) Click on the "data analysis" button under the "Data" tab (possibly you will have to add this as a tool. Go to "tools" --> "Excel Add-ins" --> select "Analysis Toolpak").
- 4.) The Data Analysis button will bring you to a list of options. Select "histogram"
- 5.) In the histogram window, your "input range" is your original data; "bin range" is your "Bin" column. Select "output range" and enter in the column where you would like the data to be...this doesn't matter too much just pick a free column.
- 6.) Finally, check the box labeled "cumulative percentage"

- 7.) When you have the cumulative percentage column calculated, calculate reverse cumulative percent by subtracting from 100 (the cumulative percent cells might be strangely formatted so you will have to be careful here).
- 8.) Plot your reverse cumulative percent on the x axis; "bin" column on the y axis.
- 9.) Make your X and Y axis logarithmic

Selecting inflection point (we can try for one inflection point for practice) and calculating populations:

- 10.) Eyeball the point of inflection for the curving data plotted in steps 1 – 8. The point of inflection is where the slope of the curve changes. Below is what we are shooting for from **Sinclair, 1974**. The small black arrow is the inflection point for this data.



- 11.) Let's say your point of inflection is at cumulative percent = 20% (like the example figure above). This means you have 20% of a high population (POPULATION A) and 100-20 = 80% of a low population (POPULATION B).

12.) To plot the points for population A, you need to normalize to that 20%. If we look at the highest value data point (approximately 180 ppm METAL), this point plots within the whole dataset at cumulative percent = 1%. Within population A (composing 20% of the population), however, this point will plot at point $(1/20) * 100 = 5\%$. This means that point 180 ppm comprises 1% of the *whole* dataset but 5% of population A. So, we can see in the figure above that we then plot the 180 ppm point at the 5% cumulative percent (open circle along line A). We continue plotting points in this way until we get a general trend of the linear line comprising population A.

You can set this up in Excel.

- 13.) YOU CAN do the same for B but it is not necessary for this exercise to save time. Skip to 14)

However, this time we will be doing everything in the reverse. So, for example, a data value of around 38 ppm is at cumulative percent of about 76%. For the B population, this would be at $100 - 76 = 24\%$ so now we use this original reverse \ cumulative percent (25%) and divide it by

the B population total (80%) to get $(24/80) * 100 = 30\%$. So 30% of population B is smaller than 38 ppm. However, remember that we now need to plot this point on our whole dataset so we need to plot 38 ppm at $100-30=70\%$. Keep doing this for the rest of the B population until you get a clear data trend.

14) Now check where your population line intersects the 100% value on the x. This is your threshold for the high population. You can now go back to your values and calculate an average of all the data above this threshold. That is your above background CO₂ flux for HorseshoeLake Caldera.

15) You can do the same using MatLAB. If you have MatLAB, you can use the code provided and modify so that the HorseshoeLake data are read and analyzed to return the probability density function values. The intersection of the two curves gives you threshold value of the two populations. You can then find that threshold value and go to your data and again calculate the average above background flux.

3. Analyze the UC Berkeley data you collected

After your group collected the data, share them with the other groups to get as many measurements as possible. Use either excel or MatLAB to distinguish your flux populations. Is there a high magma-derived CO₂ flux under the Berkeley campus?