

Summer school on hydrology – Part II

Upper Naryn, Tien Shan, Kyrgyzstan, discharge measurements and analysis

Eric.Pohl@unifr.ch



Table of contents

1. **Introduction**
2. **Programs (HOBOWare, R, RStudio)**
3. **Data preparation (water level loggers “HOBO”)**
4. **Data preparation (fluorometer “FL30”)**
5. **Fluorometer discharge calculation (RStudio)**
6. **Rating curve (water level to discharge; RStudio)**

The presentations from the summer school can be found here:

Overview discharge measurements: <https://t.ly/M0I6l>

Dye (color) tracer experiments: <https://t.ly/Qat2Y>

Introduction

During the Arabel Summer School 2023, a strategy was presented for the continuous monitoring of stream discharge [m^3/s] in the Upper Naryn, Kyrgyzstan. This strategy is based on the continuous monitoring of water level [cm] and actual discharge measurements. These two methods need to be combined by generating a so-called “rating curve”. The rating curve describes the relationship between the water level and discharge. To obtain a rating curve, it is necessary to measure the discharge for different water levels. At best, the full range of water level variability that a stream or river is experiencing is used for this. While this was not possible because winter discharge might be extremely low, it is possible to cover a large variability due to a strong diurnal melt and discharge variability.

The instruments that were and are used for water level are automatic pressure gauges for the raw water level, and automatic atmospheric pressure gauges. The measured pressure under the water is reflecting not only the water height, but the water height plus the atmosphere (the pressure that the atmosphere on top is exerting on us). Changes in the atmospheric pressure can be significant with respect the water level changes. A difference in atmospheric pressure, e.g. between 970 hPa and 1020 hPa correspond to a water height difference of ~ 50 cm! We therefore need to correct the changes in atmospheric pressure by subtracting the atmospheric pressure from the pressure measured under water. During the summer school we used Onset HOBO loggers. To read out the data, the software HOBOware is needed. It allows us also to correct the pressure data.

The water level [cm] does not provide us with the information about “How much water is flowing?”. To generate the rating curve, we need actual discharge [m^3/s] measurements. These were made with the fluorometer (FL30). The FL30 measurements do not directly tell us how much water was flowing at the measuring time. It gives us a concentration or “breakthrough” curve: How did the color concentration change over time. We use the so-called “**global method**” to calculate discharge from this curve. This method requires to know how much color/tracer (can also be salt) was injected into the water, and how strong the signal is measured by the FL30. For this, we prepared the color before going into the field by carefully measuring 50 g of color, and by preparing 50 ml vials with 50 μg color solutions. In the field, we calibrated the FL30 with one or two of the vials. In the calculation we will use this information. That is why it was important to write down how much tracer we used at what time.

The global method works by relating the amount of tracer injected to the concentration measured over time. One crucial prerequisite is that the color is evenly dispersed/mixed. This means that the concentration everywhere over a cross section must be the same. This allows us to measure only at any place, like at the shoreline of the river, without knowing how the cross section looks like. The important parameters to know are then just the amount of tracer (how much grams) and to have the concentration curve; also called the breakthrough curve. The methods is in more detail explained in the PowerPoint presentation.

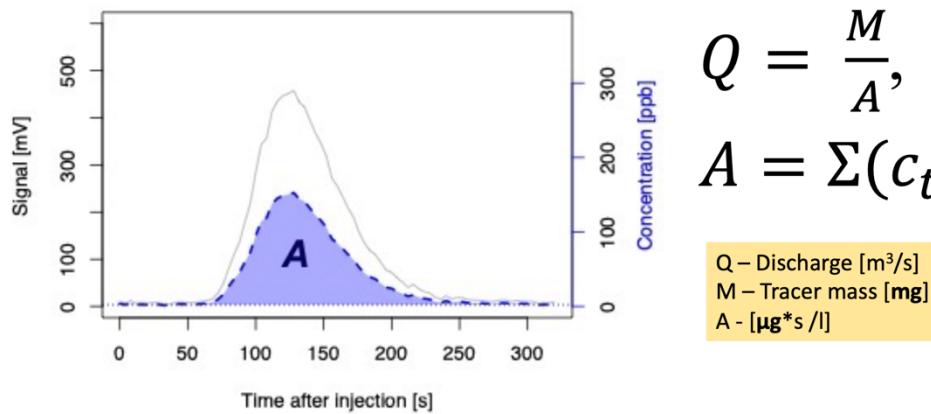


Figure 1: Discharge measurement using the global method. “Q” is discharge in m^3/s , “M” is the tracer mass in mg , and “A” is the area under the curve. The different units (mg and μg) result from the factor 1000 to go from liter to m^3 .

Programs

The main part of the analysis is happening in the freely available software **R** and **RStudio**. R is a program developed for statistics, but it can also be used for some automated calculations and to generate plots. We use this to bring together the different types of data, make sure the time format is consistent, and use its statistical methods to obtain models of how two variables relate to each other (linear regression). RStudio is a graphical user interface for R that allows for interactive plots that we will use when analyzing the concentration curves. To read the water level and atmospheric pressure loggers, the software **HOBOWare** is used. In this course, the data will be provided as comma separated text files (CSV or txt). These files can be opened in any text editor, or in OppenOffice, LibreOffice, or Microsoft Excel. After preparing this file (see “Data preparation FL30”), the needs to be saved as **.xlsx**. This can be done with Excel, LibreOffice, or OpenOffice.

The main processing takes place in RStudio because it allows us to perform the individual steps for the discharge calculation:

- 1) Background noise removal
- 2) Calibration of the signal from mV to $\mu\text{g}/\text{l}$ (=ppb)
- 3) Selection of the “breakthrough curve”
- 4) Calculation of discharge
- 5) Calculating the rating curve

This course does not require any previous programming experience. We use so-called **scripts**, which are text documents that have the R code inside them. They will be executed by RStudio, like small programs. These scripts are in text format and RStudio will transform them into computer programs that will do the calculation for us. We will, however, make some adjustments to these scripts to change **arguments**. These arguments are for example how much color tracer we injected, or the name of the spreadsheets to be read in.

Inside RStudio there are different windows. The main window shows the **script**, the bottom window the **console (it shows the output of the code)**, and on the right-hand side on the bottom the window shows **plots**. In the script window, we have the lines of code for each script. A line of code can be run or executed one by one. Under Windows and Linux, this is

done by pressing the buttons [Ctrl] + [Enter], under Mac by pressing [command] + [Enter]. Every line that is run will produce some output that will show up in the **console** or as a plot in the **plot window**. We will exercise this together.

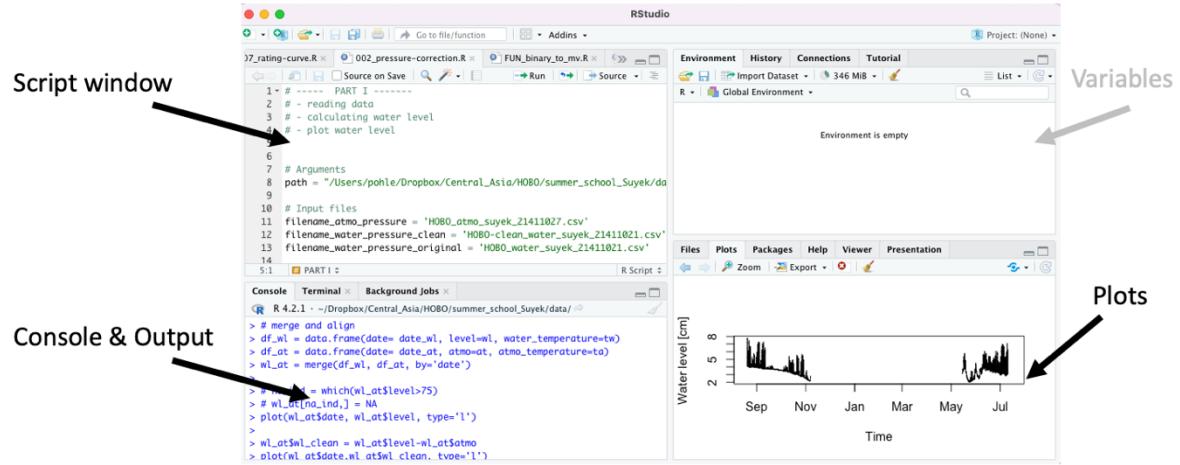


Figure 2: RStudio overview with the different windows.

Data preparation (HOBO)

The water level time series build the basis for generating the discharge time series. The preparation requires the two files: 1) water level time series, and 2) atmospheric pressure time series. Usually both time series come directly from the instruments installed in the field. The water pressure logger in the water, and the atmospheric pressure logger somewhere near by (can be some 100 m away). If the atmospheric pressure data is lost, it is also possible to use pressure data from a nearby automatic weather station or some other meteorological station. The easiest for processing is if both time series have the same measuring interval. If they do not have the same interval, one needs to interpolate the atmospheric pressure data to match the time steps of the water level data. Alternatively, one can leave out some water level data points to match the atmospheric pressure data points.

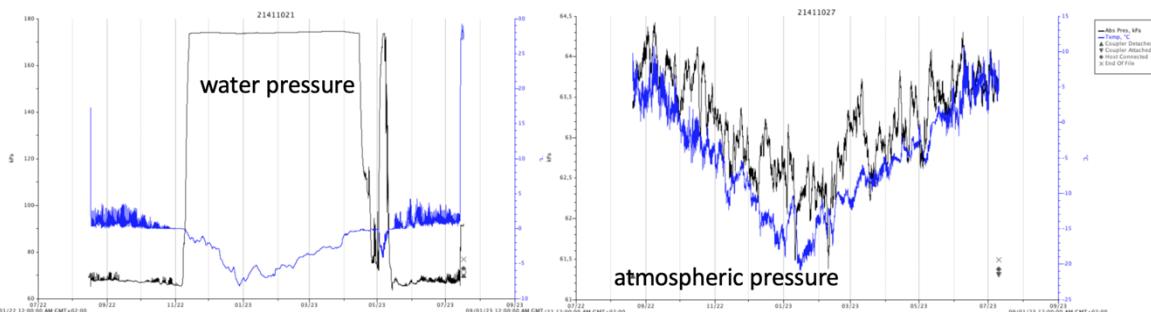


Figure 3: Example of water pressure and atmospheric pressure data of 2022-2023 at Batysh Sook (Suyek) before cleaning the data.

As can be seen in the figure above, the water pressure data can look strange in winter. This is because the water around the logger freezes and increases the pressure significantly. These data are useless for the calculation of discharge and have to be deleted for our

purposes. This is most easily done in a spreadsheet software like Excel. But be aware that Excel might randomly change the time format. **Always use a copy of the original data file and never work in your original.** A good way to clean the data is by only deleting the values for water level but to keep the other data (temperature and date). This allows later on to take the original dates and reassemble the time series with these dates if Excel broke the time format.

The actual correction consists of simply subtracting the atmospheric pressure from the water level logger. The result is only the water level. This is the final data from this step.

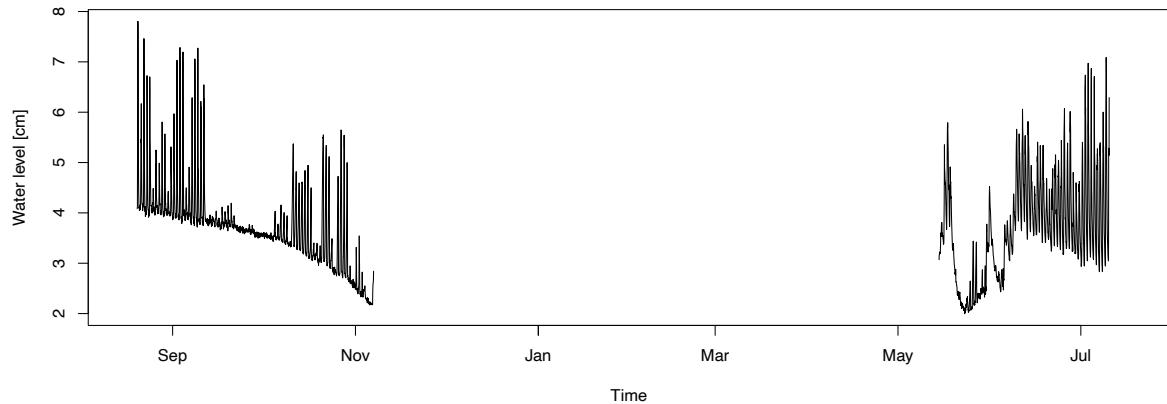


Figure 4: Water level after cleaning (frozen period) and removing the atmospheric pressure signal. This process is done in Excel (deleting the data where the water is frozen) and RStudio (merging the atmospheric and water pressure datasets, and subtracting the atmospheric pressure – script: 002_pressure-correction.R).

Data preparation (FL30)

The FL30 stores data in a specific binary format that is unreadable to us. This format needs to be translated first into a readable text format. All measurements are stored continuously in this file. Therefore, in a second step, the data can be split into multiple parts, which makes finding the individual concentration curves easier. The easiest way in my opinion is to use a spreadsheet program (e.g. Excel) to cut these measurements into individual parts. Each measurement can so be stored as a separate spreadsheet. The best for the spreadsheet names is a combination of the date and time of the measurement. After all measurements are organized in individual spreadsheets, the data is ready for processing in RStudio.

```

A 09:07:23 05:45:08 A 001 33386 001 40662 192 52961 018 38723
A 09:07:23 05:45:10 A 001 33614 001 42587 192 44230 018 38240
A 09:07:23 05:45:12 A 001 42139 001 42443 192 34674 018 37928
A 09:07:23 05:45:14 A 001 36841 001 46385 192 29070 018 38046
A 09:07:23 05:45:16 A 001 37310 001 46266 192 25523 018 38256
A 09:07:23 05:45:18 A 001 37242 001 45669 192 26722 018 38100
date Tracer1 Battery Baseline Temperature
2023-07-09 05:45:04 16.3 11.42 14.74 18.17
2023-07-09 05:45:06 16.01 11.43 14.68 18.16
2023-07-09 05:45:08 15.82 11.42 14.74 18.16
2023-07-09 05:45:10 16.11 11.41 14.77 18.15
2023-07-09 05:45:12 16.09 11.4 16.04 18.14
2023-07-09 05:45:14 16.68 11.4 15.26 18.15
2023-07-09 05:45:16 16.66 11.39 15.33 18.15
2023-07-09 05:45:18 16.66 11.39 15.33 18.15
2023-07-09 05:45:20 16.66 11.39 15.33 18.15

```

Figure 5: Fluorometer data in a normal text editor before (top) and after (bottom) translation from binary to readable data format. The tracer (color) data is now in mV.

Wrong time and date (FL30 clock) Corrected time (check notebook)

	A	B	C	D	E	F	G	H
1	date	time	time_adj	Tracer1	Battery	Baseline	Temperature	
23	09.07.23	05:45:46	15:05:46	16.24	11.35	14.83	18.29	
24	09.07.23	05:45:48	15:05:48	16.21	11.35	14.74	18.31	
25	09.07.23	05:45:50	15:05:50	16.3	11.35	14.7	18.33	
26	09.07.23	05:45:52	15:05:52	16.23	11.34	14.74	18.34	
27	09.07.23	05:45:54	15:05:54	16.08	11.34	14.68	18.36	
28	09.07.23	05:45:56	15:05:56	16.17	11.34	14.76	18.38	

All data

Data of experiment @ 14:54 2023-07-12

Figure 6: Overview of the same FL30 data as in the figure before but now in Excel. The FL30 clock is not set (old logger with battery problems). Check the field notebook to identify the measurements and the correct local time. The colored values are conditionally formatted cells. This allows finding changes in the concentration with the eye. Every experiment is copy pasted into a new spreadsheet (bottom) with a name corresponding to the date and time of the experiment. The file needs to be saved as **.xlsx** file in the end.

Fluorometer discharge calculation (RStudio)

Scripts and arguments

As outlined before, there are 5 steps in total to be performed in RStudio. This part explains the individual steps in a bit more detail. For the practical application, all **scripts** are provided in the download material. The arguments, like how much tracer fluid we injected, or the spreadsheet names need to be changed by us. The **scripts** have the file ending “**.R**” and should be automatically opened with **RStudio**. If not, one can right-click and select “Open with ...” and choose RStudio. The scripts will automatically open inside RStudio. At the beginning of each script are arguments that need to be adjusted by you.

Every script for the calculation of discharge from the FL30 requires some arguments. These are:

- 1) Name of the (Excel) file (including path/folder)
- 2) Spreadsheet name
- 3) Concentration of the original color solution in **mg/l**
- 4) Amount of tracer injected into the river in **liters**

(1) The file should be placed in a folder on your computer. This filename and file path need to be defined:

“C:/discharge/fluorometer_discharge_arabel.xlsx” (Windows)
“/Users/username/fluorometer_discharge_arabel.xlsx” (Linux/OSX)

(2)

The spreadsheet name must be the one matching the spreadsheet name in your Excel file: “230828_0945” (example: 2023-08-28 9:45h)

(3) The prepared color solutions result normally from solving a specific amount of tracer in 1 liter of water. For most cases, 50 g/liter is the concentration. Only at very high discharge

(e.g. Abramov Glacier), more tracer is used. A solution of 1 liter with a concentration of 50 g/l can last for 10 to 20 measurements depending on the discharge. The initial base solution must be put into the argument **Standard**. The value needs to be in g/liter.

- (4) The amount of tracer used for the experiment is the argument **TracerVolume**. This value needs to be in liter.

Processing

The rest of the script will go through the individual processing steps. A short description what the step does is given in the following subsections. The overview of all processing steps is:

- 1) Background noise removal
- 2) Calibration of the signal from mV to $\mu\text{g/l}$ (=ppb)
- 3) Selection of the “breakthrough curve”
- 4) Calculation of discharge
- 5) Calculating the rating curve

Background noise removal

The signal in mV shows a background signal, where we know that the concentration is 0 but the mV signal is higher than 0 (e.g. 30 mV). When we calibrate the mV signal with our calibration fluid, we must remove this false signal. To do that, the script will plot the whole experiment for us, and we can **click into the plot to define 2 points (start and end) of a period**, where we know there is **no color** in the water. After clicking two times (points will show up in the plot), the next lines will calculate the average signal strength of this period. This value is subtracted automatically from the original signal.

Calibration of the signal from mV to $\mu\text{g/l}$ (=ppb)

A bucket with 5 liters of stream water was used to calibrate the fluorometer. Into the 5 liter of river water, we added vials of 50 ml with a color concentration of 250 $\mu\text{g/l}$. This results in an approximate concentration in the bucket of $250/5 = 50 \mu\text{g/l}$. Because we also add 50 ml of water the exact concentration is 49.5 $\mu\text{g/l}$ or expressed as parts per billion (**ppb**) 49.5 ppb. When adding a second vial to the bucket the concentration is 98.04 ppb. The script allows us to identify these concentration points (up to three points). In the plot window, we can select the easily identifiable calibration points in the same way as for the selection of the background noise by clicking 2 times into the plot. The script will calculate the average mV signal of the calibration phase for one, two, or three points. After that, a linear regression determines how the mV signal corresponds to the calibrated concentrations. A plot will show this relationship. The plot can help identifying problems in the calibration: Are the points lying well on the line (good) or do they deviate (bad)?

Not every measurement has calibration points. In case one of the measurements does not have calibration points, we can use the calibration from the previous measurement. If you calculate all discharge experiments in one session, then the last identified relationship is still stored in R. If you have not done a calibration in R, then you have to open one of the measurements with a calibration and run the script to this point. Choose at best a

calibration where the water quality (especially sediment concentration) was similar to the measurement.

Finally, the mV signal is transformed into a ppb signal using the determined relationship. This ppb signal is used in the next step to calculate area of the breakthrough curve.

Selection of the “breakthrough curve”

The breakthrough curve has a characteristic shape with a sharp increase and a slow recession of the signal. In the same way we selected the background noise, and the calibration points, we now have to select the start and end point of the breakthrough curve. Once this is done, the script will already calculate the discharge for us. This is done by dividing the amount of tracer by the area under the curve (AUC) of the breakthrough curve. The amount of tracer is calculated by multiplying the amount of tracer with the concentration of the Standard.

We have to save the result of the discharge measurement now. This can be done in a new spreadsheet file (new file). It will be useful to note down the date, time, and calculated discharge. We have to do this for every measurement. The more points we have, the better.

Calculating the rating curve

After we have all discharge measurements, we can finally go into the calculation of the rating curve. The relationship between water level and discharge depends amongst other things on the river gradient and on the river cross section shape. This is why a rating curve is specific for a certain point. This point is where our water level is measured. The relationship is also not linear. This is why we will fit a polynomial curve to the measured points. A separate R script is available for this. After deriving this curve, we can finally transform the water level values into discharge values.