New data analysis
La Corona NCSU

Updated	By
10/25/2023	Agustin and Eliana
23/12/2023 Proofread	Eliana

Members
MSc. Eliana Nervi
MSc. Agustín Menta
PhD. George Chescheir
PhD. Francois Birgand

Contents

1 Study area 4
2 Rainfall data 6
2.1 General information 6
2.2 Input files and format 6
2.3 Data read and conversion 7
2.4 Quick Check: 11
2.5 Diagnostic plots (6 months of data) 12
2.6 Storage. Final Plots and spreadsheets 12
3 Weather data 14
3.1 Input files and format 14
3.2 Data read and conversion 14
3.3 Quick Check 14
3.4 Diagnostic plots 15
4 Wells data 16
4.1 Input files and format 16
4.2 Quick Check 16
4.3 Storage. Final Plots and spreadsheets: 17
5 Flumes data 18
5.1 Input files and format 18
5.2 Data read and conversion 20
5.3 Quick Check: 20
5.4 Diagnostic plots 21
6 Water Balance 22
7 Actual progress analysis 22

This document describes the new workflow proposed of field data analysis from La Corona during the visit to NCSU, having done interviews with Chip and Francois.
In Figure 1 is showed the catchment areas and equipment installed.
In Figure 2 is showed the new workflow of the data process. The new workflow process includes a Raw Data, Process and Final Data folders. Each folder is divided by each variable process. The Final Data folder also has a Historic data folder where the historic daily txt files are saved and update after every process. The Raw Data folder is only to save the folder of raw data that is send from field data. The process folder has the main objective of run the scripts of each variables to process and correct the information. In this folder the unique stable documents are the scripts but the input data and processed data are temporal and after the process has to be deleted or paste in the Final Data folder.

1 Study area

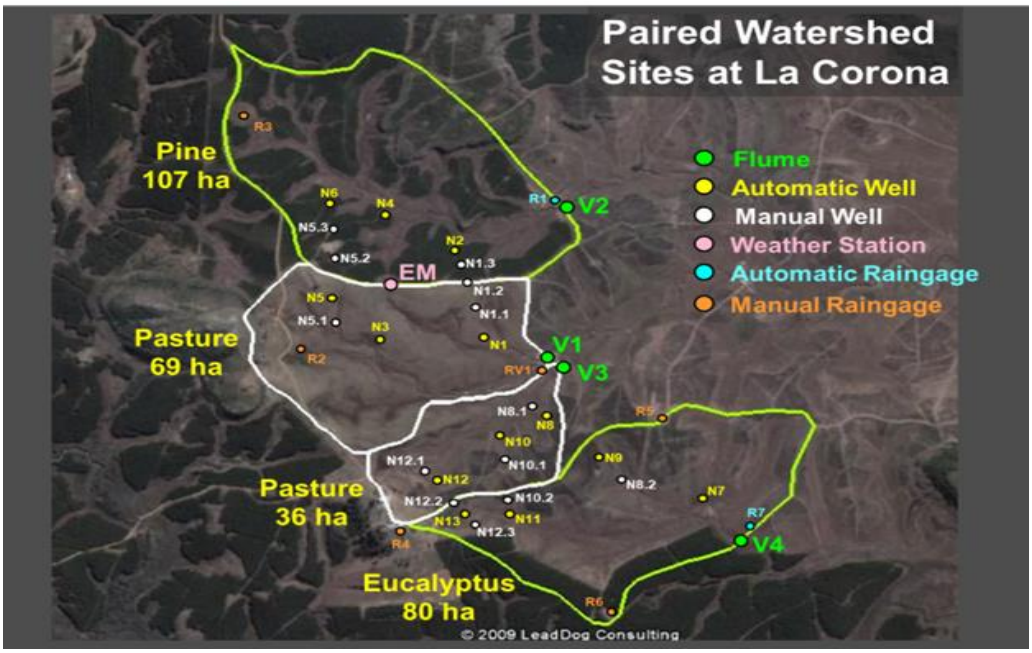


Figure 1: Catchment and equipment of La Corona

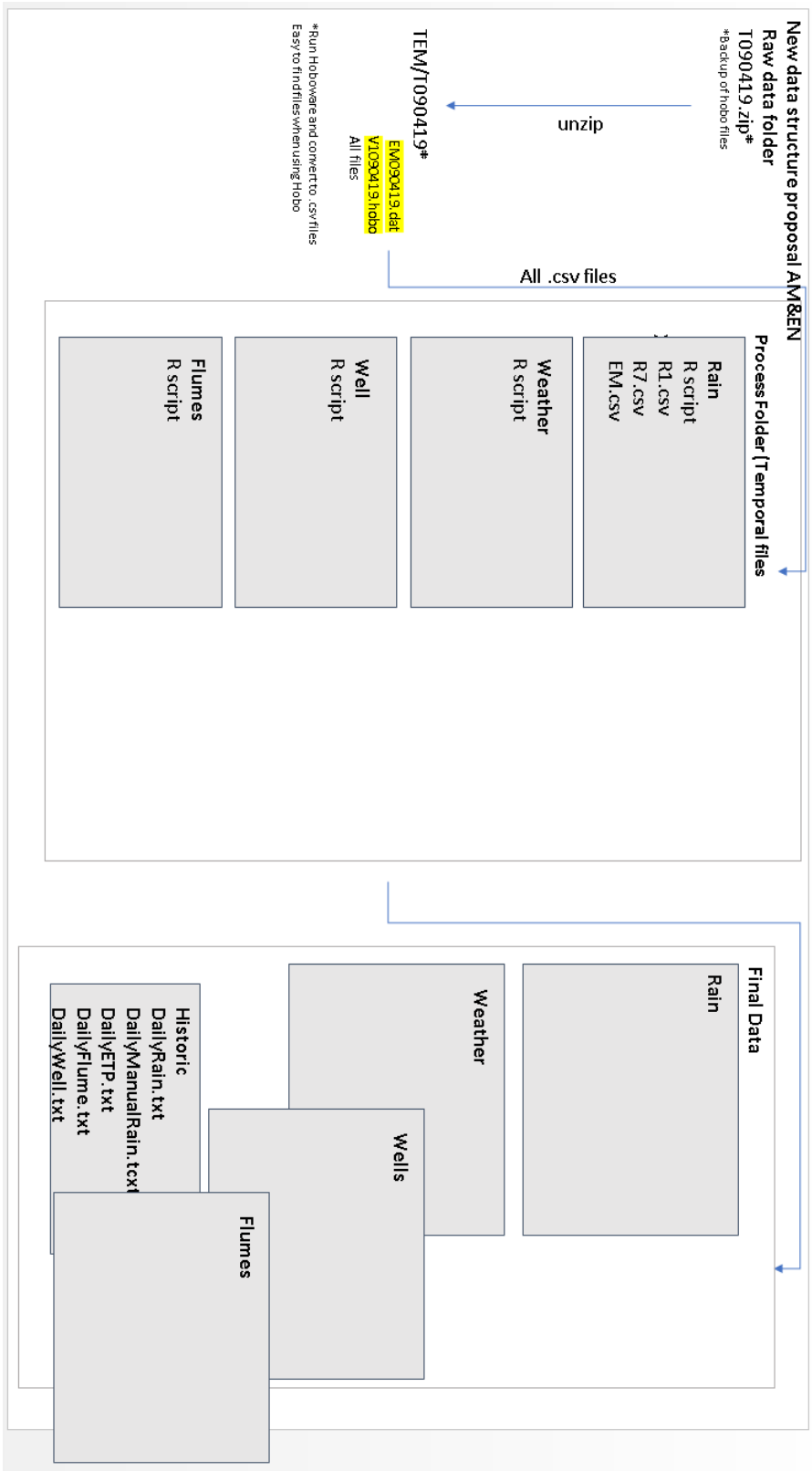


Figure 2: Proposed Workflow of data process of three main folders: Raw data, Process folder (temporal files, quick check plots and outputs) and Final data to storage the final processed files and historical data series.

All the information included in this document is in the folder where is included the raw data and processed data of each variable. It is also included an example document of Juliana field data survey.

All the scripts are in the github https://github.com/enervifa/lacorona_ncsu

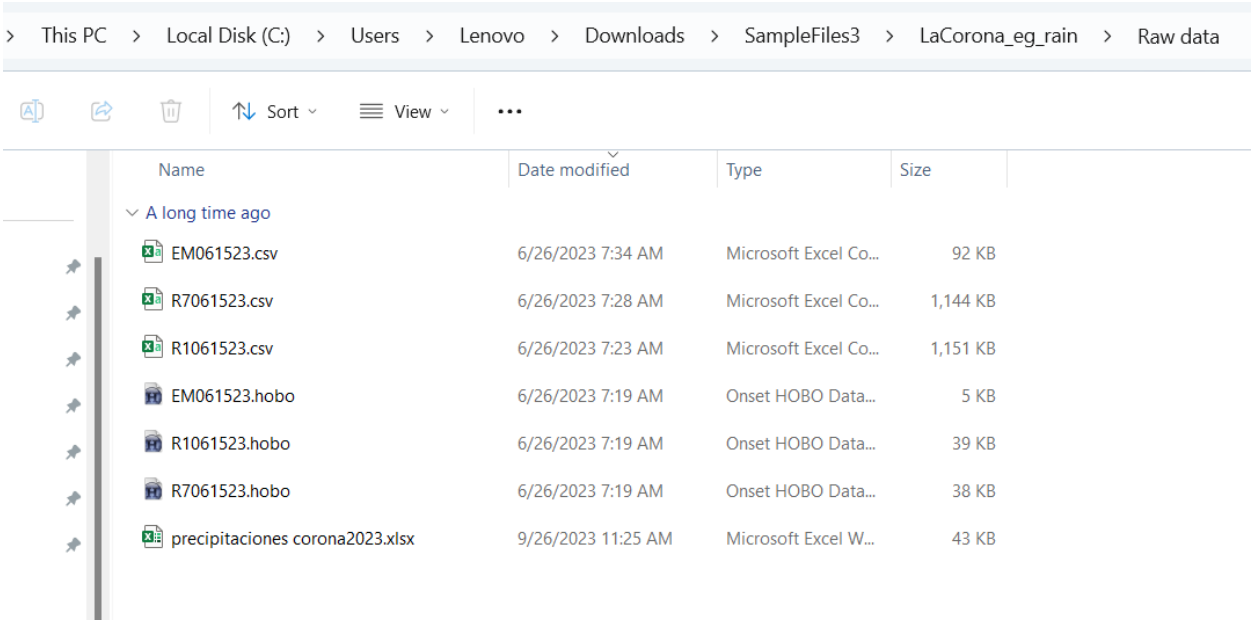
2 Rainfall data

2.1 General information

The first dataset to analyse is rain data. Automatic data is analysed first in a plot using Hoboware, and manual data is used only for checking.
Now, the EM has the most reliable rainfall data recorded every tip as a brake point logger. This logger changed in Sept 2020. Before this time, the main automatic pluviometers were R1 (Pine, catch 1) and R7 (Euca, catch 4) because the rainfall in EM was recorded by the WS every 15 min. Originally the automatic data from R7 was labelled as R4 (because it is located in catchment 4). Automatic data from R7 has been labelled as R7 since March 2023. If everything works well now, Chip will use the EM data. If not, separate datasets for catch 1&2 and catch 3&4 are used.
In this analysis rain data is not corrected by wind.
For data filling, it is used 1:1 relationship between the raingauges. Every tip in R1 and R7 it counts as 0.254 mm and in the EM station it counts as 0.1 mm. Manual rain is recorded when the people go to field and R4 (casco) is recorded more frequently with the rain event because people are living there.

2.2 Input files and format

Automatic rainfall input files, R4 (named R7 in Figure 1), R1 and EM input files, from here this is EM (only rainfall EM data). Manual rainfall is precipitaciones La Corona 2018A.xlsx



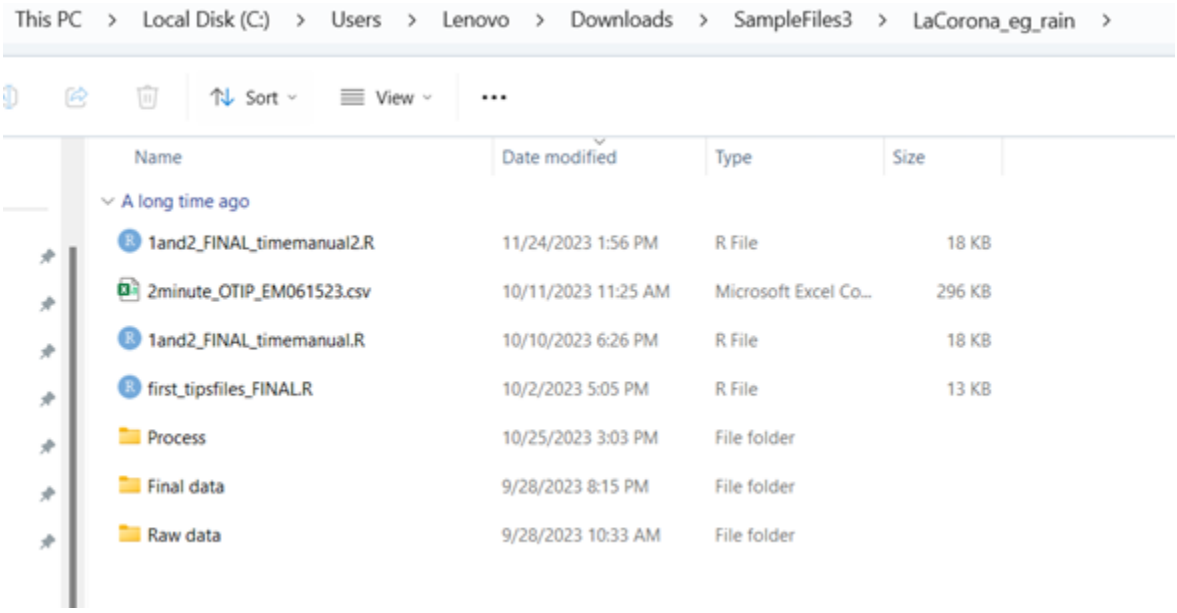
*There is one spreadsheet per year and monthly data is added in different sheets. There is a last sheet called total that adds manual rainfall monthly per station.

An example of these input files for rainfall analysis can be found in the folder

2.3 Data read and conversion

The following section describes the step-by-step with comments and useful details of the process:

1. Open R1, R7 and EM in Hoboware. Save the data in .csv format.
2. Fill out the form Rain data form *https://forms.gle/azt6nADxmnSfJRR37*
3. The folder is structured in three main folders:
    i. Raw data – to storage raw data input files (.csv files from HOBO)
    ii. Porcess - This is a live folder, dynamic one where the data is moved to be processed. This folder store the scripts.
    iii. Final data – This is the folder were corrected final data is stored after processing.



4. The script 1_quickcheck_rain.R takes these three files (is necessary to copy the name files into the script), convert to individual tips (write OTIP_files) and does a quick check using tips in mm. eg output OTIP_EM061523_P.csv
    a) If no modification is required, saved each file as in OTIP_filename_P, and run 2_cumrain.R and generate ouputs at different time scales saved in folder
    - 2minute_outputtip_EM061523
    - hourly_outputtip_EM061523
    - daily_outputtip_EM061523

b) If data is not OK, and changes are required, change manually the OTIP_file. This modification should be saved in OTIP_filename_P (processed TIP file).

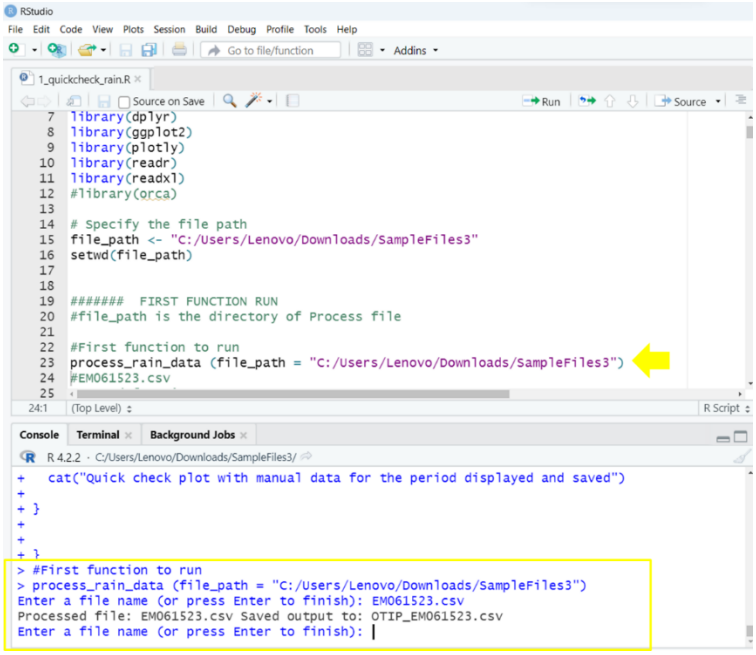


Figure 3 Screenshot of script and first function process_rain_data

5. If changes were made, run again 1_quickcheck_rain.R to check everything is ok. End the quick check and data clean process with OTIP_filename_P files. This function has an improved function called **1and2_final_timemanual.R** **(see next page)**

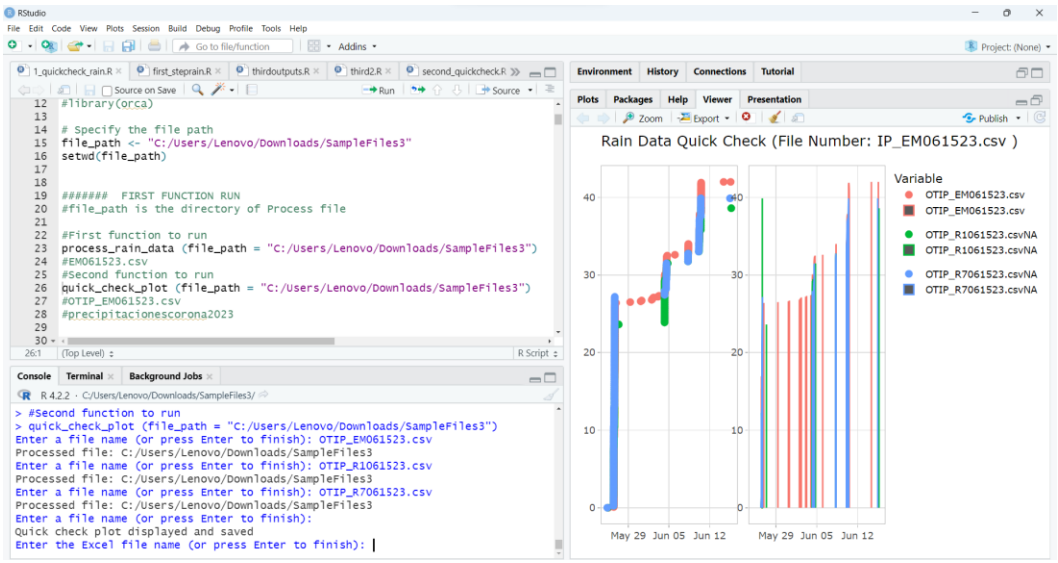


Figure 4: Screenshot of script and quick check plot

6. To add manual data, read the excel file in the Process/Rain folder by input the filename ### This could be improved in the filed by using a Google Forms instead of an excel file. Some excel files from previous years have changes un row numbers or titles, which needs to be checked before running this script.

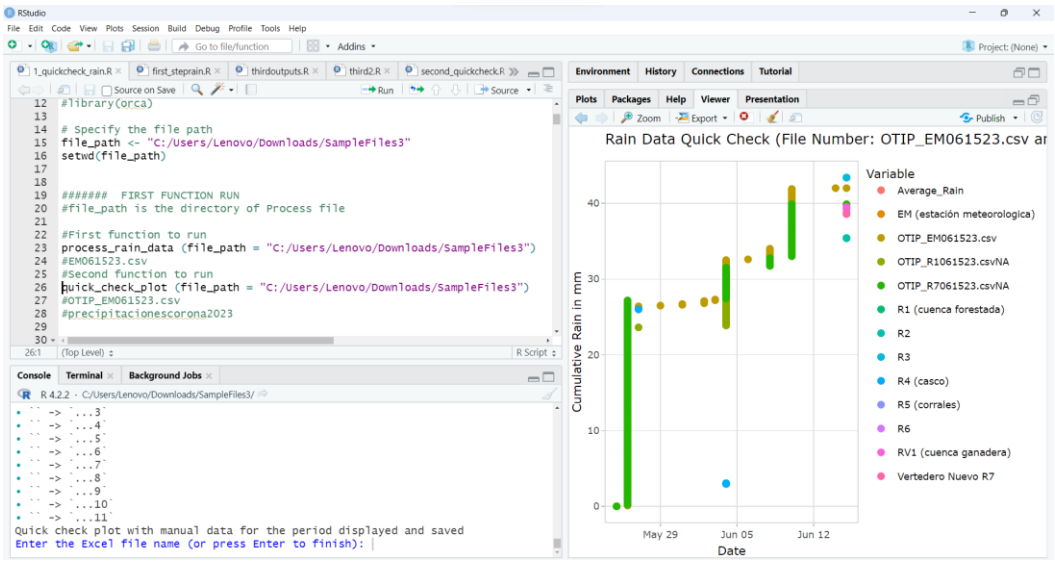


Figure 5: Screenshot of script and quick check plot with manual data

In this example, data is joined by date only, and not date and time. The script calculates an average of the data when is more than one observation for the pluviometers, wich happens often when there is a field visit. As R4 is recorded more often, the average is not considering that data from R4.

This function has an improved function called **1and2_final_timemanual.R** in which manual data is assigned a daily time of observation arbitrary at 7:00 am.

This should be adjusted by reading the time of the field visit in planilla cuenca.

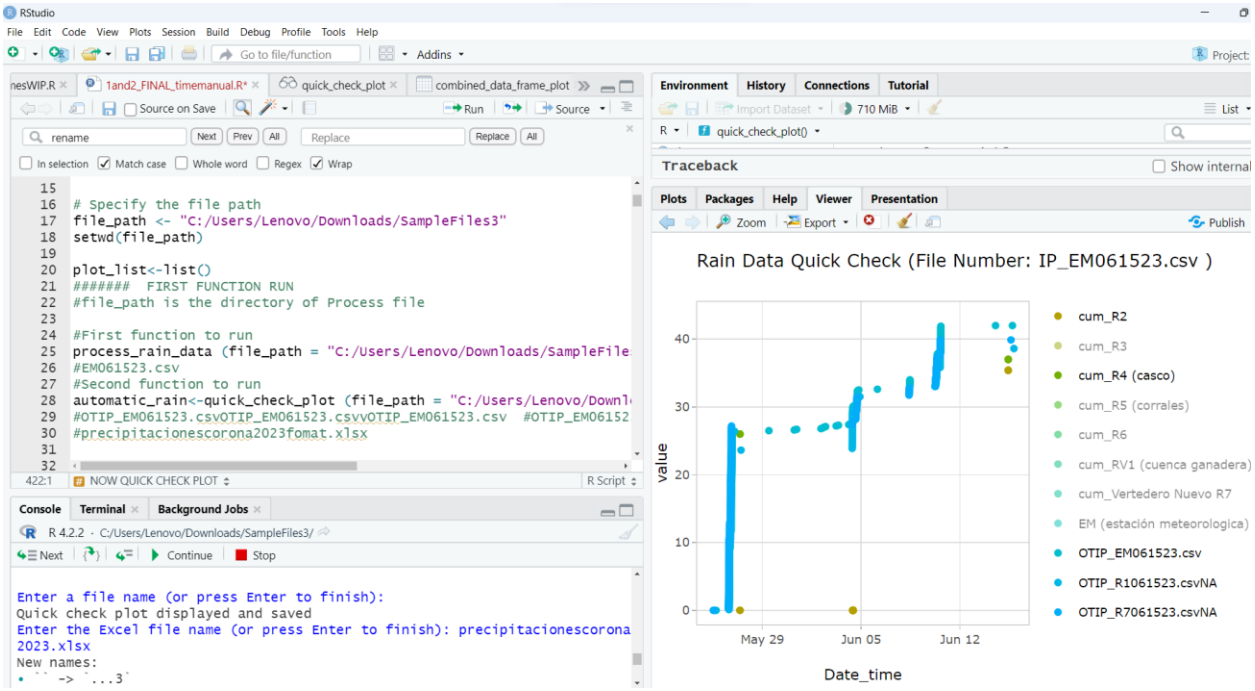
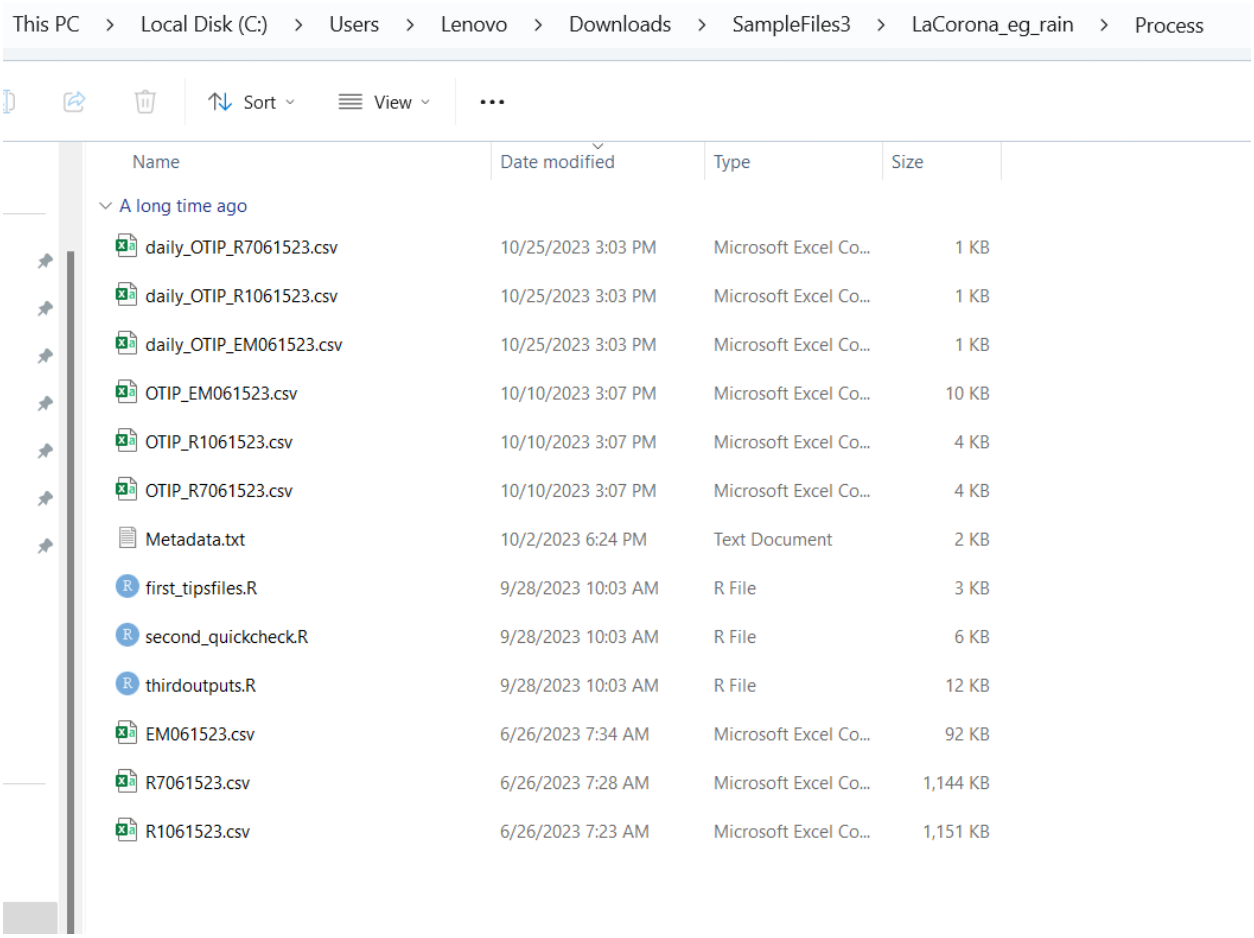
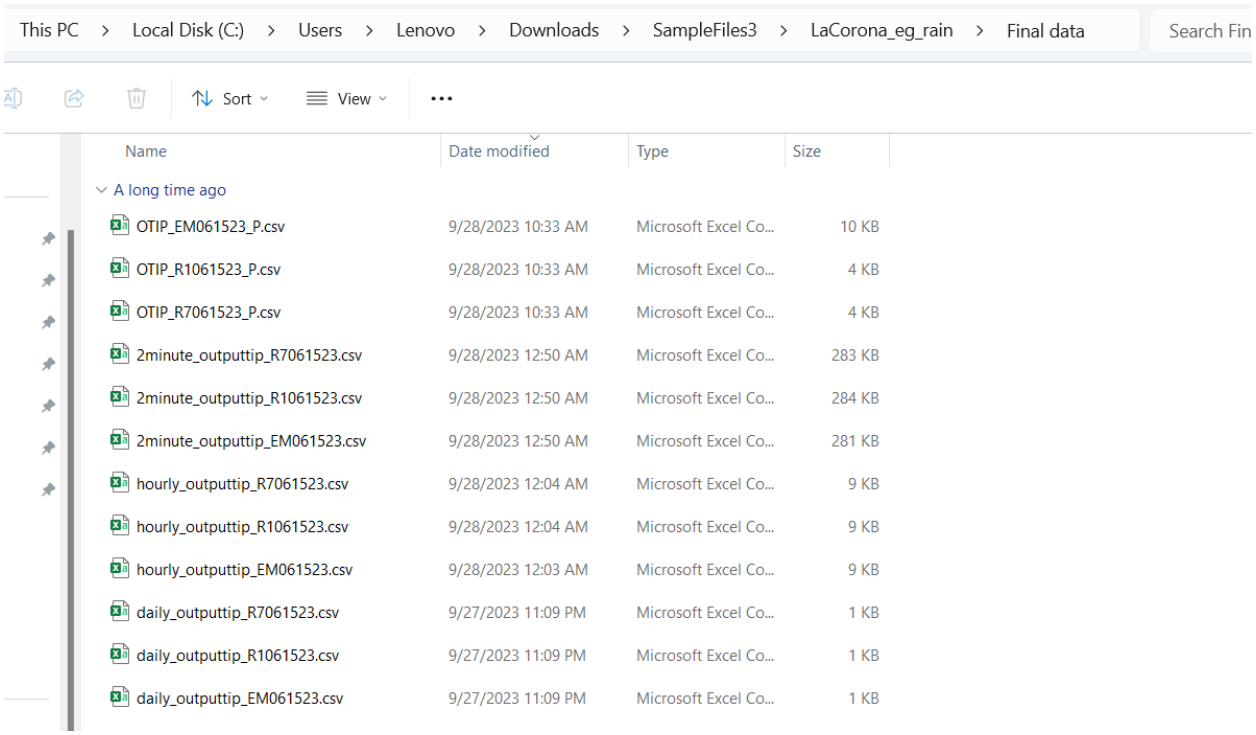


Figure 6: Screenshot of script and quick check plot with manual data defined at 7 AM

7. New files with tags 2minute_filename, hourly_filename, daiily_filename will be generatedas shown in point (3.a)
8. Run 2_cumrain.R in folder script to append the processed data series to the historical data



9. Move all generated files to Final Data folder

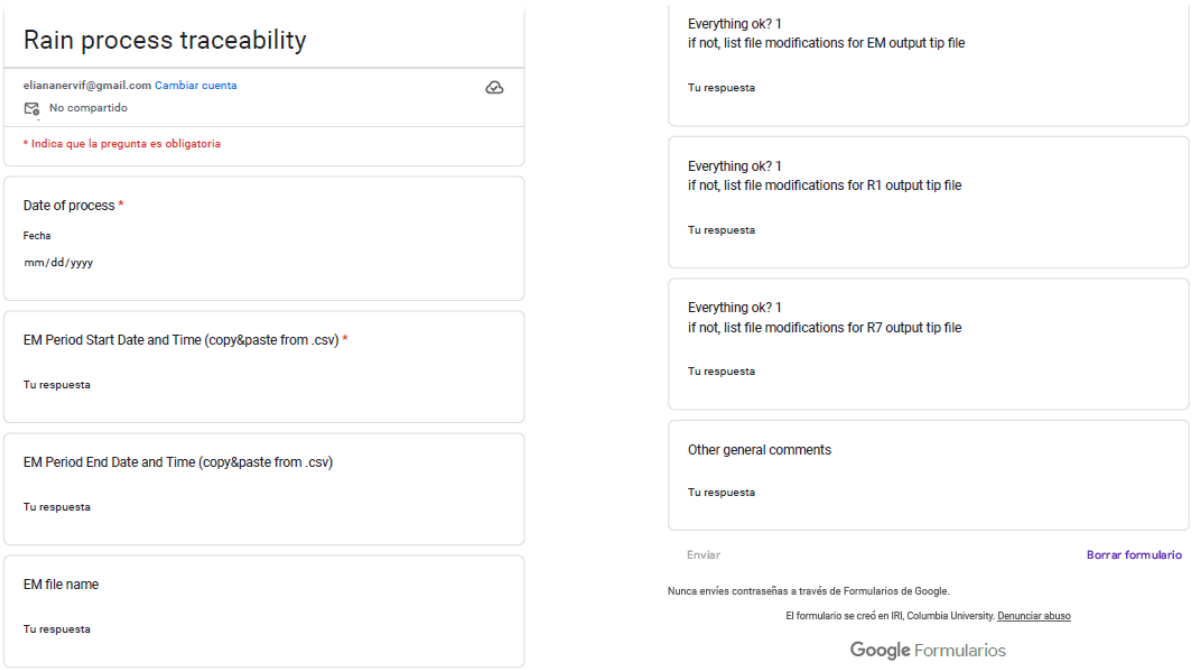


10. See data process traceability Form in *https://docs.google.com/spreadsheets/d/1k3zPik-nTle2Y716iBJjGiykP2R_5pgLYHsKD3pinCc/edit#gid=2137517585.*

2.4 Quick Check:

This step allows to compare the three rain time series and notify problems to people in the field in the short term.

In this case, the pluviometers data looks OK. However, the EM has tip counts in a period when R1 and R4 have no read. It should be a good strategy to look at the other variables of the weather station (Solar radiation, RH) before delete tips. This changes results in a change of 1 mm in the total rain registered by the EM. Fill the form *https://forms.gle/azt6nADxmnSfJRR37* to ensures the traceability of the analysis process. This will storage the metadata. In the future, a next step will include this automatically registered trough R.



Add manual data as the average of manual measurements from every dataset in precipitaciones corona. R4 pluviometer gets measured more frequently. This is important when putting all the data stacked by year.

	A	B	C	D	E	F	G	H	I	J	K
1	Fecha	R1 (cuenca forestada)	R2	R3	R4 (casco)	RV1 (cuenca ganadera)	R5 (corrales)	R6	Vertedero Nuevo R7	EM (estación meteorologica)	Observaciones
2	enero	53,4	49,8	55,2	113	59,6	61	67,2	64,2	51,8	
3	febrero	70,8	14	166,8	89	160,6	148,3	151	155,4	155,8	
4	marzo	173,6	153,4	190,3	158	162,8	164,6	192,4	189,2	153,2	
5	abril	58,6	54,2	59,6	52	52	54,8	53,2	51,8	55,6	
6	mayo	71	75,8	84,8	47	78,2	73,8	76,2	74,6	60,6	
7	junio	38,7	35,4	43,4	0	39,5	39,4	39,4	38,6	39,8	
8	julio	191,8	162,4	235,8	125	207	164,8	196,2	179,8	171,5	
9	agosto	78,6	72,4	85,2	53	78,5	72,8	73,8	74,2	77,8	
10	septiembre	143,8	144,2	147,2	140	142,4	140,6	142,2	0	140,4	
11	octubre										
12	noviembre										
13	diciembre										
14											
15	total	880,3	761,6	1068,3	777	980,6	920,1	991,6	827,8	906,5	900,9125
16											

Figure 7: Spreadsheet of precipitaciones corona

2.5 Diagnostic plots (6 months of data)

This needs to be coded

2.6 Storage. Final Plots and spreadsheets

- Daily rainfall of every rain gauge.

- Txt file with processed data and cumulative rainfall every 6 months with the manual data.
- ScatterPlots with cumulative and simply data at daily scale and correlations between the raingauges.
- Rainfall data and flume data every 2 minutes in txt file (.prn) are in the folder

3 Weather data

The second dataset is weather data. It includes in the EM: Temperature, Relative Humidity, Wind velocity, Solar radiation, and Net Solar Radiation. The backup station is the WS, it records Temperature and Relative Humidity. Both recordings are in a 15-minute timestep.

3.1 Input files and format

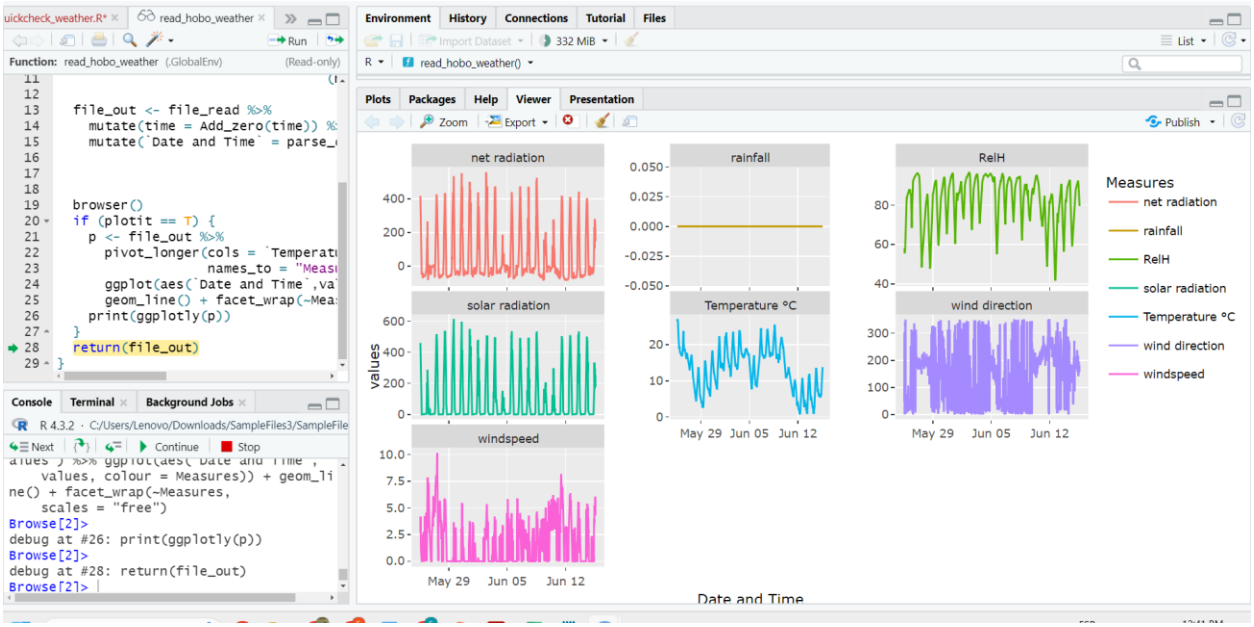
The input files are EM042223.dat (.txt comma-separated file) and WS091123.csv (from .hobo file)

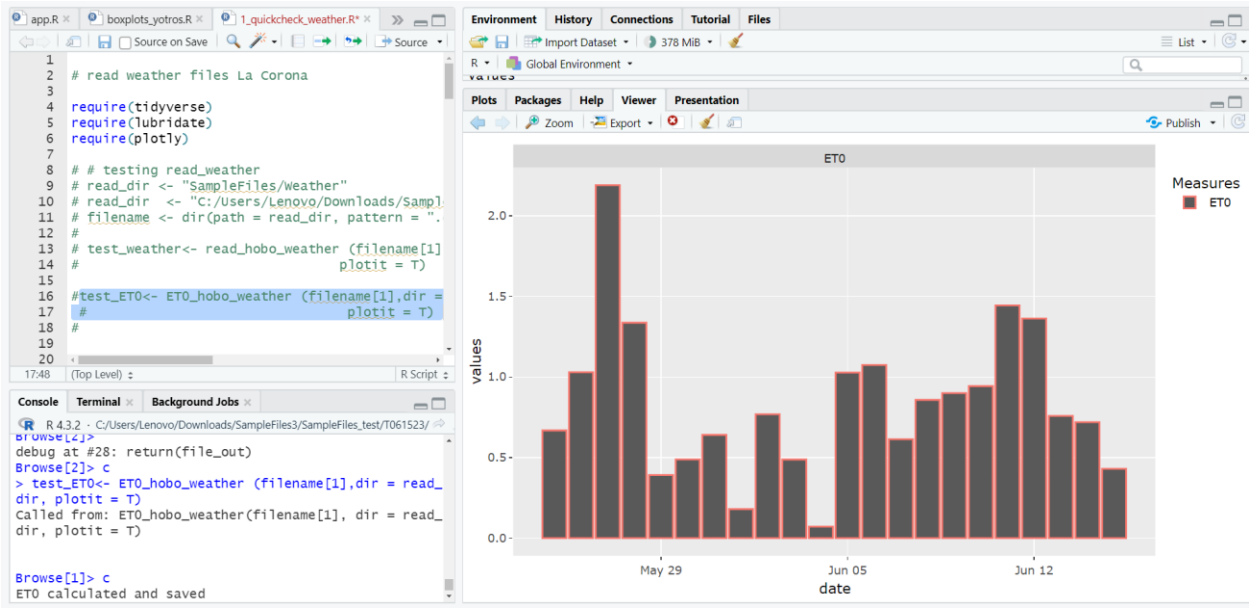
3.2 Data read and conversion

The objective of this step is to accumulate all the variables daily and calculate daily reference evapotranspiration using Penman-Monteith equation (ETo). The quickcheck of each variable, accumulation in a daily stcale and the reference ETo calculation is done in the 1_quickcheck_weather.R script
The day is considered from the data of 00:00 to 23:45 of each day.

3.3 Quick Check

1. The quick check includes two main plots:
   Plot of each variable for the period analysed and Plot EM vs WS Temp and RH





2. If something is not correct with the weather station EM
   a) The first option is to use the WS for data filling of temperature and RH.
   b) The second (last) option is to download La Magnolia daily time series for data filling considering a 1:1 correlation

3.4 Diagnostic plots

Plot ET from the EM daily in mm, verify the values are in a reasonable range, and add to a longer dataset in the folder.

Plot EM daily in mm vs WS Temp and RH is missing
##all plots are saved in html format

4 Wells data

There are 13 wells with continuous monitoring in the four catchments. The manual wells are not used for the analysis. There is no document of the construct and litologic profile of the wells.
The time step of each measure is every 1 hour and the barometer every 2 minutes because it is also used in the flumes.
There is a wellmetadata.xlsx spreadsheet where it has the information of all the wells in mm. Height is wellhead (mm). WTD is the measure of water table depth from ground (mm). Length is the depth of the well from the wellhead and Max depth is the depth of the well (mm) from ground (see Figure 8)

Well ID	Meas	Height	WTD	Length	Max Dep	
N1	Dry	547	1483	2030	1483	11:01
N1.1	Dry	400	1220	1620	1220	11:00
N1.2	Dry	1190	1590	2780	1590	10:58

Figure 8: Example of spreadsheet WellMetaData

Every Manual data measured in the automatic fields should be saved in the ManualMeasurements.xlsx spreadsheets to have a record of the historical manual data and this spreadsheet is going to be used by the R script.

4.1 Input files and format

4.2 Quick Check

To analyse the data of the wells of the catchments 1&2 paste all the csv files in the process folder and run the 1_quickcheckwells.R (in process) to plot all the wells from each pair of catchments every 1 hour and convert to daily data selecting the well depth table of 11.00 p.m
Temperature is not used in evaluation.
The manual measurement from wellhead needs to be input by the user or read automatically read from the spreadsheet ManualMeasurements.xlsx and is used in the script to plot with the continuous monitoring to check that all the data is ok.
read files and compensation needs to be coded

To process the data without using the HOBOWARE software is:

1. Read the data in the HOBOWARE and export as csv (without compensation) or read directly from R the HOBO file of the well and the barometric data
2. For compensation in R is necessary to read the absolute pressure in kPa of the well and the barometer.
    a. To calculate the water level depth from the ground is necessary to substract the last barometric pressure to the absolute pressure of the well sensor.

b. After that is necessary to change to water meters level (*100/9800) (Hi). It could be input the density of the water depending of the water temperature. But this not generate too many changes in the results.
c. WTD=MDM-(H_M –Hi). Where MDM is the manual water level depth from the ground and H_M is the difference as explained before but for the moment of the manual measurement.
d. Important: If the barometer stops before the end of the well sensor it is used the last barometric measure for the rest of the measures

4.3 Storage. Final Plots and spreadsheets:

The water table variations are used to select the hydrologic balance period between no variations. To analyse data, it is used an average between the two downstream piezometer (N1-N3 and N2 and N4) and Plot with the rain.

5 Flumes data

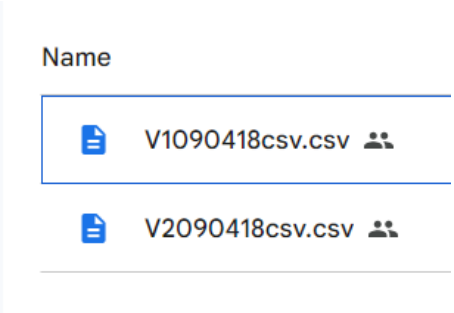
There are different loggers in each PWE. For example, there are 3 ways of measuring the flume in each catchment, and 4 in V1 and V2 because of the emergency spillway. Particularly, the HOBO U20 was installed upstream because the measurements in the HOBO U12 were often incorrect due to the accumulation of sediments in the logger pipe.
Also, the emergency spillway was built in V1 and V2 because of structural problems in the flumes that cause problems, especially after heavy rainfall.

The principal ways of estimation of discharge is firstly the Sampler (ISCO), secondly the HOBO U12, and thirdly the HOBO U20, as this latter is not able to measure low flows. The HOBO U12 is connected to the Stevens (is often good in low flows, but it does not drift (overestimate flow)).
In Eucalyptus catchments the ISCO measure stage and velocity and in Pines catchments the stage is only recorded. Velocity is only to check.

5.1 Input files and format

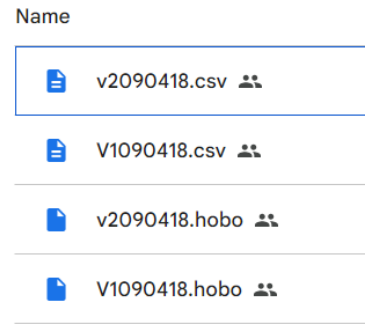
The inputs files are in the folder

Sampler ISCO
E.g V2091123.csv



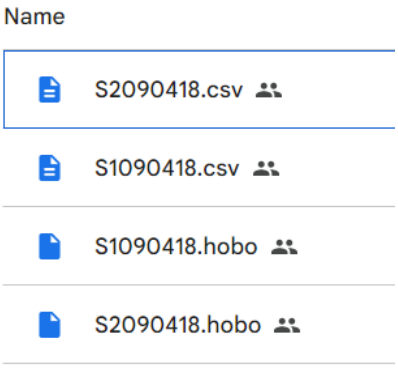
Stevens (HOBO U12)

E.g V1091123.hobo



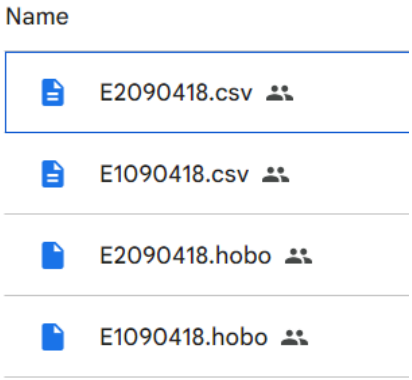
HOBO U20 OutsideWell

E.g S1091123.hobo



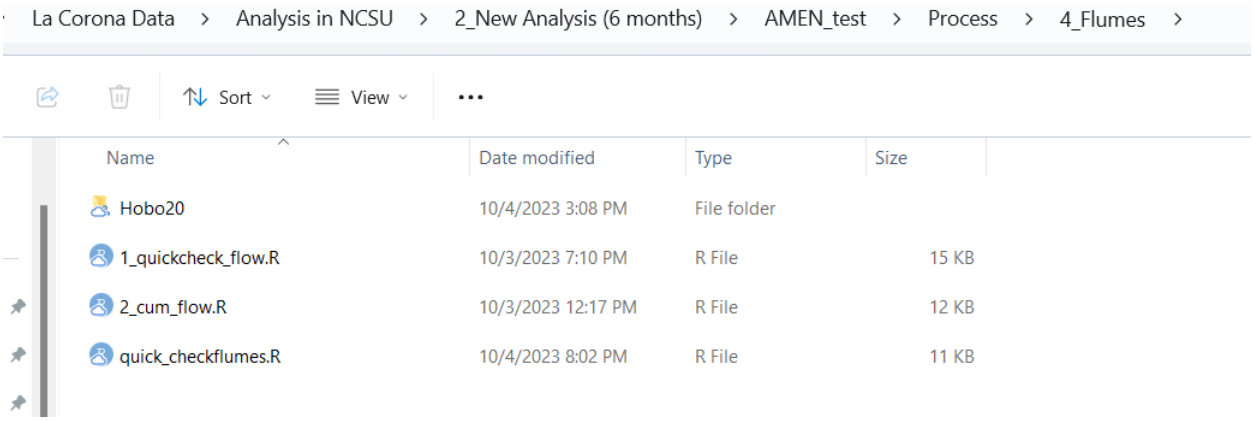
HOBO U20 Emergency Spillway

E.g E2091123.csv



All inputs files are in the same folder

5.2 Data read and conversión

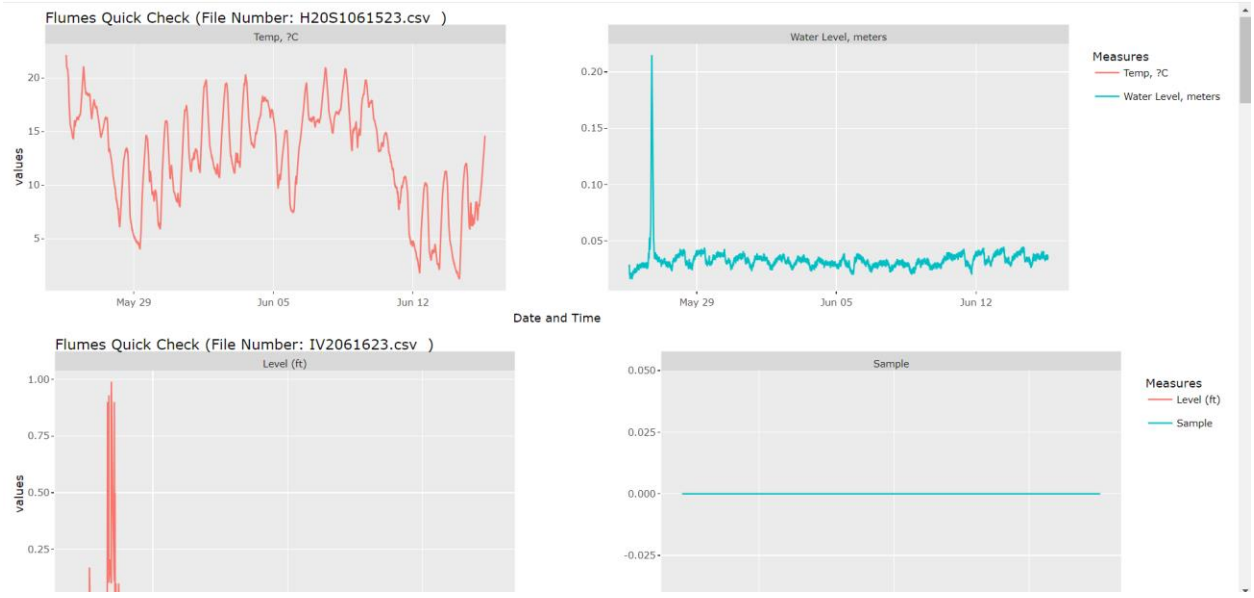


1. Convert the files to .csv using the Batch conversion function in HOBOware.
2. The data is run by 1_quickcheck_flow. R. This script plot all the flumes and the manual measurement from each catchment in meters to correct the measurements
    a. To process the data from U12, use a linear transformation to convert volts in water height in meters.
    b. The sampler needs to be converted from feet to meters.
3. After preprocess and select the most reliable measurement run the 2_cum_flow. R script. This script calculates the discharge in mm with the flume equation every 2 minutes. In this script is necessary to add the emergency spillway data, only when stage is above 0. This script also calculates the daily data.
4. The daily, 2 minutes and hourly results are saved in the folder

5.3 Quick Check:

Plot water stage in meters in 2 min timestep of three discharge loggers and select the most reliable measurement (often reliability order: Sampler, U12, U20). Flume_plotys contains the plots

of all flumes processed.



Combined_plots also plots rain to do the quick check.

5.4 Diagnostic plots

Plot rainfall and flow 2 minutes data to calculate the time response of the catchment (is very stable and the forestry catchment takes more time. Here we can see that the pasture catchment has higher peak flows.

Compare two watersheds monthly and yearly based. ## read files and code this

6 Water Balance

The water balance is done with the daily data. The period is selected with a similar water table level between almost 6 months.
Equation:

ETR=P-Q

ETR=Transpiration+Evaporation (Interception)

this step needs to be coded

7 Actual progress analysis

The Pine watersheds are processed until today but the eucalyptus catchments are only 3 or 4 years of processed.
All the raw data is less than 55 GB.

Scripts and files to read

Data/Instrument	filenames	Scripts to process
Rain/Automatic rain gauges (R1,R7,EM)	R1061523.csv, R7061523.csv,EM061523.csv	1_quickcheck_rain.R 2_cumrain.R
Rain/Automatic rain gauges and Manual Rain gauges (R2,R3,R4,R5,R6,RV1)	R1061523.csv, R7061523.csv, EM061523.csv, Precipitaciones La Corona 2023.xls	1_quickcheck_rain.R
Weather/Weather Station	EM061523.dat	1_quickcheck_weather.R
Flume/Hobou20	S1061523.csv, S2061523.csv v3p111219.csv, V4p111219.csv	quick_checkflumesWIP.R