# CEVE 101: Project 03

Jovanny Nazario (Jn77)

2024-11-25

We begin by loading the required packages.

```julia
using Revise
using RainwaterHarvesting
```

**Line 1**

Revise allows us to update code without restarting Julia.

**Line 2**

We load the RainwaterHarvesting package that contains our simulation tools.

## Setting the file path

Next, we specify the path to the rainfall data file and read it in.

```julia
filepath = "data/2.txt"
rainfall_data = RainfallData(filepath)
```

**Line 1**

This is the path to the rainfall data file for the station of interest.

**Line 2**

This creates a RainfallData object containing the rainfall data for the specified station.

```
RainfallData for station: ACARAU
  Location: -2.8858888888889, -40.118416666667
  Years of data: 50
  Total days: 18262
```

If we dig under the hood a bit, we can see that there are two main pieces of information:

- station_info: a dictionary containing information about the station, such as its name and location.
- annual_data: a dictionary mapping years to AnnualRainfallData objects, each containing the date and precipitation data for a given year.

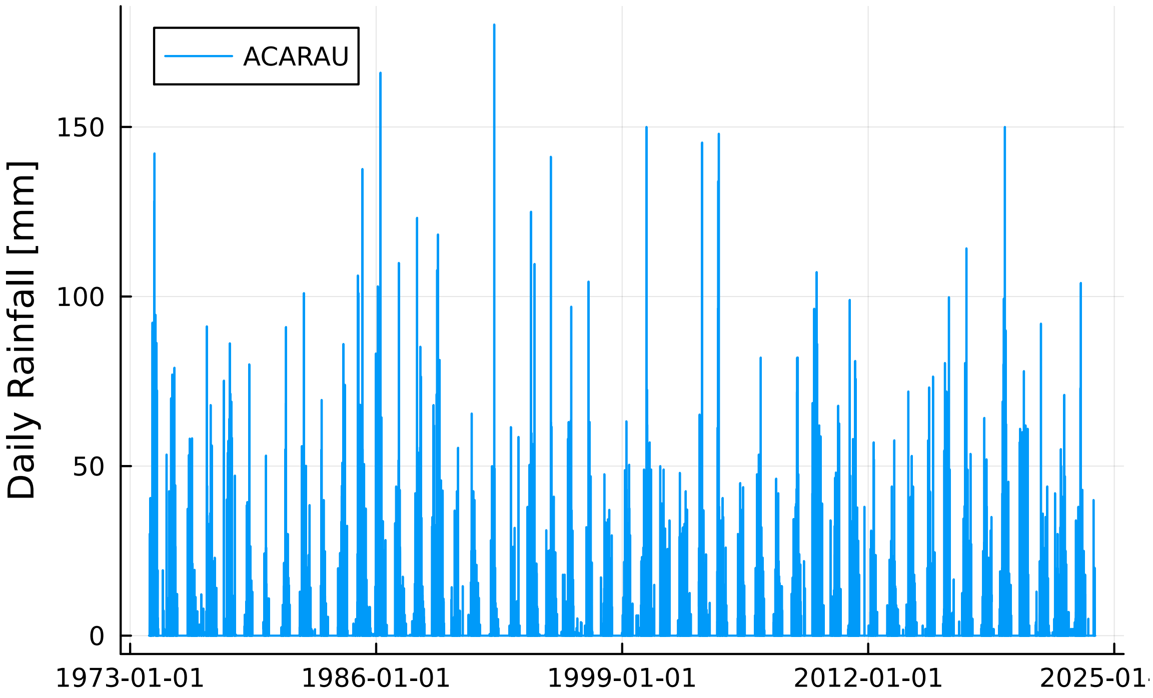
## Plotting the Data

We can plot the data for one or all of the years to verify that it looks correct.

```
plot(rainfall_data)
```

### Line 1

This `plot` function is defined in viz.jl and is a thin wrapper around the Plots.jl package.



## Discussion

After analyzing data text 2, there appears to be a long record of rainfall data! Beginning in the 1970s, there appears to be data for every year ranging between the mid 1970s and the early 2020s. There are no substantial gaps in any of the data collected, demonstrating that there was a strong collection of rainfall data. This means there is enough data to be able to identify, evaluate, and analyze if there are any substantial risks. Additionally, there appears to be no missing data within the Acarau station; research shows that there were droughts during the years: 1877–1879, 1915, 1931, 1973, 1983, 1993, and 1998. However, Acarau appears to have data recorded for all of these dates, demonstrating even further the strength data collected(Source 1).

```
Source 1: https://www.sciencedirect.com/science/article/pii/S0022169424002440#:
~:text=The%20droughts%20of%201877%E2%80%931879,solutions%20in%20response%20to%
20drought.
```

Missing data during drought years can create numerous issues on meaningful risk assessment and data analysis. For instance, missing data can produce bias in rainfall trends; rainfall data can be overestimated or underestimated if data is missing during drought years. Additionally, they can create inaccuracies in predicting future drought years and their severity which can cause even more complications for the surrounding area, both people and environment. This could further lead into inadequate preparation for these drought years if they cannot be predicted accurately. Finally, mitigation measures cannot be properly created if the severity of the problem cannot be measured due to insufficient rainfall data. Ultimately, missing data makes it difficult to assess these wide-ranging challenges and develop proper solutions to ensure the safety and wellbeing of the people and environment. The station in Acarau, however, does not face this issue due to its sufficient collection of data with no apparent drought years.

## Understanding the Theoretical Model

The modeled mathematical equations make certain assumptions that skew some of the data being evaluated. For instance, the consumption model assumes that the average household begins using water after the 150th day; this is based on the idea that the rainy season lasts for about the first 150 days of the year. This must mean that the average household has other sources of water to utilize. Additionally, the consumption model assumes that each household in the area consumes 74.1 liters per day during the dry season, however, reports convey "water consumption in Brazilian households is around 70 to 120 liters per person, per day, depending on the income of the family and on the geographic region and climate"(Source 2). Furthermore, the equation assumes exactly 74.1 Liters of water is used everyday which is almost impossible to measure perfectly and thus, will vary day to day. Also, the modeled equation does not take into account the surrounding area; possible blockage could occur by objects such as trees overhead nor does it take into account the maturity of the people in the household. For example, children will consume less water than adults because their needs are not the same as adults. Finally, the modeled equation fails to account for changes in weather patterns due to climate change or the warming of the earth and the effects this may have on the water collection system.

```
Source 2: https://www.clasp.ngo/wp-content/uploads/2021/01/Brazil-Water-Efficency-Report-1.pdf
```

After identifying these limitations and assumptions, the mathematical equations can be modified to relax some of these issues. For example, updating or changing the average household consumption of water to a more recent approximation and integrating the number of people and their maturity(age) in a household can drastically improve the accuracy of the consumption of water. A model would look as such: C_t = {95.0A+ 70.0C day of year > 150 0 day of year ≤ 150}, where A refers to the number of adults in the household and C refers to the number of children in the household. Changes in weather patterns must also be addressed; thus in order to integrate it into the modeled equations, a percentage variable must be introduced that assess changes in weather patterns and compares it to rainfall data from one to three years prior. For instance the inflow equation can be updated with a variable W(be multiplied to the equation) to make the prediction of "inflow" more accurate.

# Model Implementation and Testing

## Defining the Model Parameters

We can define the model parameters with sensible default values. You will have the opportunity to change these values later.

```
param = ModelParameters(
    runoff_coefficient=0.85,
    roof_area_m2=60.0,
    first_flush_mm=2.7,
    tank_capacity_L=17000.0
)
```

**Line 1**

This creates a `ModelParameters` object with the specified parameters.

**Line 2**

The runoff coefficient ($\eta$) is a measure of how much of the rain that falls on the roof ends up in the tank.

**Line 3**

The roof area ($A$) is the area of the roof that the rain falls on.

**Line 4**

The first flush volume ($V_f$) is the volume of rain that is discarded because it is dirty.

**Line 5**

The tank capacity ($V_{\text{max}}$) is the maximum volume of the tank.

## Discussion

An increase in the roof_area_m2 from 45.0 to 60.0 will allow for a greater amount of rainfall to be collected into the tank without having to drastically change the roof parameters. This 33% increase in roof area will decrease the number of years with failures from 3 out of 50 to 1 out of 50 while not creating too much extra work needed to be done. Additionally, since there is more roof area, more water needs to go into being the first_flush diversion, thus the first flush was increased proportionally from 2.0 to 2.7 in order to make sure the water is sanitary and of good quality. Finally, since there is any increase of inflow water, the tank capacity has been very slightly increased by 1000 liters in order to accommodate this change. After calculating the reliability percentage of these modifications: (49 successful years/50 total years)*100 = 98.0%, it can be seen that these changes have a positive impact on the overall success of the water capturing and storage system.

Nevertheless, there is still a current limitation in the model; the model fails to address the concern for the time it would take to make these changes to the parameters of the roof area, first flush, and tank capacity. This limits the ability to predict the effect of the construction period on the rainfall data collected. Additionally, the current rainfall data uses trends from the previous parameters to predict rainfall trends. Thus, these new parameters may not be as effective when predicting rainfall data in 2024 and later years when weather patterns are varying due to the warming of the earth. In order to address this gap, the rainy season should be studied further to determine whether there are significant changes to predictions in future rainfall data collection. In order to integrate it into the modeled equations, a percentage variable must be introduced that assess changes in weather patterns and compares it to rainfall data from one to three years prior. This will enhance the accuracy of the rainfall data predictions and help predict any severe changes or failures before they occur to mitigate any damage to the households that depend on this source of water during the dry season.

## Running the Simulation for One Year

Let's run the model for the year 1981.

```
rainfall_1981 = rainfall_data.annual_data[1981]
results_1981 = run_timesteps(rainfall_1981, param)
p1 = plot(results_1981)
```

**Line 1**

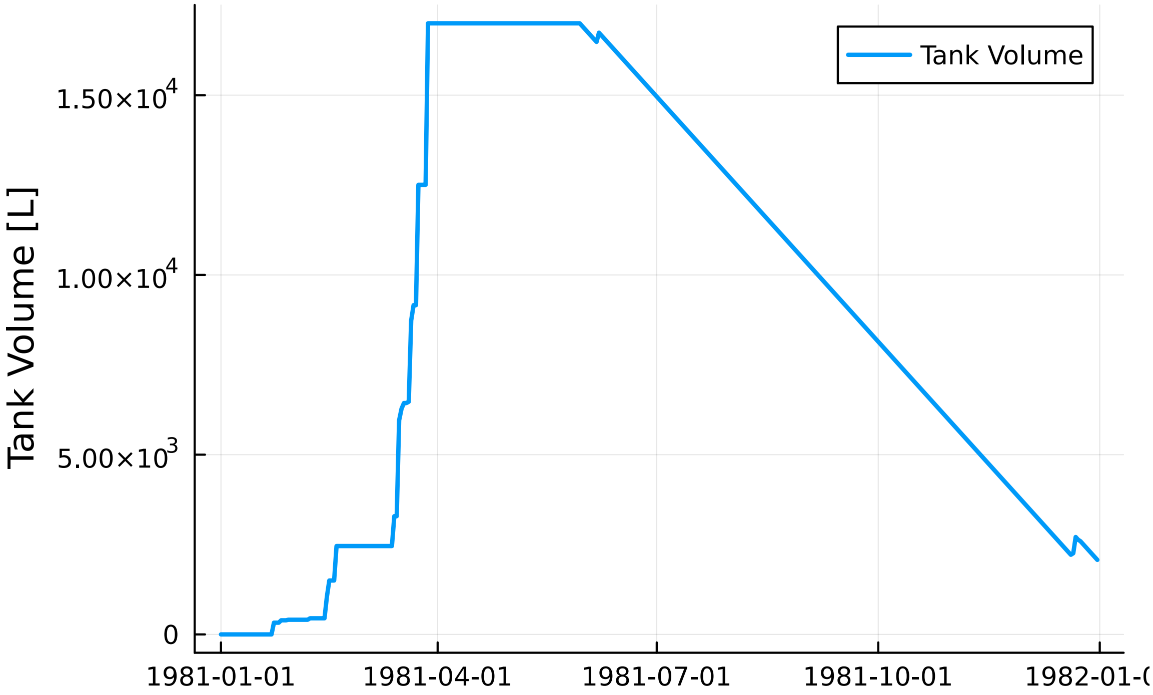
We select the annual rainfall data for the year 1981.

**Line 2**

We run the simulation for the year 1981.

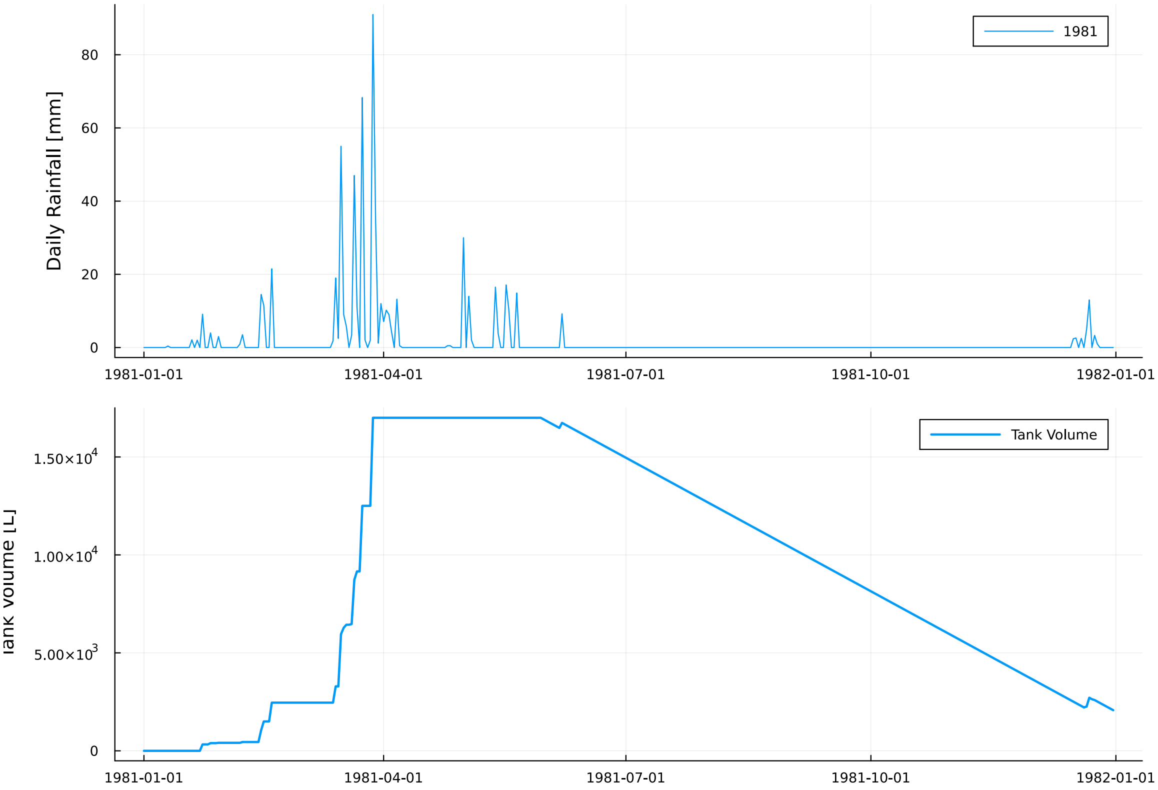
**Line 3**

We plot the results, again using a `plot` function defined in viz.jl.



To compare rainfall and tank volume:

```
p2 = plot(rainfall_1981)
plot(p2, p1, layout=(2, 1), size=(1000, 700), link=:x)
```



Observe how the tank's water level responds to rainfall events. Note any periods when the tank runs dry or overflows.

## Reliability Analysis

We can run simulations for all years, one at a time, and then check how many times the tank runs dry. You can run simulations for all years at once with the following code:

```
all_years = sort(collect(keys(rainfall_data.annual_data)))
all_results = [run_timesteps(rainfall_data.annual_data[year], param) for year
in all_years]
any_failures = [!isempty(result.failure_dates) for result in all_results]
println("Number of years with failures: ", sum(any_failures), " out of ",
length(all_years))
```

### Line 1

We get all the years in order.

**Line 2**

We run the simulation for each year.

**Line 3**

We check if any failures occurred.

**Line 4**

We print the number of years with failures and the total number of years.

```
Number of years with failures: 1 out of 50
```