1. Made the values obtained by the node more accurate by using a lookup table. ESP ADC very noisy, reduced noise to get improved readings.
2. Calculated flowrate from total flow readings and timestamps:

Flowrate = water flowed(kL)/time passed(hr)

**Calculate hours between two times**:

To present the difference between two times as a **decimal number**, use this formula:

**=(*End time* - *Start time*) \* 24**

Supposing that your start time is in A2 and end time in B2, you can use a simple equation B2-A2 to calculate the difference between two times, and then multiply it by 24, which is the number of hours in one day.

Source: https://www.ablebits.com/office-addins-blog/calculate-time-excel/

1. Outlier removal: Find optimal number of clusters using Kmeans, by checking Silhouette coefficient, calinski harabasz score, and davies bouldin score. Which came out to be 2. Divide the plots into clusters, find median and standard deviation of each cluster. Define outliers as points at a distance of more than 3 standard deviations from the median of respective cluster the point belongs to.
2. Smooth data: Take a moving average of the elements of a vector using a fixed window length that is determined heuristically. (tried 5, 10)
3. Normalize data: Compute the [z-score](https://in.mathworks.com/help/matlab/ref/double.normalize.html#mw_e6886c44-2923-4074-844a-f3e1a447359d). Center data to have a mean of 0, and scale data to have a standard deviation of 1.