July 7th Presentation, Part 2

Arya Shahdi

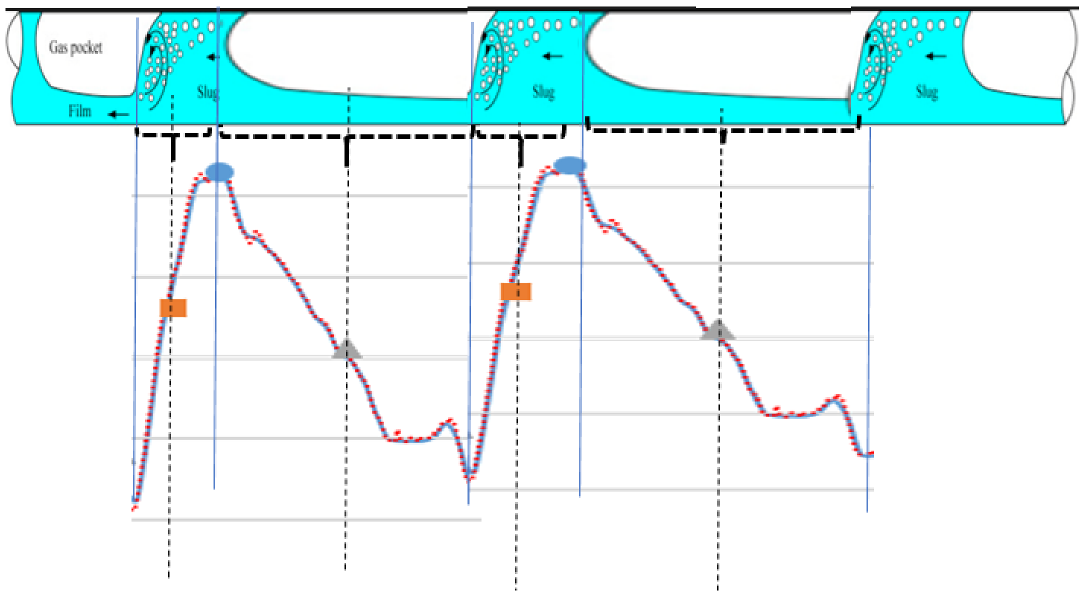
Flow Chart and validation of methodology

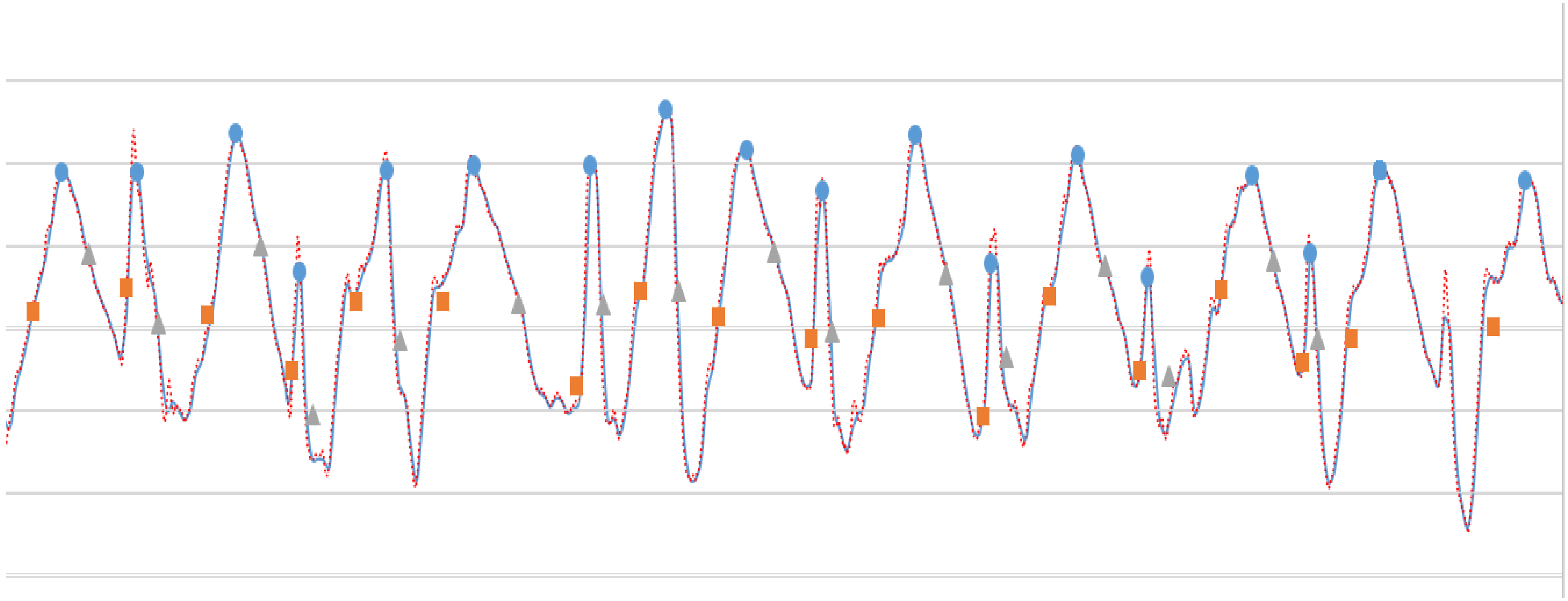
Frequency

- As it is explained by the flow chart, the developed code is able to obtain $Frequency, L_s, H_{LLS}$. To briefly explain the code and the procedure behind it, it is better to start with the used methodology to find peaks and bottoms among data. At first, the code will scheme through the data and find all the minimum and maximums (local and global) and then ignores the ones that are considered to be noises. There are two numbers that are associated with this global maximum picking. In the portion of the code where it separates global maximum from the noises (local maximums), it peaks those maximums that are above a certain limit which is called "Lower Limit Peak" and then move backward through the data till a minimum that is less than a certain value is met. The second limitation is called "Upper Limit Bottom". Now at the point, those peaks that had already the mentioned characterizations (explained above) will be considered as slugs for frequency determination. It is noteworthy to say that the reported frequency is highly dependent on the two limitations. Further experiments should be done to determine these two limits. In another word, LLP is the minimum voltage of a slug and UPB is the maximum voltage of film.
- Now that the number of slugs and their associated times are known, the rest of the code will focus on calculating the other characterizations of the flow including $L_s, H_{LLS}, H_{LTB}, L_f$.

Slug and film hold up and length

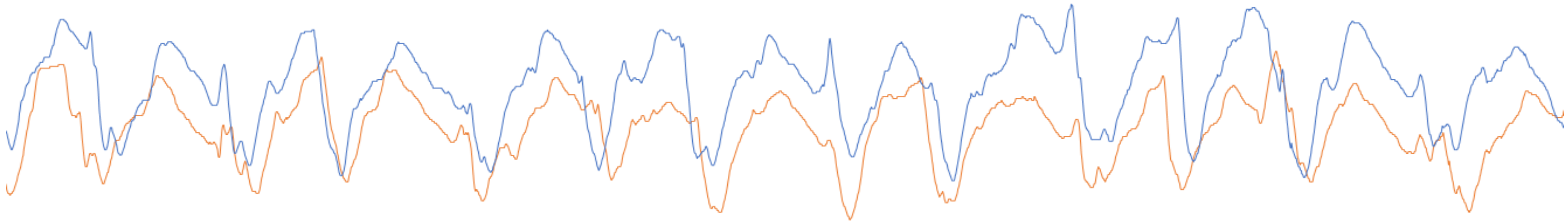
- This section will be explained by an example. In below plot, you can see that each peak (associated to one slug) is determined by blue circles. The orange squares are the averaged H_{lls} and the gray triangles are average film hold up.
- For each slug, from the peak back to the previous minimum, normally a steep line is seen. It is believed that that line represents each slug. And comparatively, the right line (from the peak to the next bottom is film and the averaged hold up is for film section.
- For each slug, the time associated with film and slug are known and if transitional velocity is known, the length can be determined. So V_{tb} should be firstly measured. The next section, the methodology used to get the V_{TB} is explained.


Gas pocket
Film
Slug
Slug
Slug



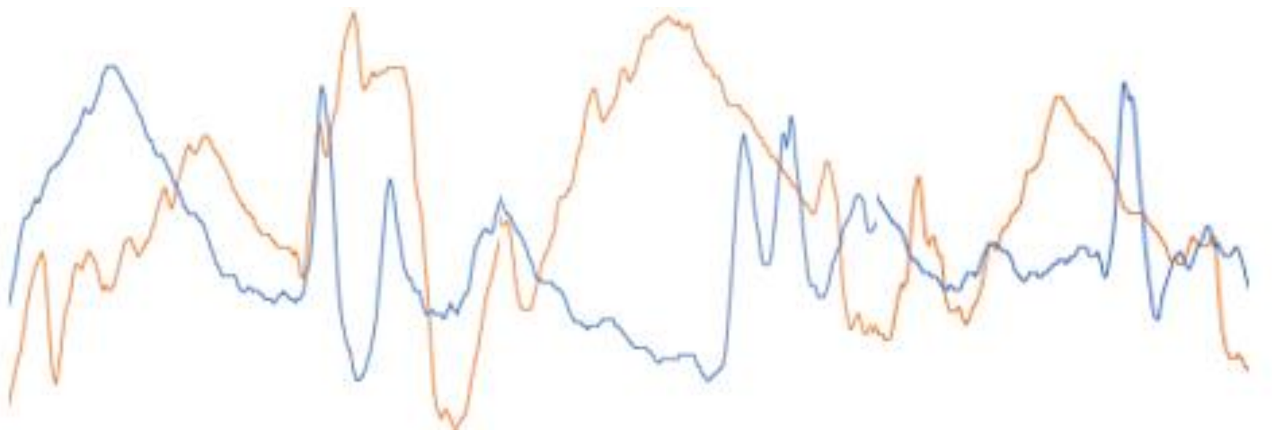
Transitional Velocity Determination

- I put several hours to find how transitional velocity can be determined by writing a code. However, since the characteristic of a slug will change considerably from capacitance sensors, peaks and slug noises cannot be traced by censors. However, if an integrated optimizing code is developed, it is very likely to capture the velocity. For the time being, I manually tried to match the data and get the time difference and while the length between capacitance sensors are known, the velocity could be calculated.



Transitional Velocity Determination

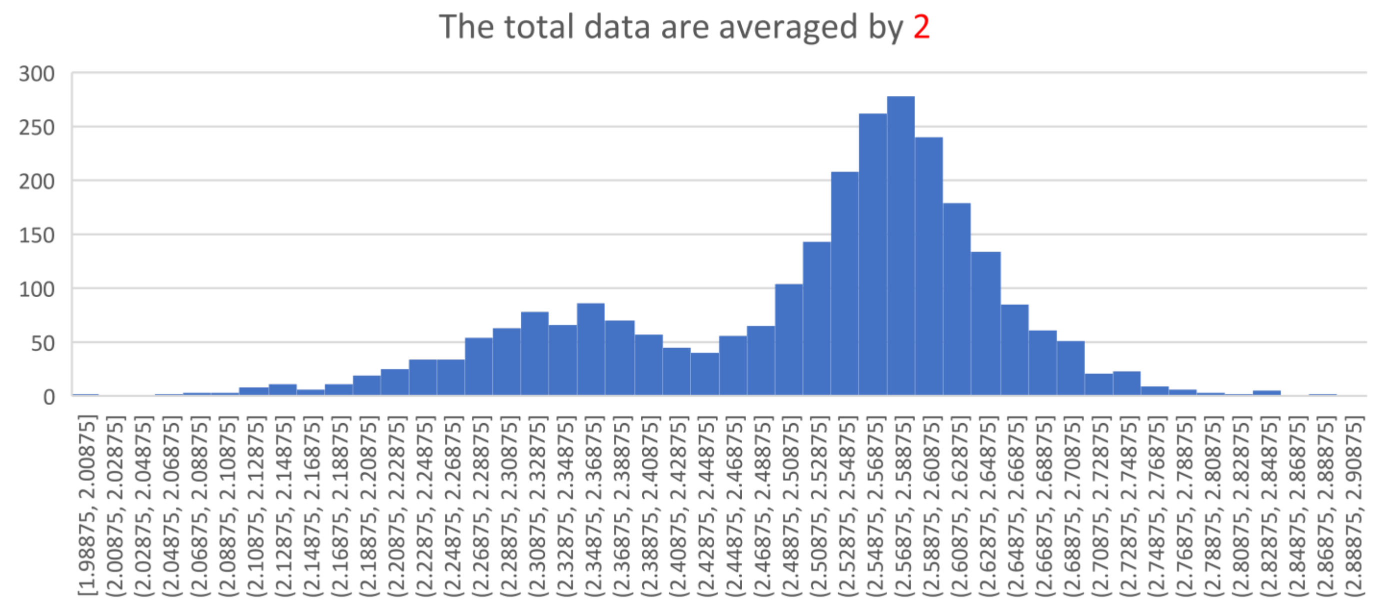
- The next plot will show some irregularities even though the above perfect match between data are seen for the dame file.



BUT, the approximation is fairly acceptable and is somehow valid

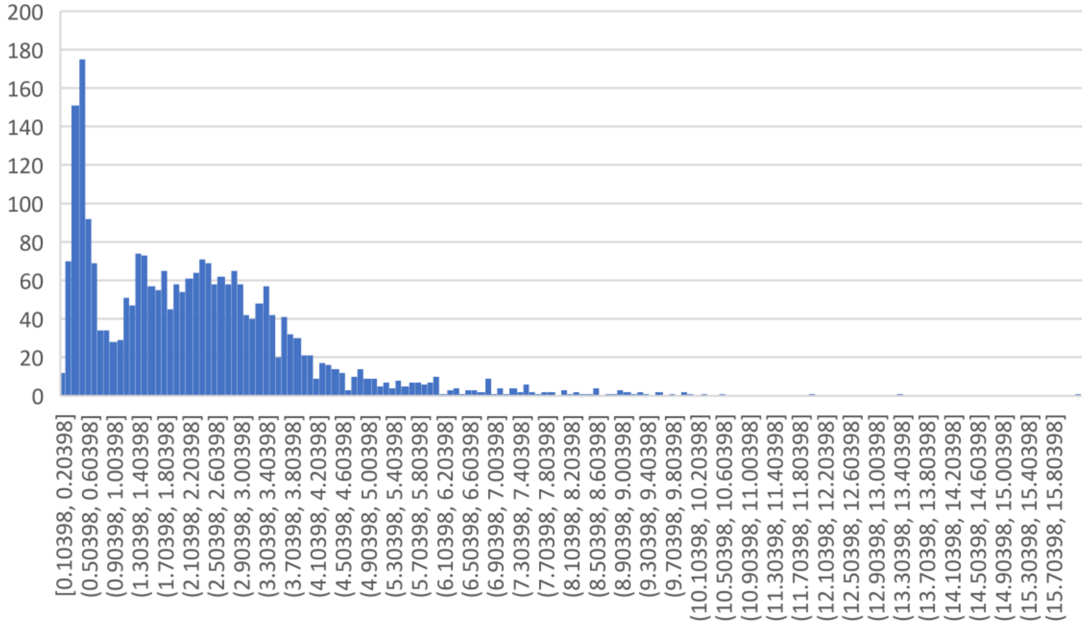
Slug and proto slug characterization

- The data are disturbed by flashes that are known as Proto-Slugs and there is no available assigned approach to distinguish proto-slugs from the actual slugs captured by simulation. This was the main limitation to obtain reliable information to be used in mechanistic models.
- The speculation was to have the slug length and slug hold-up to be lower for proto-slugs. By plotting the histogram of slug hold-ups for all slugs, it was found out that two humps are present. An example is present here



Slug and proto slug characterization

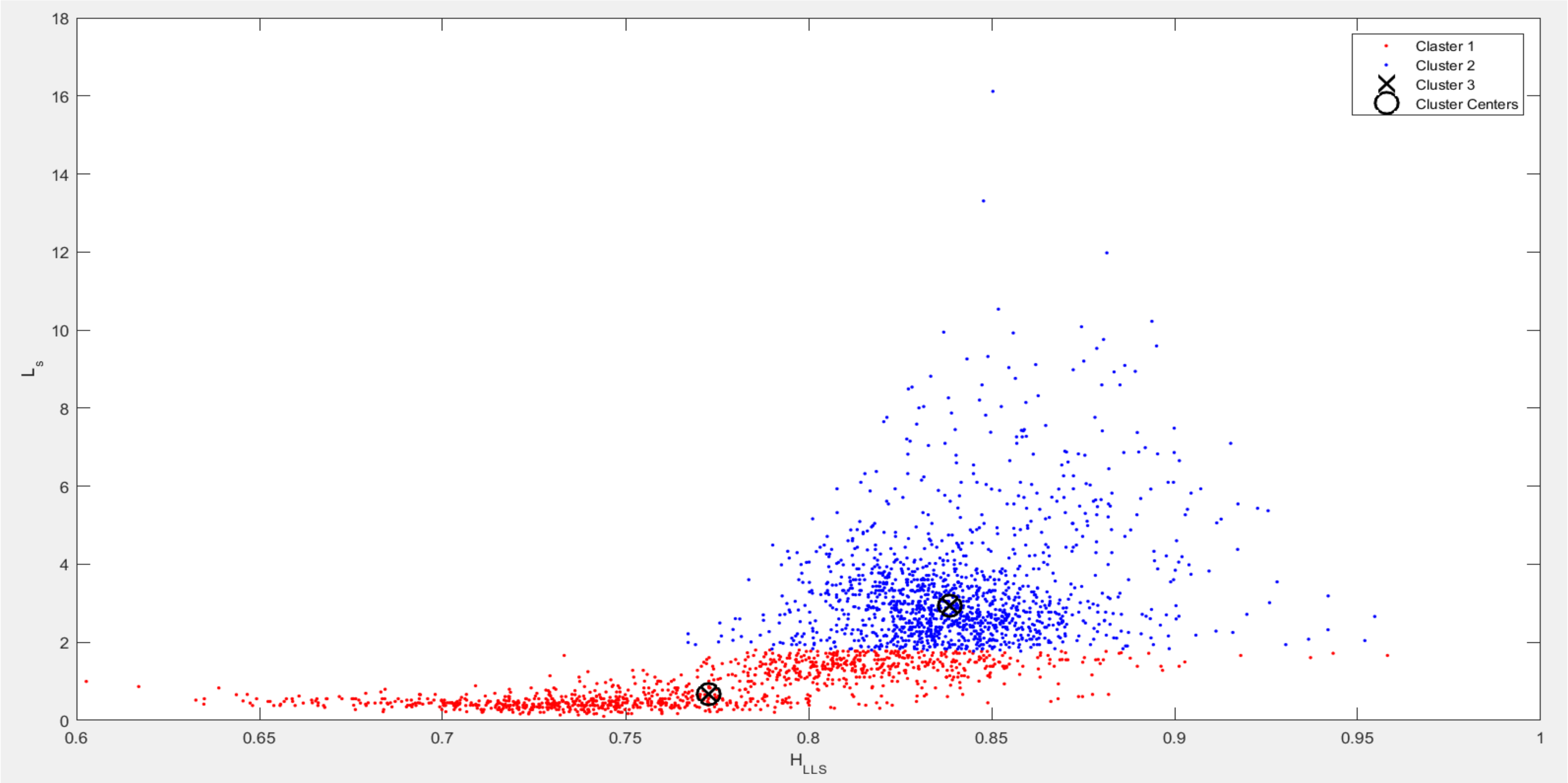
- Still with the above data, it was hard to distinguish all the data for slugs and proto-slugs.
- Another speculation was to have slug length of proto-slugs considerably lower than the slug length of slugs. But the histogram was not as distinctive as slug hold up.



Slug and proto slug characterization

- Now that we know neither of two histograms are helpful enough to characterize and distinguish the slugs and proto-slugs, I have come up with a method that includes clustering method using K-means (Lloyd’s algorithm) to categorize the data into two sets of clusters to have one cluster describing slugs and the other one proto-slugs.
- The developed code was given with a set of data with two dimensions, first one was H_{LLS} and the other one was L_s.
- This unsupervised (unlabeled) method of optimization clusteres the data and reportes the center of each cluster as the representative of that category (which are really close to the mean of that cluster).

Slug and proto slug characterization



Slug and proto slug characterization

Cluster Centers (By MATLAB)		
	H_LLS	L_s
Cluster 1 (Center)	0.772754	0.665472
Cluster 2 (Center)	0.838565	2.9461

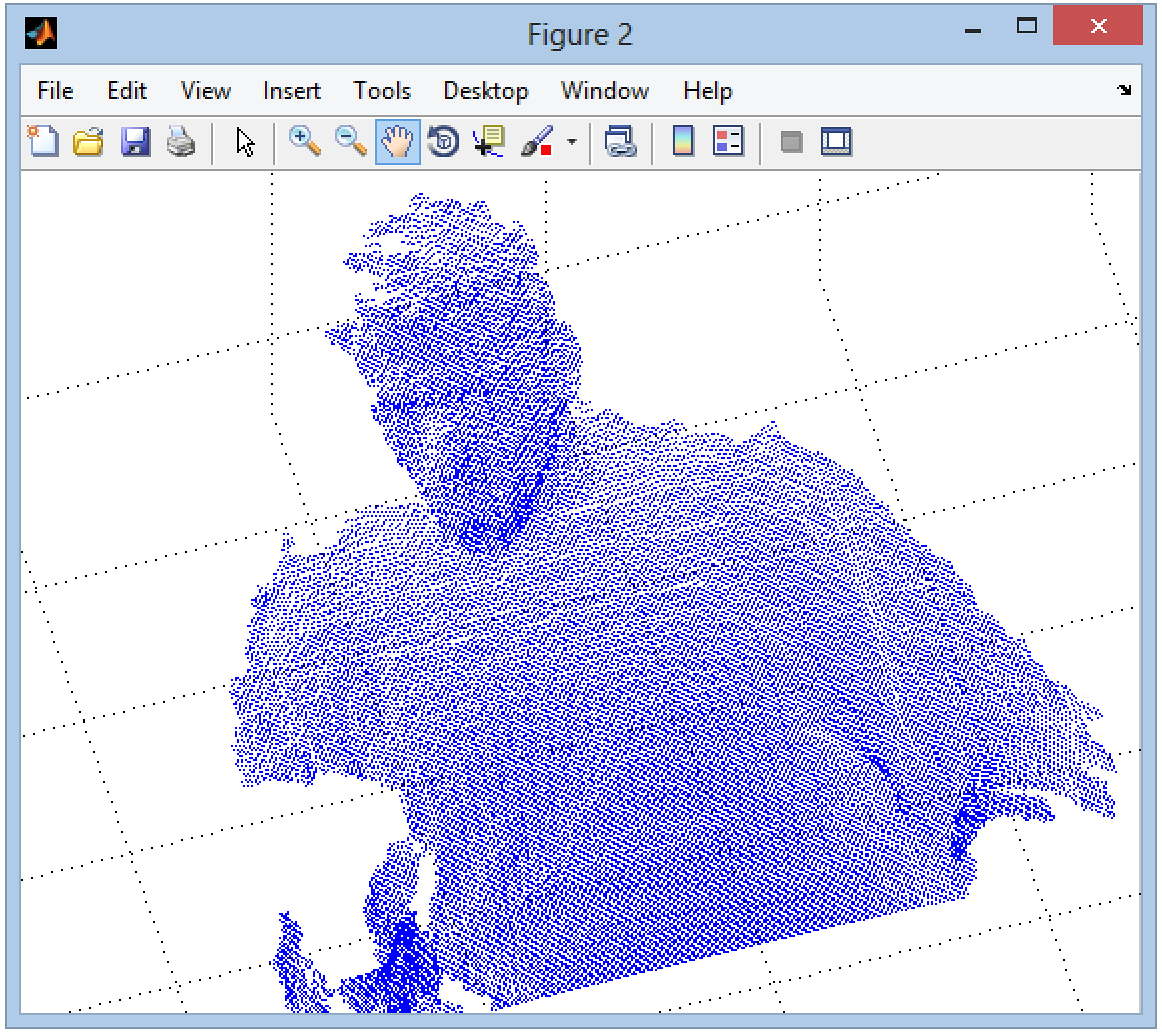
Averages		
	H_LLS	L_s
Cluster 1 (Center)	0.772135	0.872802
Cluster 2 (Center)	0.841142	3.406232

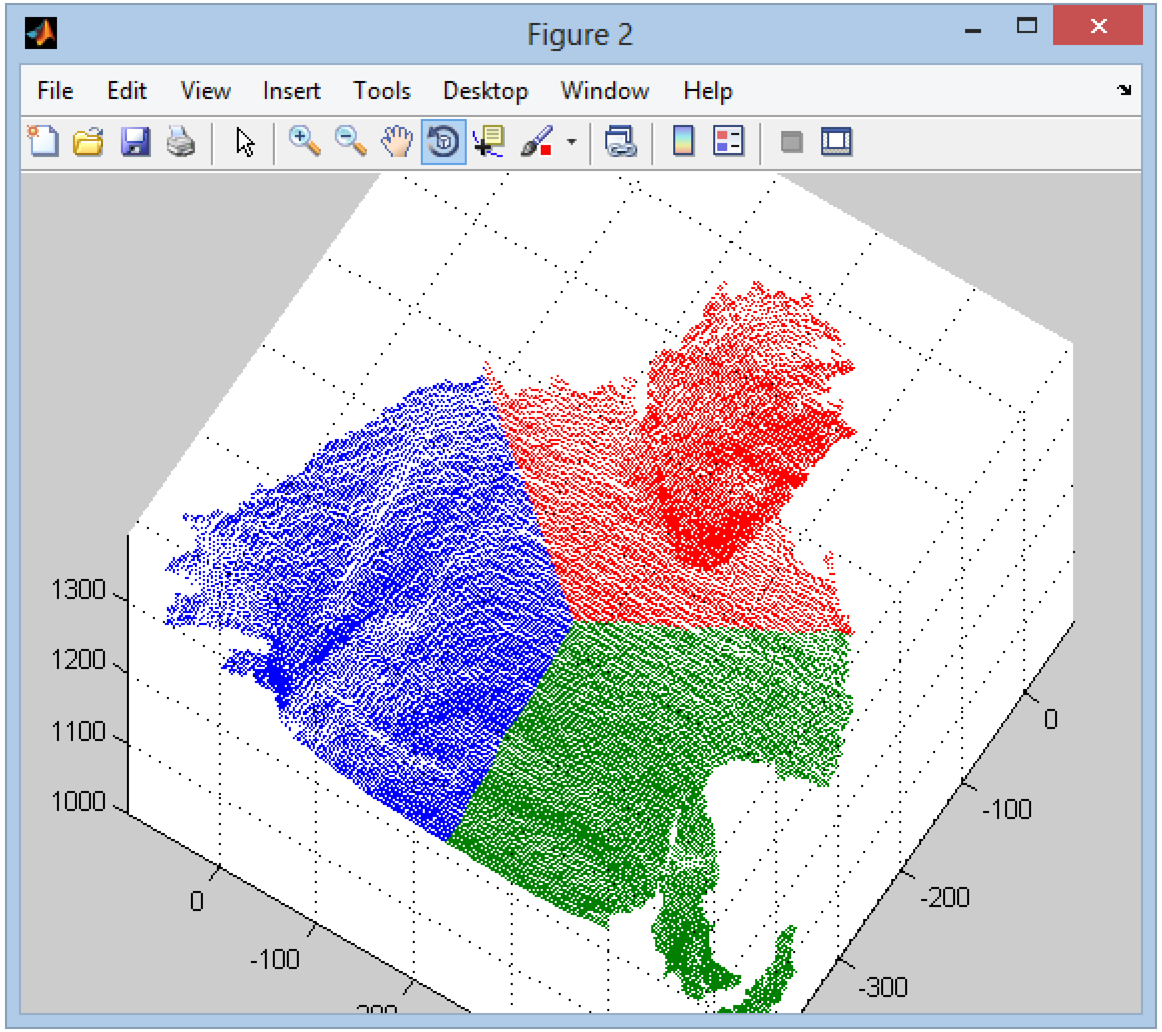
K-means Clustering Method

- Firstly, the number of required clusters are chosen as an input and then two points are chosen within the range of data and then all the data are introduced to those centers one by one and the cluster which has the less distance to that introduced data will be winner and that data will be associated to that center (or cluster) and when all the data are labeled to their clusters, the total distances of data to their presentative center of cluster will be calculated. For the next step, the location of the center will be changed and the new distance summation is calculated and compared to the previous ones and if it was less, it will be reported as the optimum center location and classification.

https://datasciencelab.files.wordpress.com/2013/12/p_n2000_k15_.gif

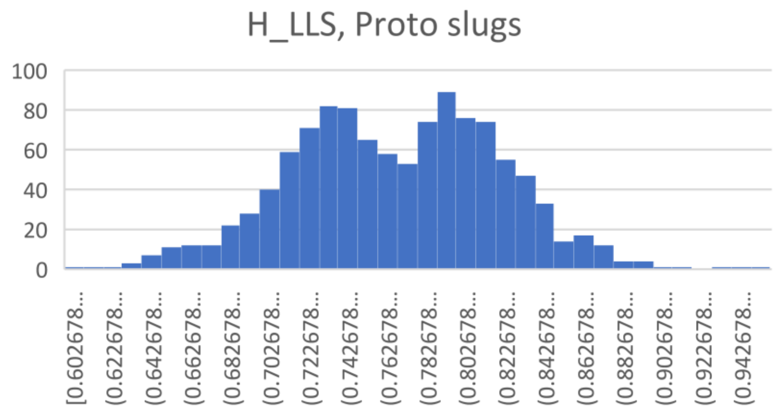
K-means Clustering Method

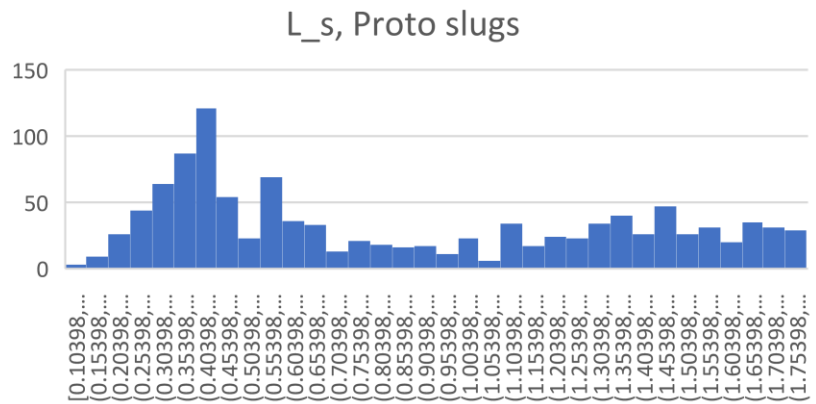


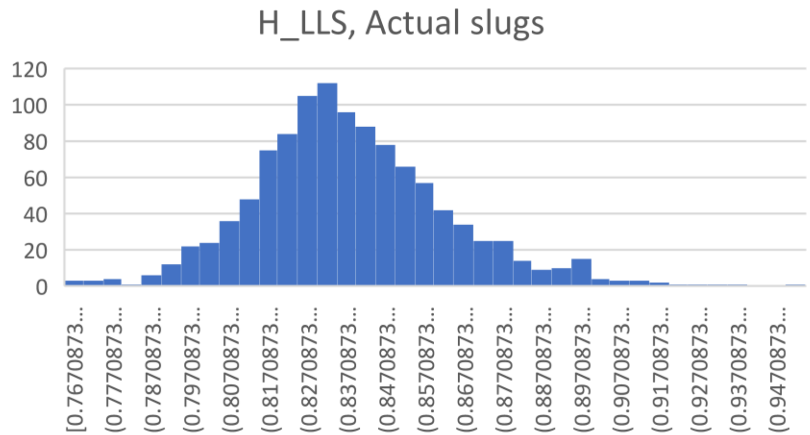


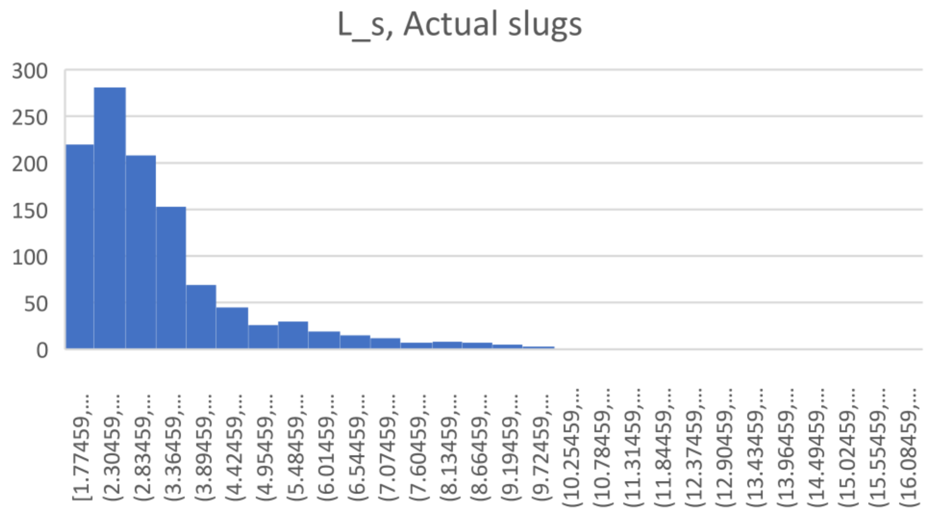
K-means Clustering Method Application

- Now that all data are separated into two clusters and classifications, we can count the number of data associated with two clusters and then report the number of actual and proto slugs. The following histograms of H_{LLS}, L_s are for cluster one and two.

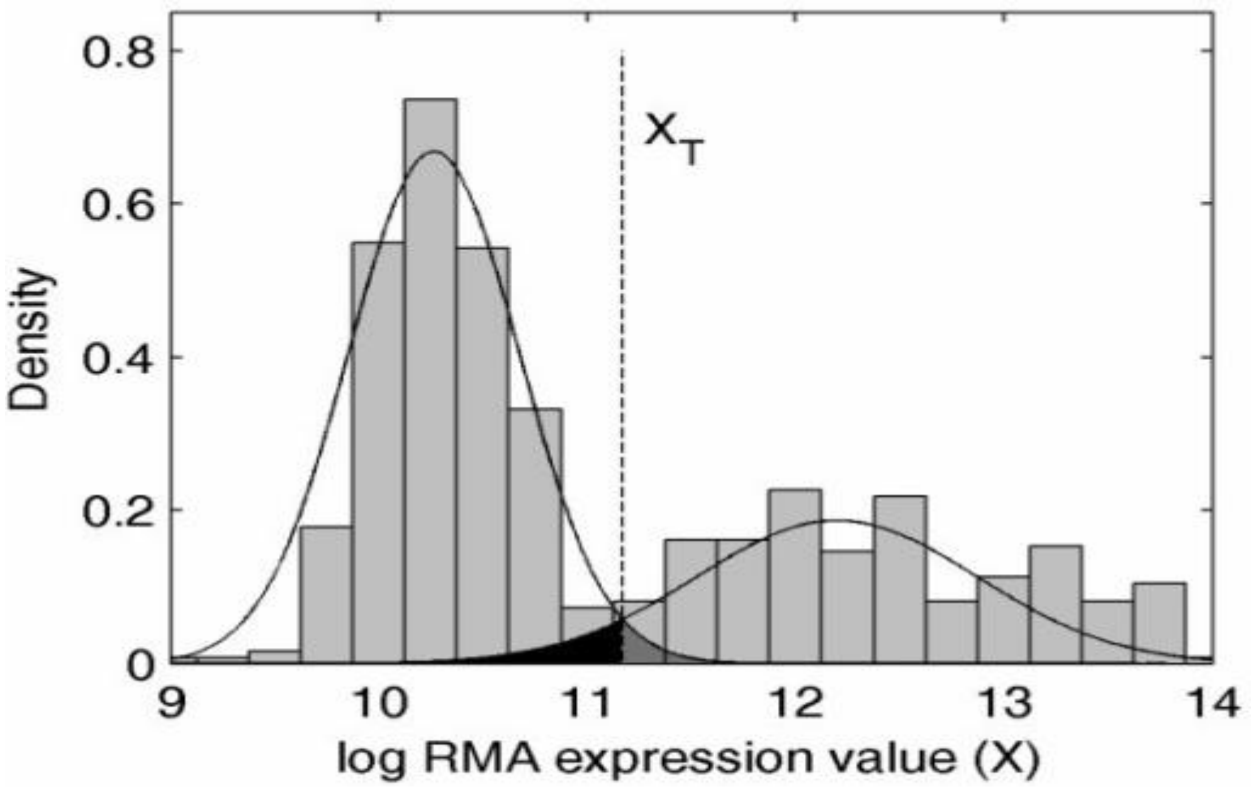






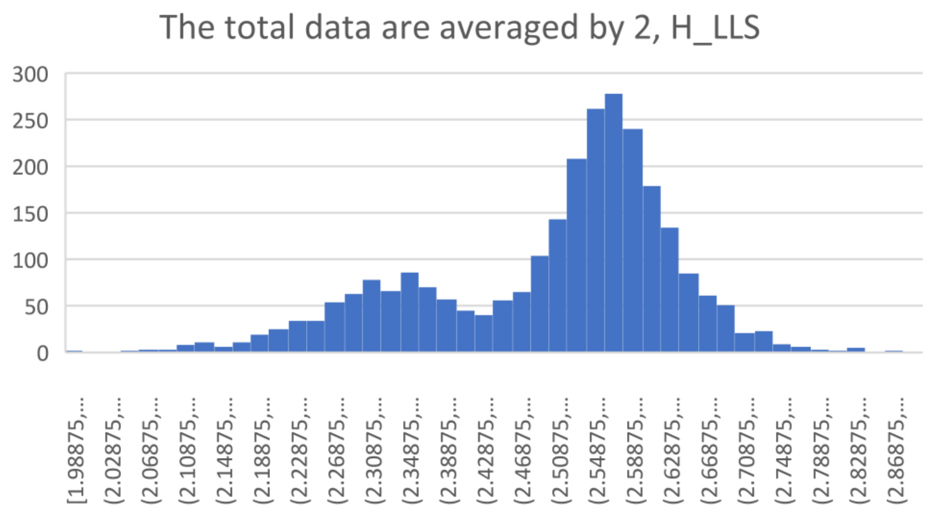


Bimodal Normal Distribution characterization



Bimodal Normal Distribution characterization

- We found a code in R-language that does the separation and we got the average H_{LLS} for all slug holdups. Interestingly, those came out to be very similar two those calculated from clustering method.



HLLs		
	Cluster1	Cluster 2
Clustering method	0.772754	0.83856548
Bimodal method	0.780504	0.8369034
error%	0.992914	0.21294915

Further Recommended Investigation

- LLP (the minimum voltage of a slug) and UPB (the maximum voltage of film) are input for the code and they are chosen by the user. A method should be developed to certify which two limitations are valid. One recommended method is to have the capacitance data and record the video of the slugs. Then match the film and capacitance censors and detect the proper slugs and their corresponding peak and bottom within the data.
- Two experiment with the same flowing condition should be implemented and results should be consistent. As many experiments as possible should be done and corresponding flow characterizations should be reported and compared to closure relationships and recorded videos.
- Still V_{TB} is an input from the user and should be calculated based on a judgment call. A methodology is required to develop to include it in the code.
- Clustering is just one way of characterizing the data. More statistical analysis is required to describe the data