Liposome Flux Assays in the Context of NavAb meet R

Motivation

- The difficulty of using the FLUOstar plate reader is analysis of data.
- The lab involved utilized MATLAB for conversion of a .DAT file into a csv file with normalized data for further analysis.
- The main issue is clearance for the MATLAB program resided with the post-doc (not important to me anymore) and the professor (also in my mental black hole), leaving me with the task of analyzing the data.
- Moreover, any differences in the lab procedure entailed recycling the code to accommodate for example loading samples vertically vs horizontally.
- Finally, importing .DBF files into Excel was not possible as data was imported in a distorted manner.
- The only solution was to develop R code to implement analysis of data using the .DBF output as opposed to the .DAT output.
- While the concern of recycling code wasn't really mine, I was also ambitious to see if I could achieve recursive code without needing to rewrite code if lab personnel performed the same task differently.
- As will be seen recursion was not achieved although still a work in progress but so far it seems my solution might lie within shiny apps after some clean up of the transformed .DBF file.

Importing Cleaned Data Sets into R

- The procedure for cleaning up the data sets will be provided in jupyter notebook as Rmarkdown seems to force organization and as said earlier, this is not my most sophisticated work.

```r
#import my principal data set
navab227_norm2 <- read.csv(file = 'navab227_norm3.csv', header = T)
navab221_avgnorm <- read.csv(file = 'navab221_norm3.csv', header = T)
head(navab227_norm2)
```

```
##   WELLNUM Timeadj Time E01 E02 E03  E04  E05  E06   E07   E08   E09 F01 F02 F03
## 1       1       0   30 273 299 259 2155 2090 2024 10533 10170 10243 352 279 187
## 2       2      30   60 266 294 253 2132 2082 1992 10429 10159 10166 344 273 184
## 3       3      60   90 266 295 255 2121 2080 1994 10389 10098 10116 338 265 175
## 4       4      90  120 270 291 257 2123 2088 1971 10302 10130 10080 345 274 185
## 5       5     120  150 263 284 255 2113 2097 1971 10312 10081 10106 343 275 183
## 6       6     150  180 268 289 249 2112 2089 1967 10239  9953 10007 346 269 184
##    F04  F05  F06   F07   F08   F09 G01 G02 G03  G04  G05  G06  G07  G08  G09
## 1 1821 1813 1805 11077 11079 10672 187 163 168 1608 1760 1601 9280 9236 8882
## 2 1806 1783 1789 10980 10916 10571 182 171 168 1613 1760 1587 9232 9111 8761
## 3 1792 1777 1780 10939 10928 10544 180 157 159 1606 1765 1589 9209 9110 8772
## 4 1812 1798 1779 10974 11011 10545 179 156 158 1603 1752 1591 9180 9104 8746
## 5 1802 1790 1789 10946 10891 10499 178 165 161 1606 1753 1583 9085 9074 8742
## 6 1812 1779 1799 11003 10989 10597 179 159 163 1622 1758 1598 9102 9062 8754
##       E2uA     F2uA     G2uA    E2norm    F2norm    G2norm    E20uA    F20uA
## 1 2089.667 1813.000 1656.333 1.0000000 1.0000000 0.9861751 10315.33 10942.67
## 2 2068.667 1792.667 1653.333 0.8917526 0.8880734 0.9723502 10251.33 10822.33
## 3 2065.000 1783.000 1653.333 0.8728522 0.8348624 0.9723502 10201.00 10803.67
## 4 2060.667 1796.333 1648.667 0.8505155 0.9082569 0.9508449 10170.67 10843.33
## 5 2060.333 1793.667 1647.333 0.8487973 0.8935780 0.9447005 10166.33 10778.67
```

```
## 6 2056.000 1796.667 1659.333 0.8264605 0.9100917 1.0000000 10066.33 10863.00
##      G20uA    E20norm   F20norm    G20norm     E0uA     F0uA     G0uA    E0norm
## 1 9132.667 0.45405112 1.0000000 0.39800285 277.0000 272.6667 172.6667 1.0000000
## 2 9034.667 0.34964655 0.9078376 0.18830242 271.0000 267.0000 173.6667 0.7500000
## 3 9030.333 0.26753670 0.8935410 0.17902996 272.0000 259.3333 165.3333 0.7916667
## 4 9010.000 0.21805329 0.9239214 0.13552069 272.6667 268.0000 164.3333 0.8194444
## 5 8967.000 0.21098423 0.8743937 0.04350927 267.3333 267.0000 168.0000 0.5972222
## 6 8972.667 0.04785209 0.9389839 0.05563481 268.6667 266.3333 167.0000 0.6527778
##      F0norm   G0norm  X                X.1         X.2         X.3
## 1 1.0000000 0.953125 NA
## 2 0.8731343 1.000000 NA
## 3 0.7014925 0.609375 NA standard error 2uM
## 4 0.8955224 0.562500 NA                Exp         Neg         Pos
## 5 0.8731343 0.734375 NA        0.047957719 0.050837673 0.051479985
## 6 0.8582090 0.687500 NA
```

```
head(navab221_avgnorm)
```

```
##   WELLNUM Time  A01  A02  A03  A04  A05  A06  A07  A08  A09  A10  A11  A12  B01
## 1       1    0 4532 4977 4675 6757 6859 6556 7783 7957 7912 9212 9056 9312 4365
## 2       2   33 4447 4885 4618 6673 6739 6506 7639 7827 7770 8996 8898 8984 4312
## 3       3   66 4391 4842 4544 6634 6655 6473 7586 7769 7714 8848 8758 8979 4281
## 4       4   99 4342 4768 4512 6501 6624 6435 7511 7717 7633 8776 8731 8933 4265
## 5       5  132 4353 4791 4496 6566 6552 6481 7528 7666 7658 8818 8647 8884 4280
## 6       6  165 4338 4741 4489 6505 6559 6415 7482 7692 7636 8768 8641 8884 4270
##    B02  B03  B04  B05  B06  B07  B08  B09  B10  B11  B12  C01  C02  C03  C04
## 1 4396 4651 6986 6784 6480 8275 7828 7695 9124 8960 8775 3817 3694 3753 5760
## 2 4310 4576 6880 6727 6370 7949 7665 7593 8935 8803 8630 3758 3635 3713 5645
## 3 4262 4507 6804 6677 6343 7898 7623 7508 8879 8759 8470 3710 3576 3663 5561
## 4 4248 4466 6749 6593 6294 7883 7571 7496 8738 8732 8405 3693 3557 3654 5564
## 5 4237 4471 6791 6648 6289 7877 7541 7460 8773 8709 8430 3683 3572 3610 5510
## 6 4222 4446 6810 6534 6247 7850 7461 7420 8661 8617 8389 3698 3562 3609 5530
##    C05  C06  C07  C08 C09  C10  C11  C12     P5um     N5um     E5um    Pnorm5
## 1 5807 5381 6751 6671 145 7889 7778 7267 7644.667 8953.000 9193.333 1.0000000
## 2 5710 5344 6618 6502 145 7786 7654 7133 7524.333 8789.333 8959.333 0.9320918
## 3 5644 5271 6579 6450 139 7769 7568 7051 7462.667 8702.667 8861.667 0.8972912
## 4 5661 5300 6562 6405 143 7761 7552 7038 7450.333 8625.000 8813.333 0.8903311
## 5 5567 5190 6536 6358 162 7728 7521 6986 7411.667 8637.333 8783.000 0.8685102
## 6 5575 5232 6531 6333 187 7705 7434 6975 7371.333 8555.667 8764.333 0.8457487
##      Nnorm5    Enorm5      T2a      T3a      T4a      T5a    T2norm    T3norm
## 1 1.0000000 1.0000000 4728.000 6724.000 7884.000 9193.333 1.0000000 1.0000000
## 2 0.9124777 0.8923478 4650.000 6639.333 7745.333 8959.333 0.9008895 0.9129242
## 3 0.8661319 0.8474160 4592.333 6587.333 7689.667 8861.667 0.8276154 0.8594446
## 4 0.8245989 0.8251802 4540.667 6520.000 7620.333 8813.333 0.7619653 0.7901954
## 5 0.8311943 0.8112253 4546.667 6533.000 7617.333 8783.000 0.7695892 0.8035653
## 6 0.7875223 0.8026376 4522.667 6493.000 7603.333 8764.333 0.7390936 0.7624272
##      T4norm    T5norm     E3um     N3um     P3um    Enorm3    Nnorm3    Pnorm3
## 1 1.0000000 1.0000000 6724.000 6750.000 5649.333 1.0000000 1.0000000 1.0000000
## 2 0.8624793 0.8923478 6639.333 6659.000 5566.333 0.9129242 0.9327917 0.9365605
## 3 0.8072727 0.8474160 6587.333 6608.000 5492.000 0.8594446 0.8951256 0.8797452
## 4 0.7385124 0.8251802 6520.000 6545.333 5508.333 0.7901954 0.8488429 0.8922293
## 5 0.7355372 0.8112253 6533.000 6576.000 5422.333 0.8035653 0.8714919 0.8264968
## 6 0.7216529 0.8026376 6493.000 6530.333 5445.667 0.7624272 0.8377646 0.8443312
##       E2um     N2um     P2um    Enorm2    Nnorm2    Pnorm2  X
## 1 4728.000 4470.667 3754.667 1.0000000 1.0000000 1.0000000 NA
```

```
## 2 4650.000 4399.333 3702.000 0.9008895 0.9229107 0.8965291 NA
## 3 4592.333 4350.000 3649.667 0.8276154 0.8695965 0.7937132 NA
## 4 4540.667 4326.333 3634.667 0.7619653 0.8440202 0.7642436 NA
## 5 4546.667 4329.333 3621.667 0.7695892 0.8472622 0.7387033 NA
## 6 4522.667 4312.667 3623.000 0.7390936 0.8292507 0.7413229 NA
##                  X.1    Enorm2.1    Nnorm2.1 Pnorm2.1
## 1
## 2
## 3
## 4 standard error 3uM
## 5                Exp         Neg         Pos
## 6        0.047957719 0.050837673 0.051479985
```

Visualization of Fluorescence

- A plot of average normalized fluorescence (y) vs Cycle No (a proxy for time) in normal or signal conditions.
- There is some biology behind this plot and this can only be understood by comparing conditions were potassium NOT Cesium is placed within the liposome.
- The biophysics might be a lot to take in but basically, we cant make sense of the data without comparisons.
- Images to explain the difference between K inside vs Cs inside the liposome

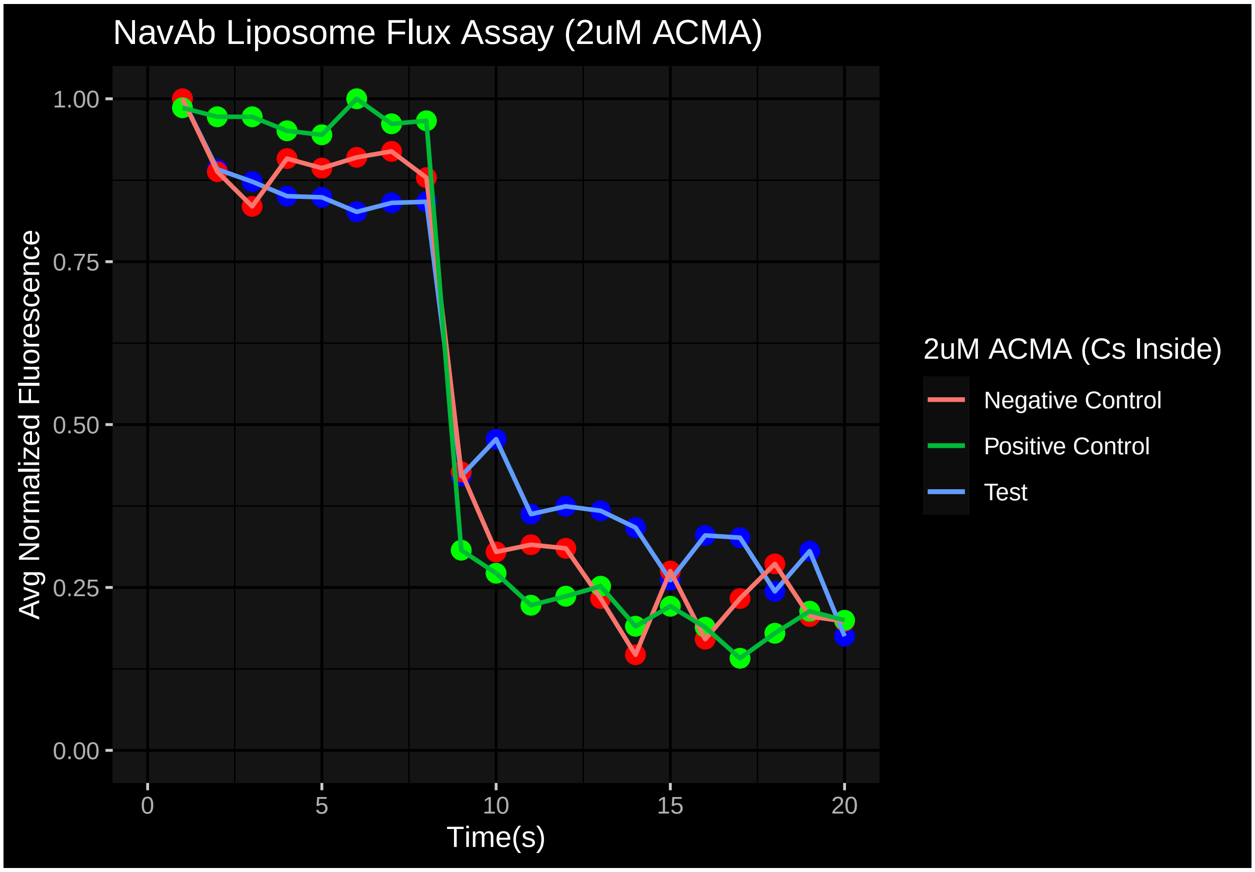
```r
# Plot in signal zone for Cs inside
library(ggplot2)
library(ggdark)
library(ggthemes)
p1 = ggplot(navab227_norm2, aes(x=WELLNUM)) +
  geom_point(aes(y=E2norm), size=3, color="blue") +
  geom_point(aes(y=F2norm), size=3, color="red") +
  geom_point(aes(y=G2norm), size=3, color="green") +
  geom_line(aes(y=E2norm, color='Test'), size=0.8) +
  geom_line(aes(y=F2norm, color='Negative Control'), size =0.8) +
  geom_line(aes(y=G2norm, color='Positive Control'), size =0.8) +
  dark_mode()+
  xlim(0,20) +
  ylim(0,1) +
  labs(title = 'NavAb Liposome Flux Assay (2uM ACMA)',
       x = 'Time(s)', y='Avg Normalized Fluorescence', color='2uM ACMA (Cs Inside)')
```

```
## Inverted geom defaults of fill and color/colour.
## To change them back, use invert_geom_defaults().
```

```r
# call plot
p1 #or print(p1)
```

```
## Warning: Removed 20 rows containing missing values (geom_point).
## Removed 20 rows containing missing values (geom_point).
## Removed 20 rows containing missing values (geom_point).
```

```
## Warning: Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
```

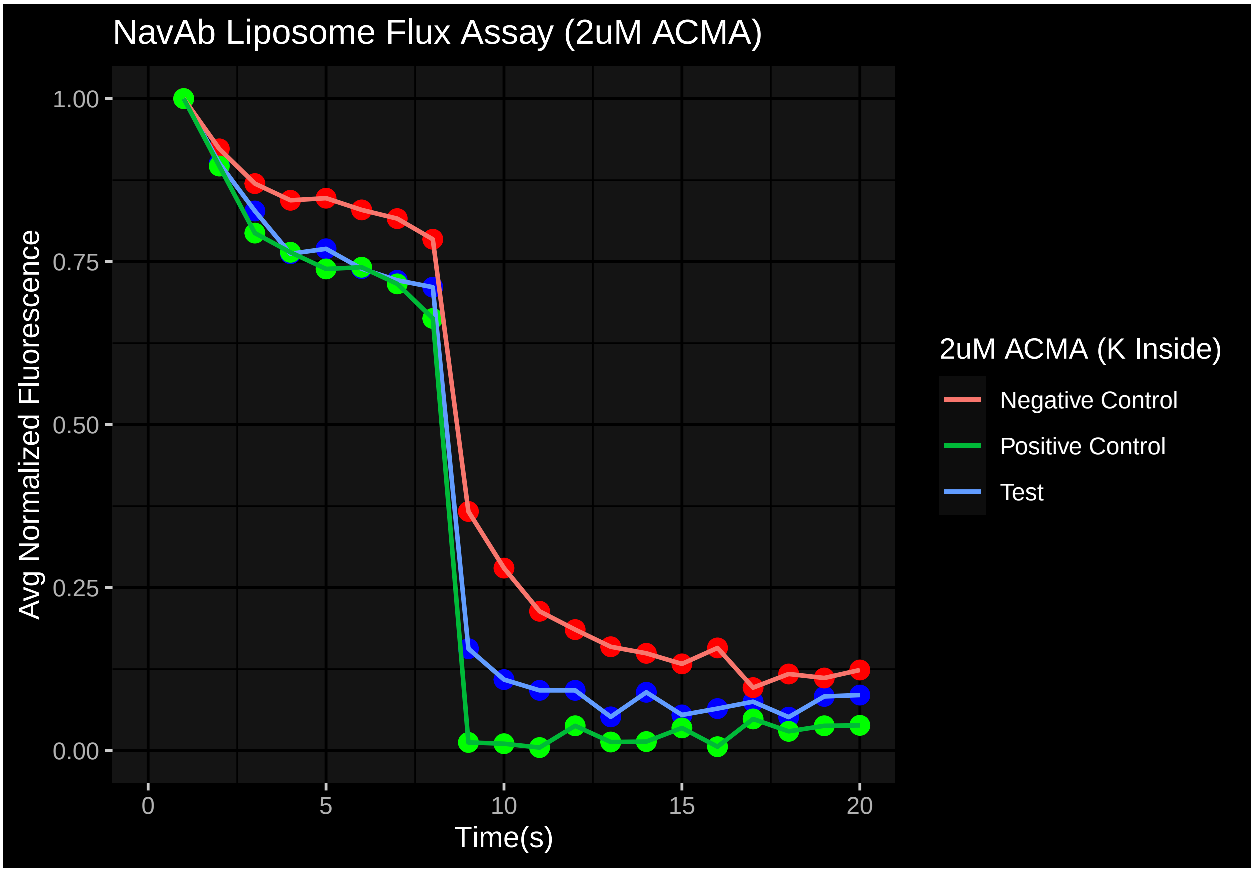



```
#Plot in signal zone for K inside
#E,N,P = exp,neg and pos respectively
#Ensure to take notes, i was working for hours and had no idea
#why i called them this
p2 = ggplot(navab221_avgnorm, aes(x=WELLNUM)) +
  geom_point(aes(y=Enorm2), size=3, color="blue") +
  geom_point(aes(y=Nnorm2), size=3, color="red") +
  geom_point(aes(y=Pnorm2), size=3, color="green") +
  geom_line(aes(y=Enorm2, color='Test'), size=0.8) +
  geom_line(aes(y=Nnorm2, color='Negative Control'), size =0.8) +
  geom_line(aes(y=Pnorm2, color='Positive Control'), size =0.8) +
  dark_mode()+
  xlim(0,20) +
  ylim(0,1) +
  labs(title = 'NavAb Liposome Flux Assay (2uM ACMA)',
       x = 'Time(s)', y='Avg Normalized Fluorescence', color='2uM ACMA (K Inside)')

# call plot
p2 #or print(p1)
```

```
## Warning: Removed 20 rows containing missing values (geom_point).
## Removed 20 rows containing missing values (geom_point).
## Removed 20 rows containing missing values (geom_point).

## Warning: Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
```



- The spoiler is when you look at the plots, you can tell the increased volatility in signal when Cs is inside relative to K inside and this is likely because Navab-Cs interactions are likely more non-specific making the result with Cs inside less predictable. - Another spoiler, within the signal zone (2uM, 3uM, 5uM are currently tested values) the results above are more predictable than not. I will reveal the signal zone at the end which is my opinion because the lower boundary is not well determined. - More in Beamer.

Enough about the signal zone (Let's talk about noise)

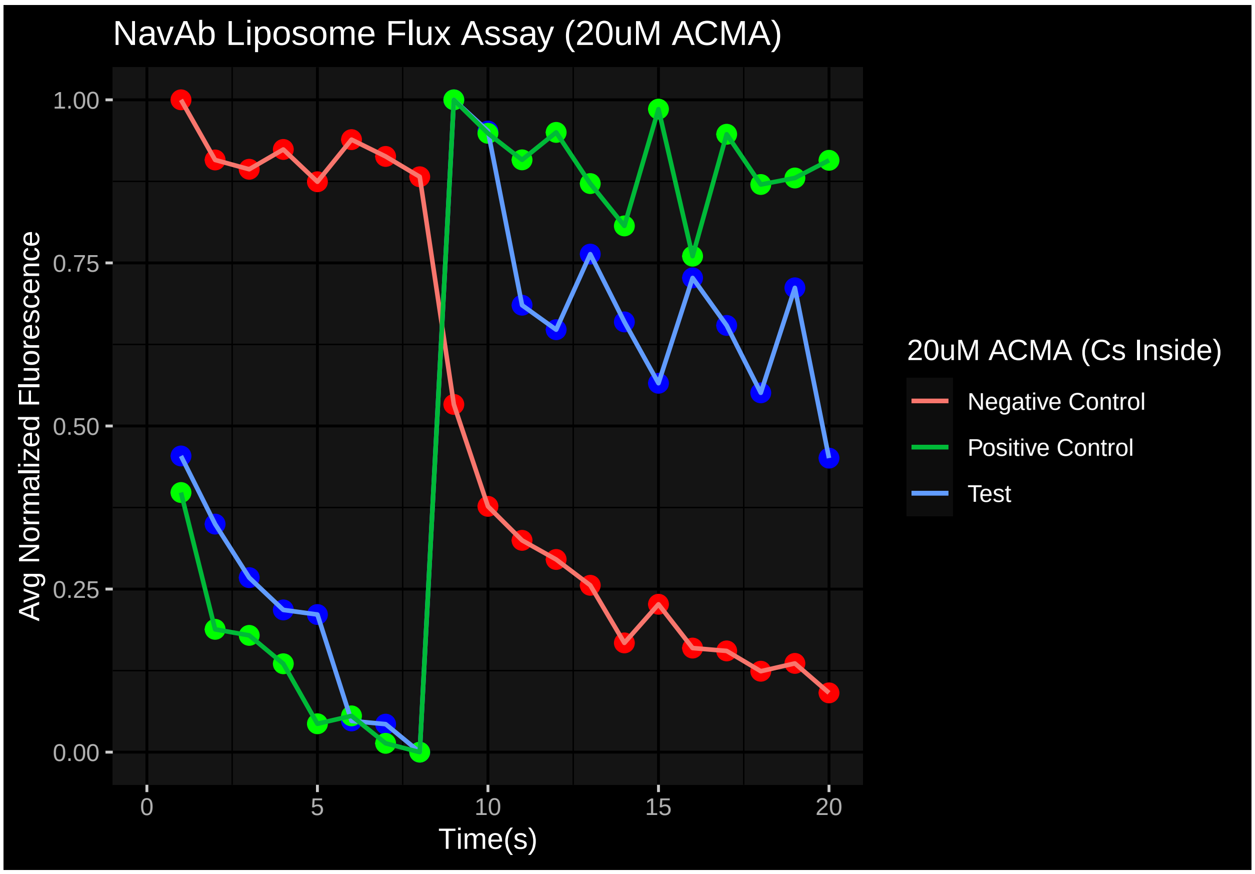
- Out of the signal zone or at least at the boundary of the signal zone (>=) I notice an inversion of some parts of our results from the signal zone and in other cases, extreme unpredictability (especially the lower boundary).
- Upper boundary of signal zone towards the noise zone

```
#plotting in the noise zone at >=(20uM-inf) with Cs inside
p3 = ggplot(navab227_norm2, aes(x=WELLNUM)) +
  geom_point(aes(y=E20norm), size=3, color="blue") +
  geom_point(aes(y=F20norm), size=3, color="red") +
  geom_point(aes(y=G20norm), size=3, color="green") +
  geom_line(aes(y=E20norm, color='Test'), size=0.8) +
  geom_line(aes(y=F20norm, color='Negative Control'), size =0.8) +
  geom_line(aes(y=G20norm, color='Positive Control'), size =0.8) +
  dark_mode()+
  xlim(0,20) +
  ylim(0,1) +
  labs(title = 'NavAb Liposome Flux Assay (20uM ACMA)',
       x = 'Time(s)', y='Avg Normalized Fluorescence', color='20uM ACMA (Cs Inside)')
```

```r
# call plot
p3 #or print(p1)
```

```
## Warning: Removed 20 rows containing missing values (geom_point).
## Removed 20 rows containing missing values (geom_point).
## Removed 20 rows containing missing values (geom_point).

## Warning: Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
```



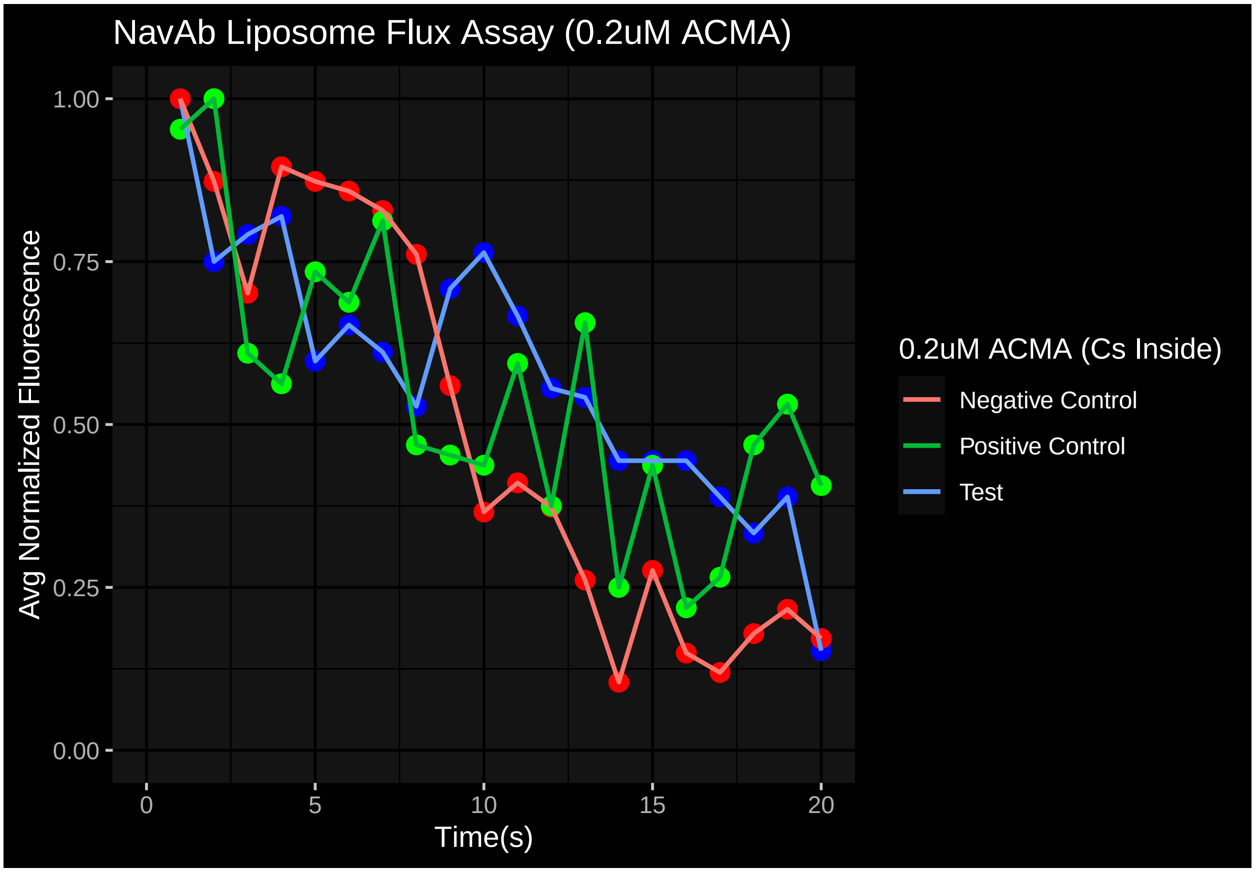
- Lower boundary of signal zone towards the noise zone

```r
#plotting in the noise zone at <=(1-2) with Cs inside
p4 = ggplot(navab227_norm2, aes(x=WELLNUM)) +
  geom_point(aes(y=E0norm), size=3, color="blue") +
  geom_point(aes(y=F0norm), size=3, color="red") +
  geom_point(aes(y=G0norm), size=3, color="green") +
  geom_line(aes(y=E0norm, color='Test'), size=0.8) +
  geom_line(aes(y=F0norm, color='Negative Control'), size =0.8) +
  geom_line(aes(y=G0norm, color='Positive Control'), size =0.8) +
  dark_mode()+
  xlim(0,20) +
  ylim(0,1) +
  labs(title = 'NavAb Liposome Flux Assay (0.2uM ACMA)',
       x = 'Time(s)', y='Avg Normalized Fluorescence', color='0.2uM ACMA (Cs Inside)')
```

```r
# call plot
p4 #or print(p1)
```

```
## Warning: Removed 20 rows containing missing values (geom_point).
## Removed 20 rows containing missing values (geom_point).
## Removed 20 rows containing missing values (geom_point).

## Warning: Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
```



ACMA concentration within the signal zone is predictable

- Within the signal zone varying ACMA concentration drives similar behavior in the system.
- We could expect differences but that is not the reality

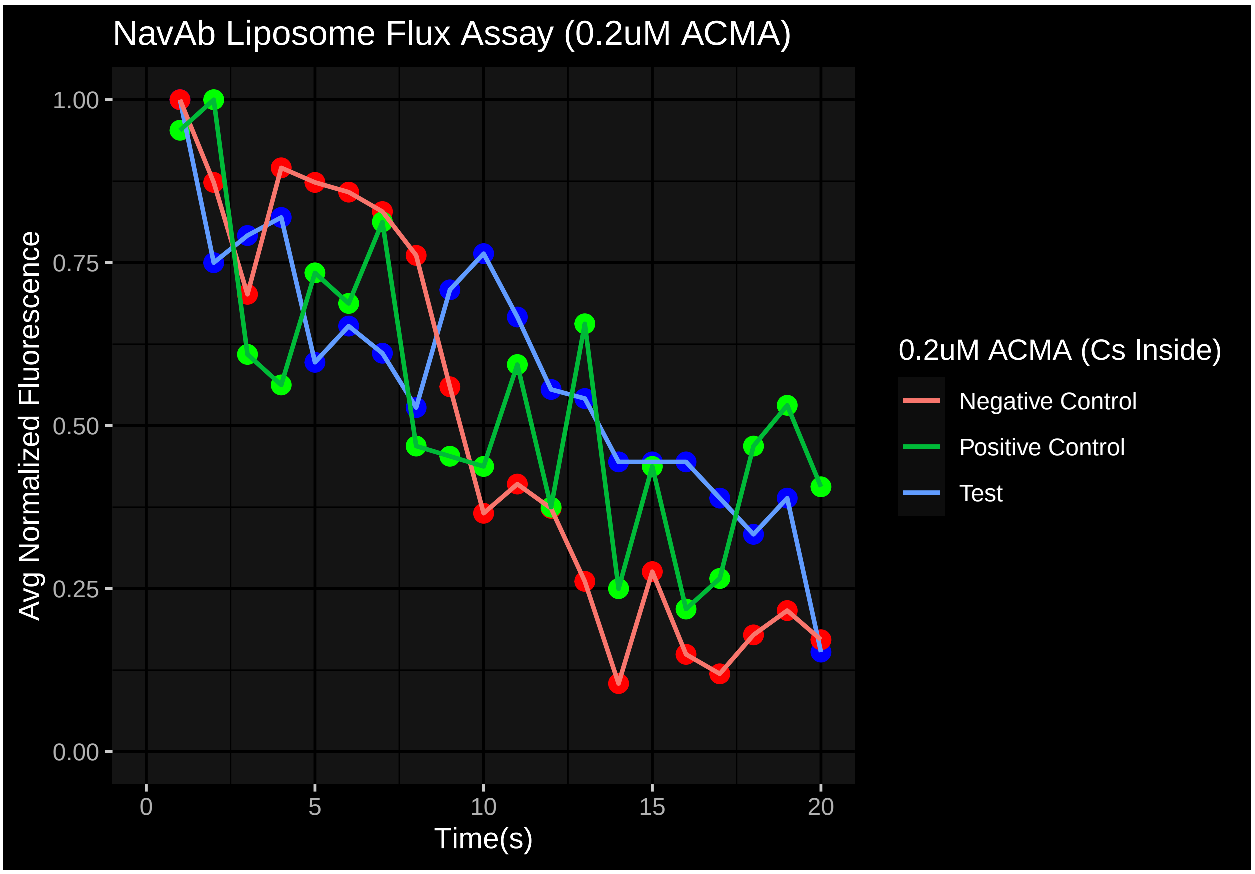
```r
#plotting in the noise zone at <=(1-2) with Cs inside
p4 = ggplot(navab227_norm2, aes(x=WELLNUM)) +
  geom_point(aes(y=E0norm), size=3, color="blue") +
  geom_point(aes(y=F0norm), size=3, color="red") +
  geom_point(aes(y=G0norm), size=3, color="green") +
  geom_line(aes(y=E0norm, color='Test'), size=0.8) +
  geom_line(aes(y=F0norm, color='Negative Control'), size =0.8) +
  geom_line(aes(y=G0norm, color='Positive Control'), size =0.8) +
  dark_mode()+
  xlim(0,20) +
  ylim(0,1) +
```

```
  labs(title = 'NavAb Liposome Flux Assay (0.2uM ACMA)',
       x = 'Time(s)', y='Avg Normalized Fluorescence', color='0.2uM ACMA (Cs Inside)')

# call plot
p4 #or print(p1)
```

```
## Warning: Removed 20 rows containing missing values (geom_point).
## Removed 20 rows containing missing values (geom_point).
## Removed 20 rows containing missing values (geom_point).

## Warning: Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
```




```
# The driver of control point of this signal noise zones (ACMA)
# concentrations of ACMA drive signal-noise zones advocating balance

p5 = ggplot(navab221_avgnorm, aes(x=WELLNUM)) +
  geom_point(aes(y=T2norm), size=3, color="blue1") +
  geom_point(aes(y=T3norm), size=3, color="blue2") +
  geom_point(aes(y=T4norm), size=3, color="blue3") +
  geom_point(aes(y=T5norm), size=3, color="blue4") +

  geom_line(aes(y=T2norm, color='2uM ACMA'), size=0.8) +
  geom_line(aes(y=T3norm, color='3uM ACMA'), size =0.8) +
  geom_line(aes(y=T4norm, color='4uM ACMA'), size =0.8) +
  geom_line(aes(y=T5norm, color='5uM ACMA'), size =0.8) +
```

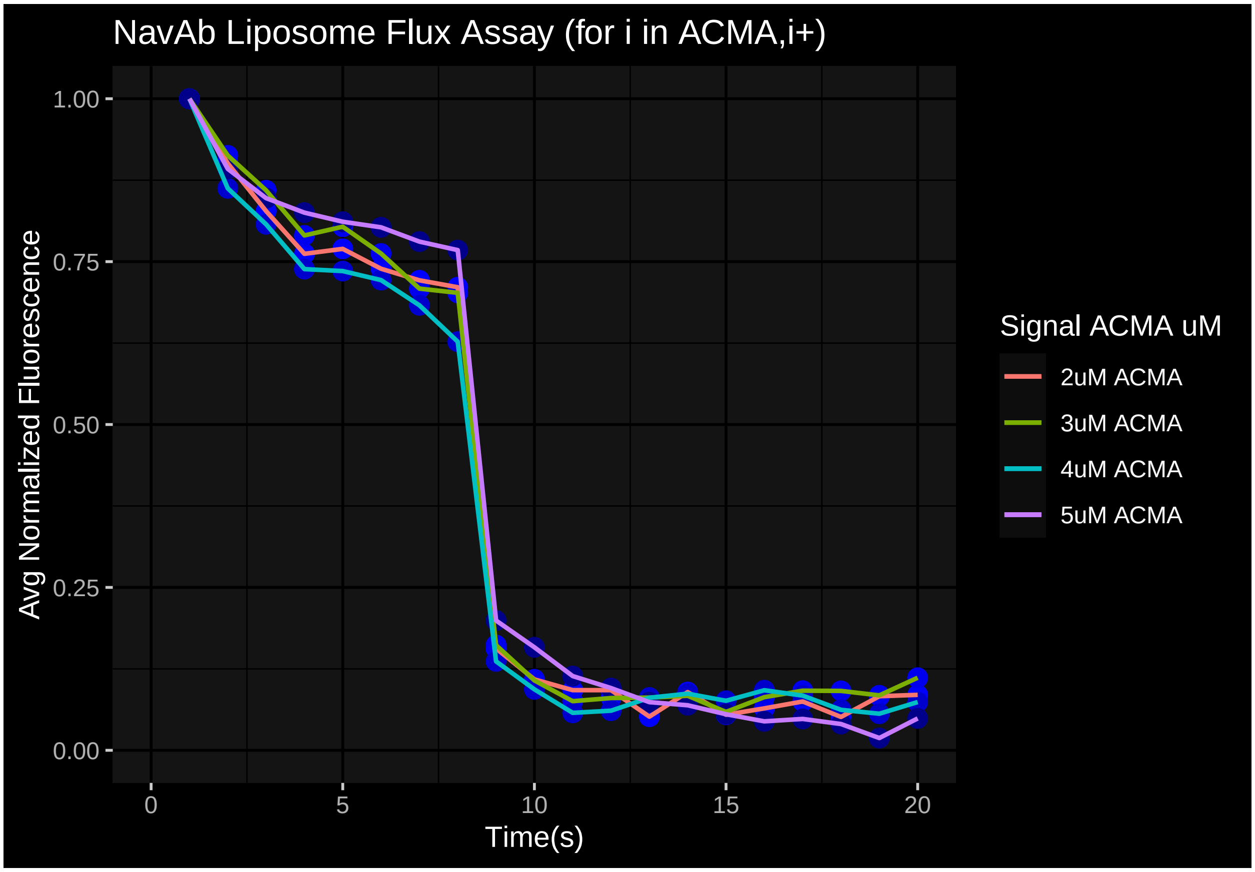
```
  dark_mode()+
  xlim(0,20) +
  ylim(0,1) +
  labs(title = 'NavAb Liposome Flux Assay (for i in ACMA,i+)',
       x = 'Time(s)', y='Avg Normalized Fluorescence', color='Signal ACMA uM')

# call plot
p5 #or print(p1)
```

```
## Warning: Removed 20 rows containing missing values (geom_point).

## Warning: Removed 20 rows containing missing values (geom_point).
## Removed 20 rows containing missing values (geom_point).
## Removed 20 rows containing missing values (geom_point).

## Warning: Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
## Removed 20 row(s) containing missing values (geom_path).
```



Takebacks

- Very complex and cerebral project
- Signal-Noise ratios are important to know when understanding a system
- A failure to know this prevents good interpretation of results
- Final spoiler, my opinion is the signal ration is 2<= x >=20

- I am more certain of the upper boundary than I am of the lower and now that I say that even the upper boundary is not well known.
- This is very rudimentary coding and I am curious to see a more sophisticated approach from anyone
- Also the next step is automation of DBF to CVS conversion using R shiny apps.
- I am very proud of this project and check out my beamer.