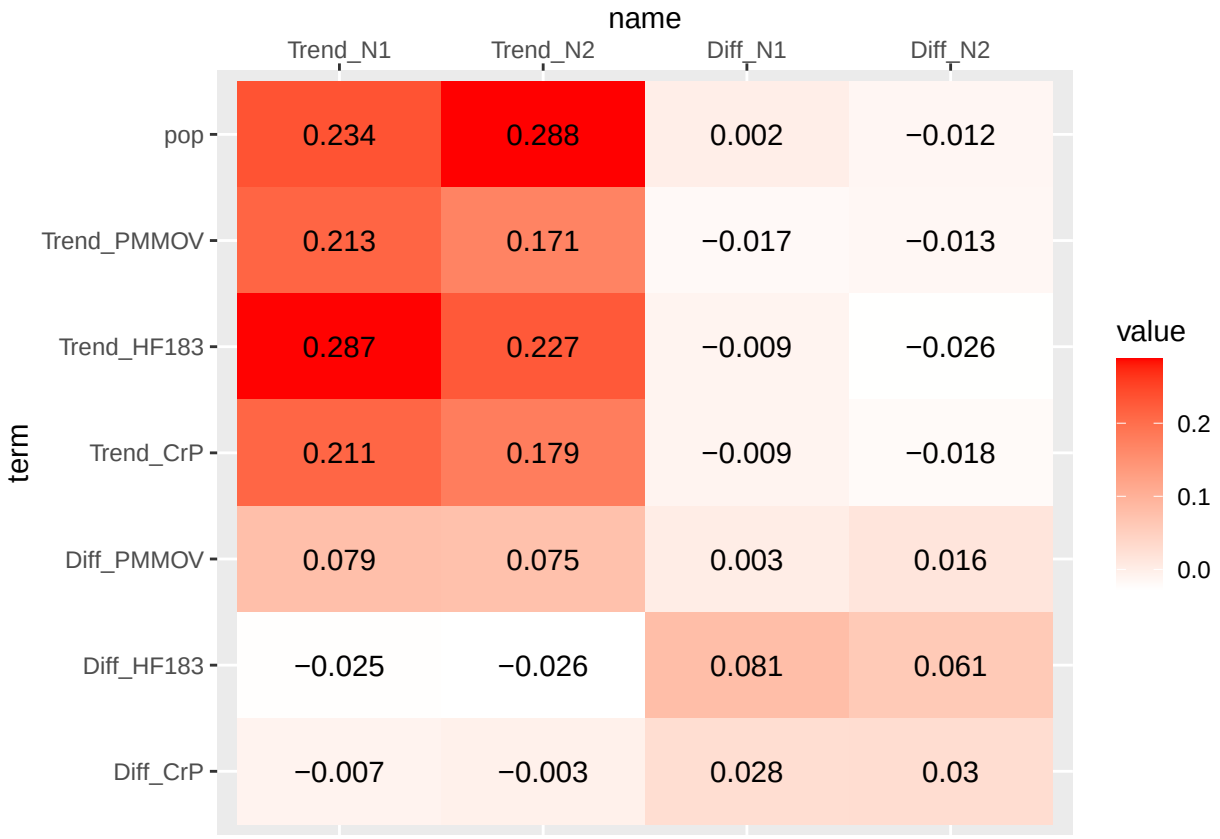
Covid-19 Wastewater Flagging Method Optimization Using a Minimum Flagging Difference of 10%

Marlin Lee, Kyllan Wunder, Abe Megahed

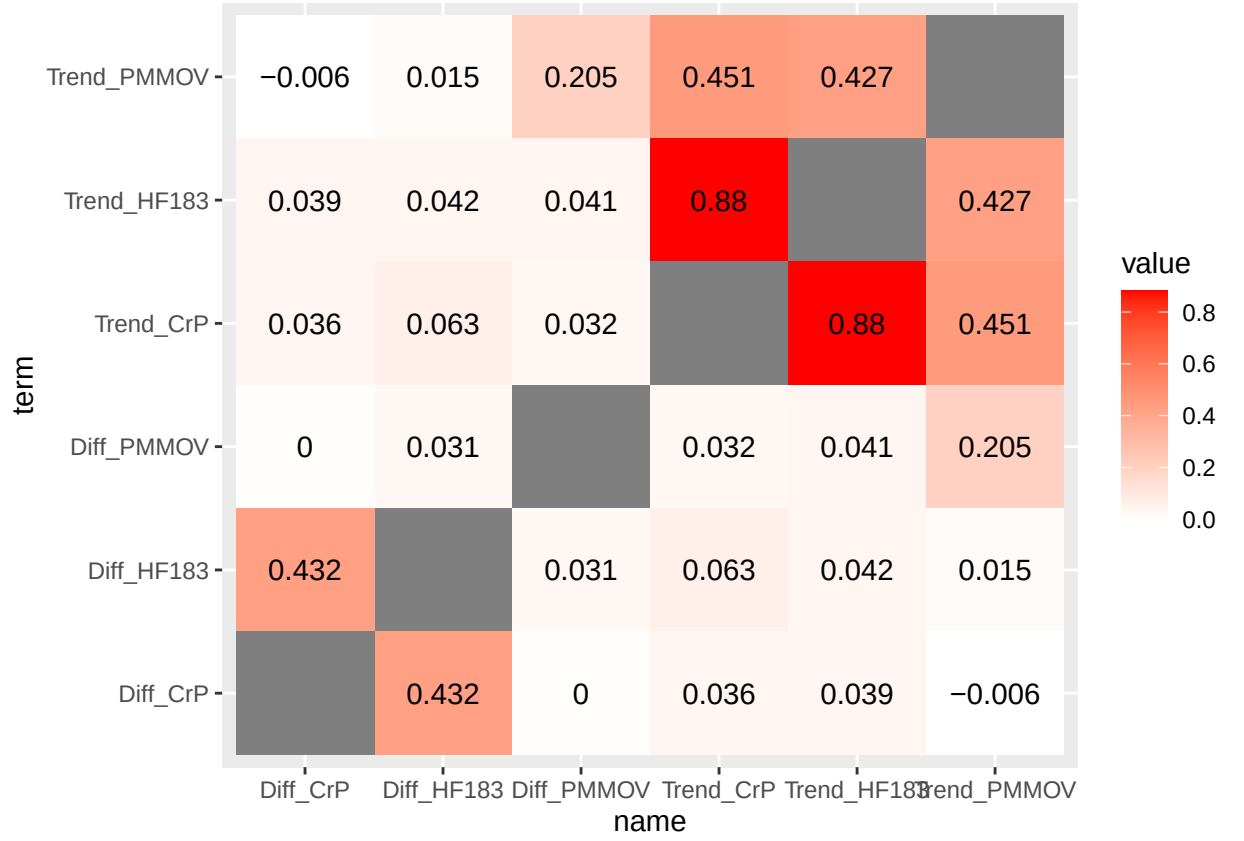
University of Wisconsin Data Science Institute - October 2022

using HFD to get a good understanding of the relationships between the Covid-19 measurements and the other covariates. we did this by breaking the signal down into two components. (1) the trend of the data. This is created using a loess smoothing. (2) the detrended data. We do this on the log of the data because the noise appears to be log normal meaning a more clear trend. The main goal of this document is to show the main takeaways of those plots in an easier way to see. The main takeaways are, 1) the covariates don't relate to the covid signal any more than the population does. 2) The detrended data seems to be largely independent

The heatmap below shows the correlation between the Covid-19 signals and the covariate components. The baseline these covariates are meant to control for is population so it acts as the baseline. We see the data it has a .25 correlation which is mildly meaningful. This is around the correlation the other covariants have. None of the detrended data correlates meaningful supporting the view that they are mostly caused by noise.



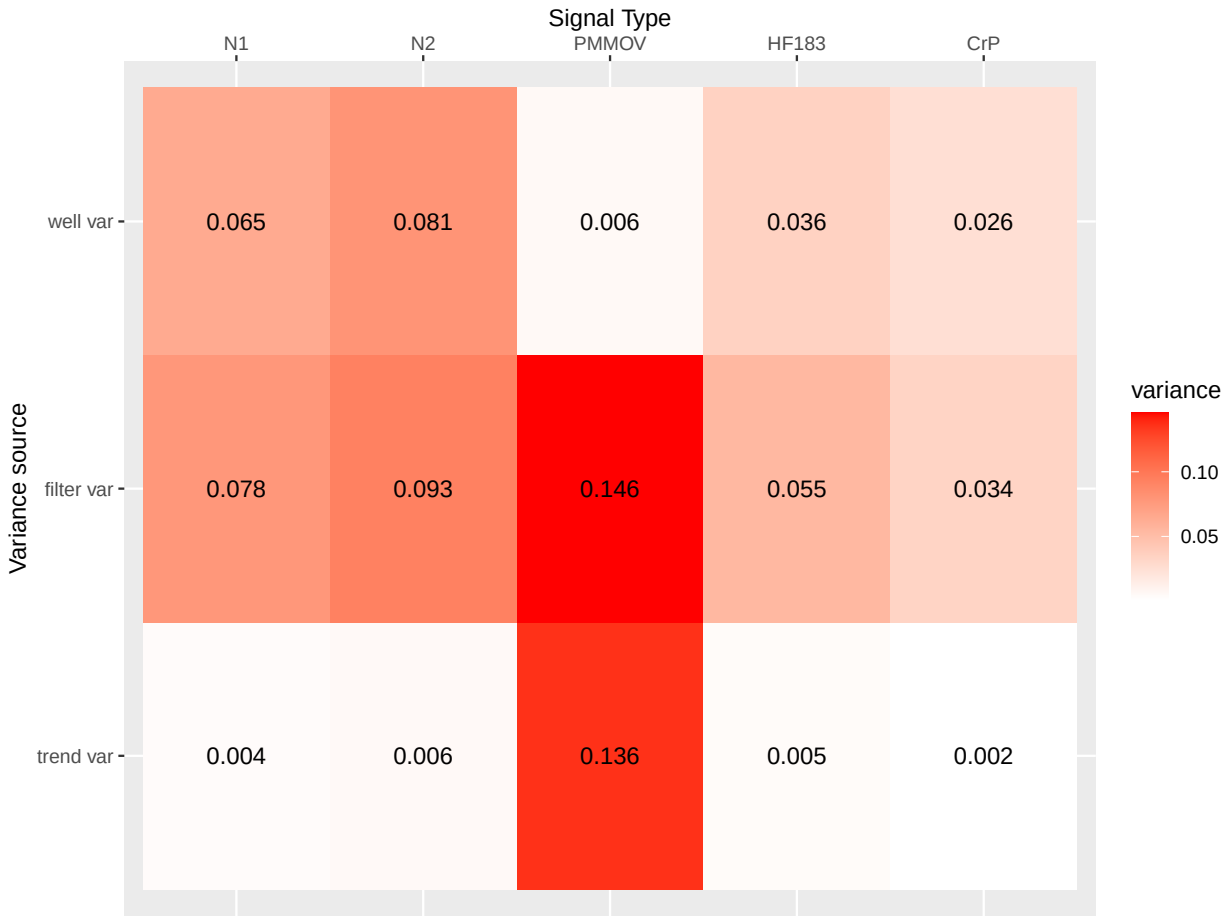
looking at the relationship between HF183 and CrP you see a distinctly strong relationship in both the trend and the noise. We don't have enough understanding of the underlying process to know if this is expected.



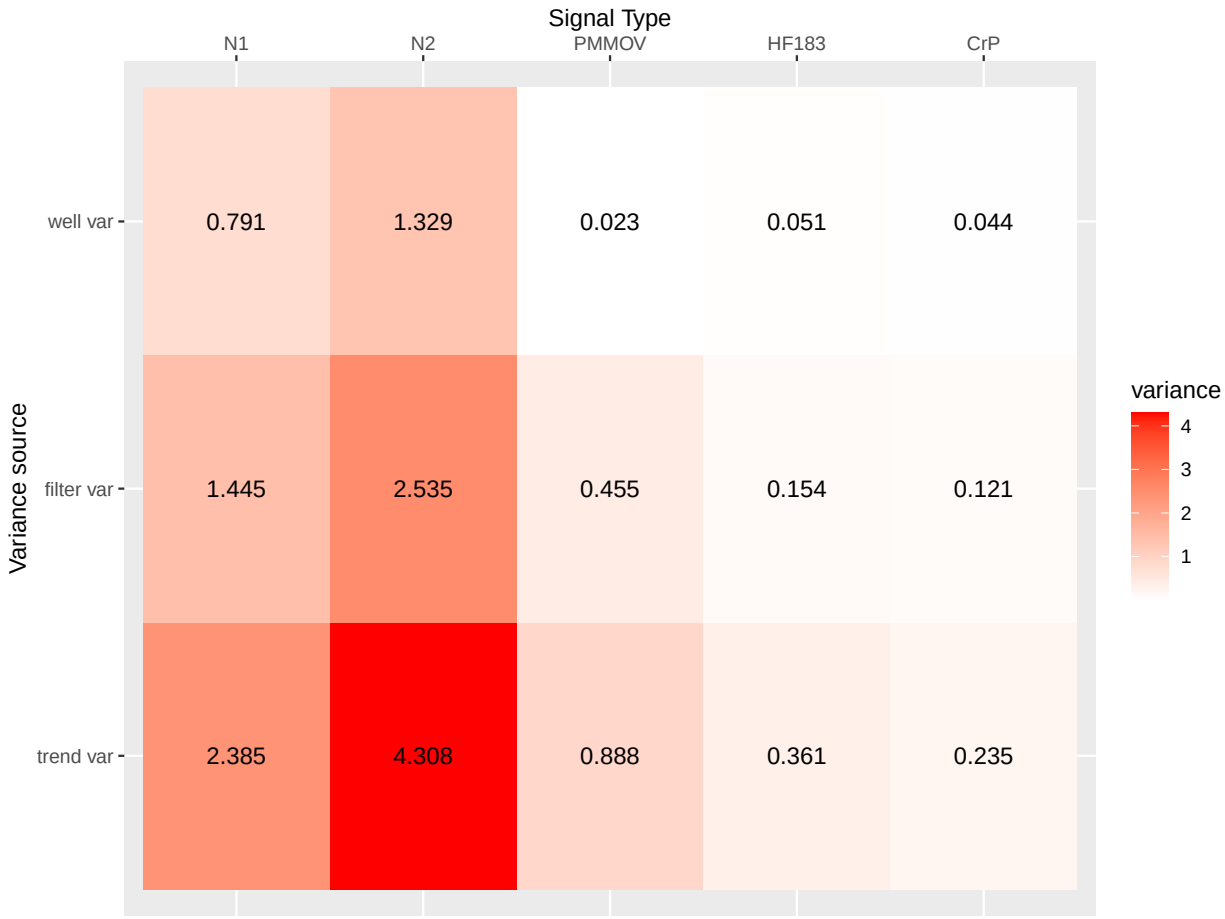
For a baseline we showed the relationship between N1 and N2 that show an extremely strong correlation clearly due to them measuring the same thing.

```
##                    log      base
## trend corr   0.9268483 0.9824389
## detrend corr 0.3407199 0.3446168
```

We also want to look at the variance of these processes. We can capture variance at three steps of the process. At the lowest level we have variance in the technical replicates. Next we have variance at the well replicates. finally we have variance from the detrended data level. each level contains the lower levels variance. bellow is a table showing each signals three level of variances.


Signal Type
N1
N2
PMMOV
HF183
CrP
Variance source
well var
filter var
trend var
0.065
0.081
0.006
0.036
0.026
0.078
0.093
0.146
0.055
0.034
0.004
0.006
0.136
0.005
0.002
variance
0.10
0.05

```
## [1] 0.077241
```


Signal Type
N1
N2
PMMOV
HF183
CrP
Variance source
well var
filter var
trend var
0.791
1.329
0.023
0.051
0.044
1.445
2.535
0.455
0.154
0.121
2.385
4.308
0.888
0.361
0.235
variance
4
3
2
1