

# Covid-19 Wastewater Data Downsampling

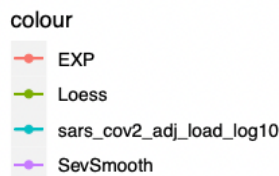
Marlin Lee, Steve Goldstein, Kyllan Wunder, Abe Megahed  
University of Wisconsin Data Science Institute - July, 2022

In this analysis, we look at using downsampling to simulate results that would be obtained using a lower wastewater sampling rate. Wastewater facilities in different communities sample and analyze water at different frequencies and it is not always feasible to sample wastewater at a high frequency. Some facilities sample wastewater each day while others may sample only once or twice per week.

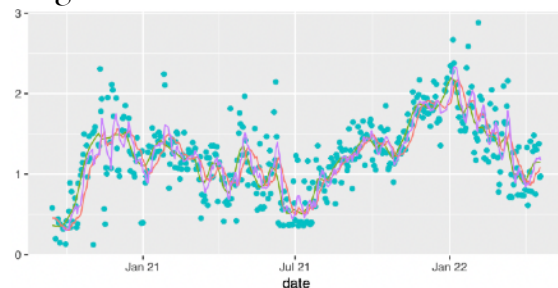
This analysis seeks to answer whether the DHS method is robust when reducing the number of measurements. We do this by running the analysis with the different smoothing on both the original Madison data and the data downsampled to two days per week.

## The Data

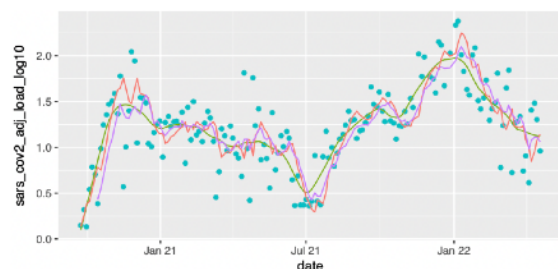
The graphs to the right display the original data and the data downsampled to 2 days per week. The major features of the data is preserved in the 2 day sampling.



Original



Downsampled to 2 Days / Week:

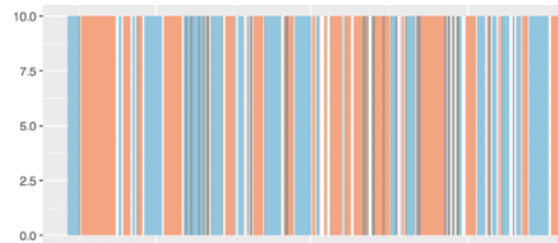


## Exponential Smoothing

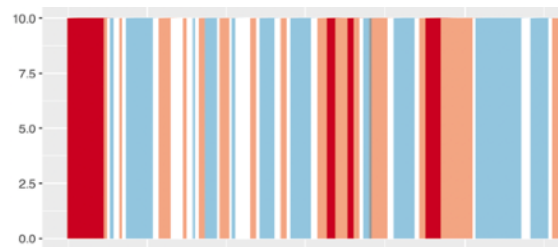
The 2 day downsampling captures the three main regions of moderate increase, one to the far left and two at the middle right as well as a smaller region of moderate increase to the middle left and one to the far right.

The two day sampling captures a conspicuous set of regions of major increase (in red) that are lacking in the the original data. This is because the major increase regions are denoted by slope and by sampling the data at a smaller rate, we increase the slope between data points. If we applied a normalizing factor to account for the higher slope due to the lower sampling rate, then the flagged regions would more closely match.

Original



Downsampled to 2 Days / Week:

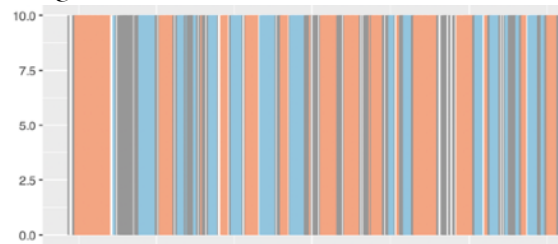


## Loess Smoothing

The Loess smoothing captures the three main regions of moderate increase similar to the exponential smoothing, one to the far left and two at the middle right. More regions are tagged as fluctuating than in the exponential smoothing.

The two day sampling preserves the main regions of moderate increase and smooths over the fluctuations to yield broader regions of moderate decrease in between the main increasing regions. As with the exponential smoothing, the region to the far right is conspicuously highlighted as a major increase.

Original



Downsampled to 2 Days / Week:

