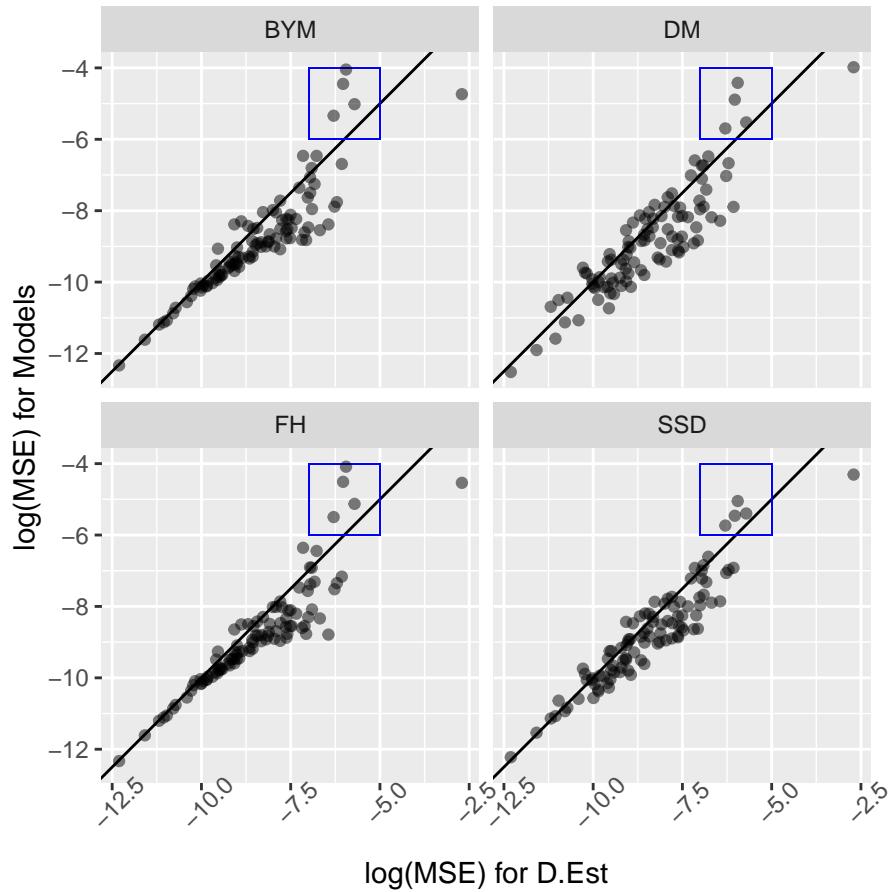


# NC Simulation Study: MSE Analysis by Area

Sho Kawano | September 5, 2024

## MSE from Direct Estimates vs. Models by Area



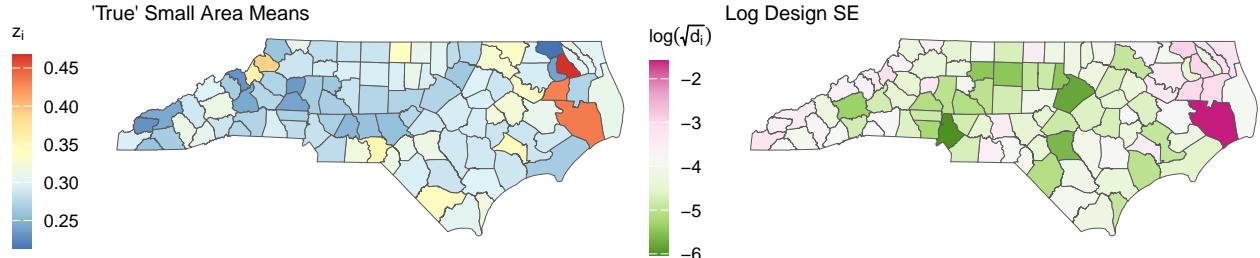
Note that the points shown are *averages by area for the 300 simulations*. If a model outperformed the direct estimate for a given area, a point would lie *below* the line. We can see that the SSD model achieves a MSE reductions across the board.

model	Min	Q1	Mean	Median	Q3	Max
SSD	-0.0038172	-5.60e-06	0.0005883	2.87e-05	0.0001773	0.0534502
DM	-0.0094423	-1.07e-05	0.0004478	2.50e-05	0.0001916	0.0482715
FH	-0.0141873	3.60e-06	0.0004389	3.49e-05	0.0002069	0.0562298
BYM	-0.0148831	1.70e-06	0.0004327	3.79e-05	0.0002156	0.0581630

There are a few areas where using a model increases MSE (enclosed in rectangle) relative to the direct estimate. The increase in MSE is *lower* for SSD model than all other models. This is a major factor in

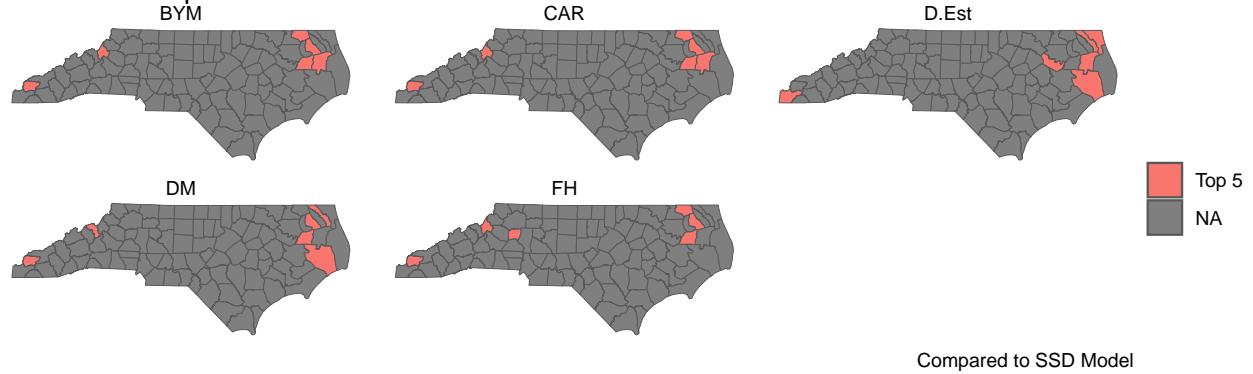
why the SSD model outperforms others. This can also be seen from the summary statistics of the MSE reduction relative to the direct estimate (on original scale) shown above:  $MSE - MSE_{D.Est}$ .

## Spatial Distribution of MSE Decreases/Increases Relative to SSD Model



First is a map of the ‘true’ small area means in this simulation study and the log design SEs. We can see that some counties along the western edge and northeastern corners of the state have the highest/lowest values as well as high design SE’s. Intuitively, the top 5 counties in terms of MSE reductions (relative to the SSD model) are mostly located in these places.

### Locations of Top 5 MSE Reduction Counties



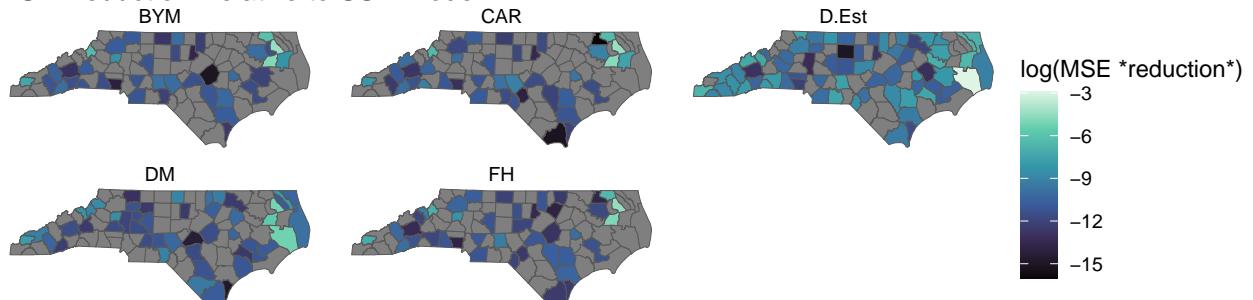
### Where did SSD Model decrease MSE?

We can see the spatial distribution of MSE reduction relative to SSD Model below. The *lighter* the color, the higher the MSE reduction is and vice versa.

Note on log MSE reductions:

- $-3$  log MSE reduction is approximately  $\approx 0.223^2$  on the original scale
- $-6$  log MSE reduction is approximately  $\approx 0.05^2$  on the original scale
- $-9$  log MSE reduction is approximately  $\approx 0.01^2$  on the original scale
- $-12$  log MSE reduction is approximately  $\approx 0.0025^2$  on the original scale

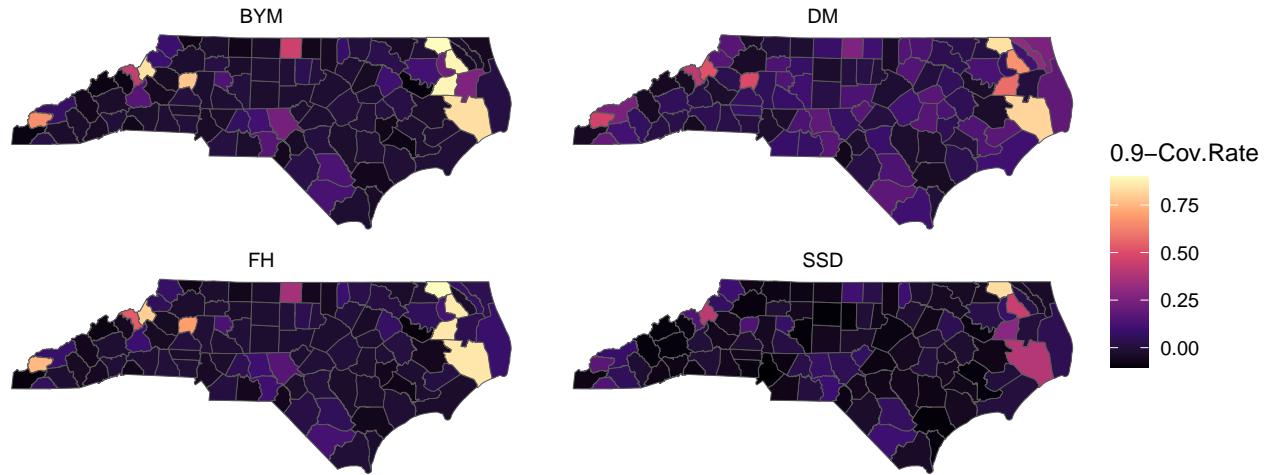
### MSE Reduction Relative to SSD Model



We can see that the highest MSE reductions by the SSD are achieved in the northeastern corner of the state. These are the same counties where the models generally struggle with compared to the direct estimate. This echoes what we saw in the first plot. Outside of the high reduction counties, the patterns differ a bit but there are still some similarities.

## Coverage Analysis

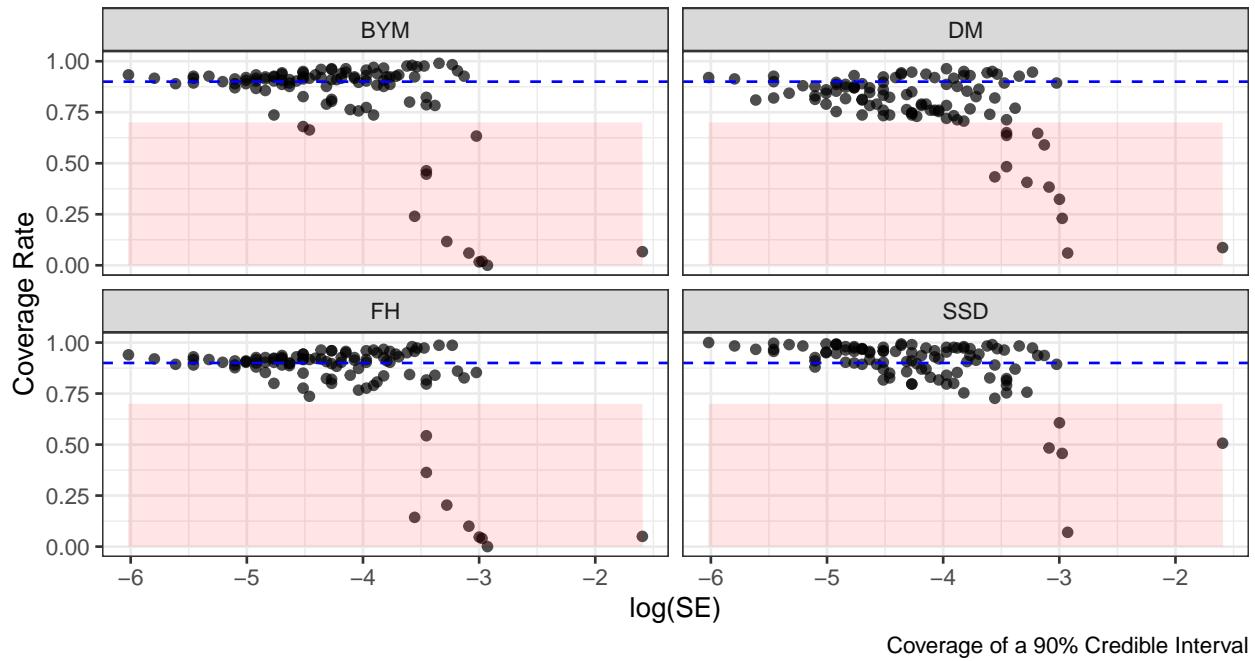
**Undercoverage of a 90% credible interval by model**



Here we plot the undercoverage of the average coverage rate in each area. We define undercoverage as  $0.9 - \text{Avg. Coverage Rate}$ . If a model, on average achieves 90% coverage in a given area, the value should be zero.

We can see from this plot that undercoverage is very high in the same counties (located in the northeastern and western edges of the state) that had accuracy issues. The models without random effects selection (BYM and FH models) struggle mightily in these areas. The DM model struggles a bit less than the other models in these areas but has higher undercoverage in other parts of the state. This can also be seen in the plot below.

## Log Design SEs vs. Coverage by model



Here is a plot of the log design SEs compared to the coverage rate of a 90% credible interval from each model. The dotted line is 90% (the target coverage rate), and the red area indicates coverage below 70%. We can see that the coverage rate for many models start to drop as the SE gets larger. There is a clear difference when you compare the SSD to the other models.

Below is a table with the minimum, 5th, 10th, and 25th percentile values of the coverage rate (rounded to the 4th digit). In fact, we can see that the BYM and the FH have horrible coverage for the lowest 5% of counties but has better coverage for the 25th percentile than the DM. This is what results in the DM having very low coverage overall. Again, the SSD really outperforms the other models.

model	Min	5th Percentile	10th Percentile	25th Percentile	Mean
BYM	0.00	0.1142	0.6603	0.8492	0.8281
DM	0.06	0.4055	0.6457	0.7517	0.7894
FH	0.00	0.1412	0.7637	0.8425	0.8334
SSD	0.07	0.7207	0.7960	0.8667	0.8941

## Summary

### MSE

- There are areas in the eastern parts of the state where models struggle in terms of accuracy where design SE is large. The SSD performs the best in these areas compared to the other models.
- Reductions over the direct estimate are achieved in most counties
- Reductions over the other model estimates are concentrated in a few counties where high and small area means are adjacent to each other. This is seen in the northeastern/eastern part of the state and some counties along the western edge of the state.

### Coverage

- The same areas that have accuracy issues also cause coverage issues for the other models

- BYM and the FH have horrible coverage for the lowest 5% of counties (compared to the DM & SSD) but has better coverage for the 25th percentile than the DM.