COVID-19 Case Forecast Model Evaluation

Between Specific Model Types

Richa Jain, July 2021

Agenda

INTRODUCTION
- Hypothesis & what does this project entail

METHODS
- Step by Step details

RESULTS
- What outputs were produced

DISCUSSION
- What did I find

Introduction

- **Hypothesis:**
  - When there is a sudden change in population mobility patterns, forecasts from models that include mobility data will be more accurate than those from models that do not include such data

Methods

Created a Table of Model Characteristics

- Looked for models that included mobility data
- Determined what data each model used
- Determined the model type

Determined the Inclusion Criteria

- Location
- Time Period
- Target
- Confirm all model information for selected models

Evaluation Graphs & Charts

- Truth data
- Forecasts
- WIS
- Relative WIS

Inclusion Criteria

01	**Location**	• California (06) • New York (36)
02	**Time Period**	• December 5 2020 - December 19 2020 (CA) • December 17 (NY)
03	**Target**	• Incident Cases
04	**Key Metric**	• Relative WIS = ○ average_WIS_mobility/average_WIS_non-mobility • Relative WIS (baseline) = ○ average_WIS_after/average_WIS_before

Models and Model Characteristics

Model	Case Data	Model Type	Social Distancing Assumptions?	Mobility Data?	Notes
COVIDhub-baseline	JHU CSSE	Median prediction at all future horizons	no	no	
LANL-GrowthRate	JHU CSSE	Statistical dynamical growth model	no	no	
COVIDhub-ensemble		Unweighted average or median of submitted forecasts	no	no	
RobertWalraven-ESG	JHU CSSE	SEIR model	no	no	
IowaStateLW-STEM	NYT, Johns Hopkins, Covid Tracking Project, USA Facts	Nonparametric space-time disease transmission model	no	yes	
UVA-Ensemble	CDC	AR, ISTM, SEIR model	no	yes (Baidu)	
JHU_CSSE-DECOM	JHU CSSE	Empirical machine learning model	no	yes (SafeGraph)	

How were these models selected?

1 — 2 — 3

Case Forecast

Determined which models submitted case forecasts

Mobility Data?

Separated models with and without mobility data

Time Period

Narrowed down models that were submitting during selected time

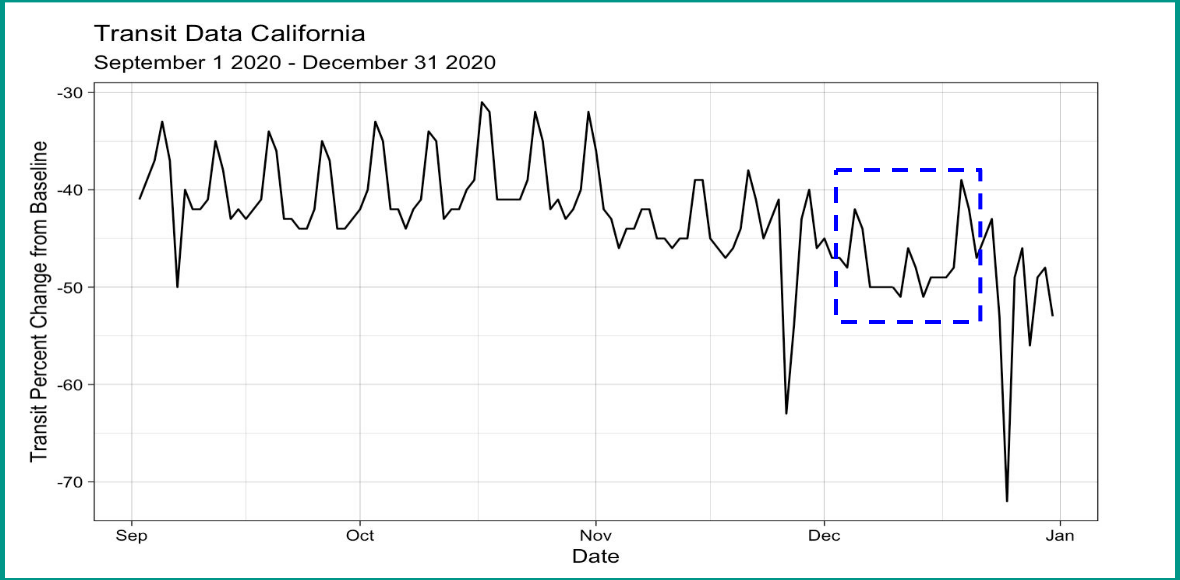
CALIFORNIA

Time Period Selection

California

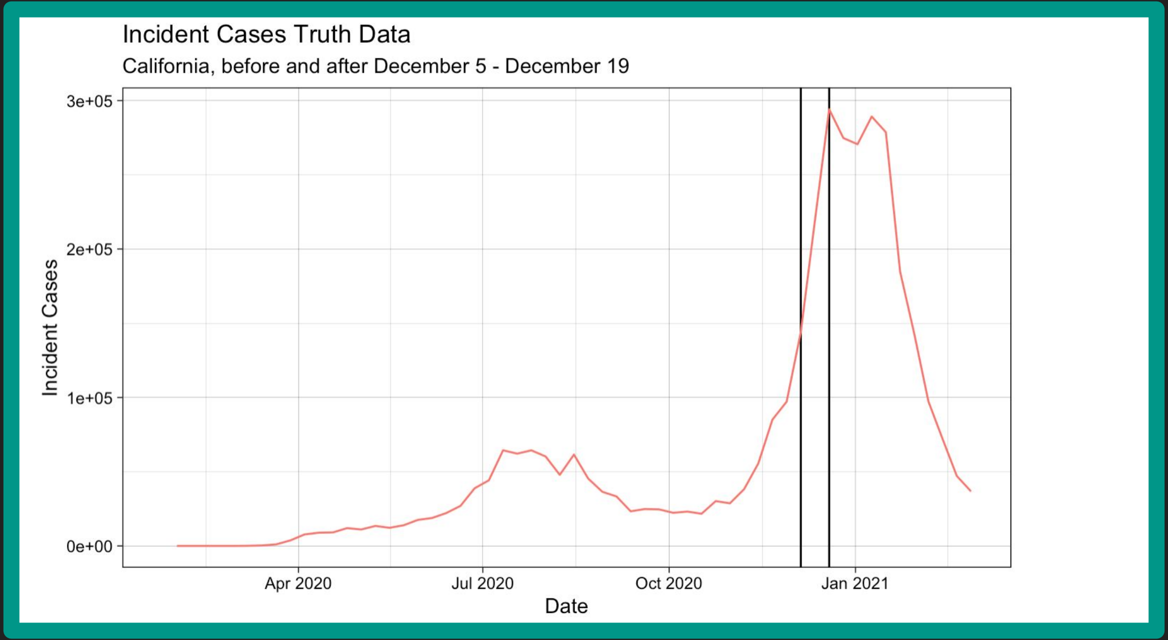
Date: December 5 2020 - December 19 2020

- Transit Mobility Change



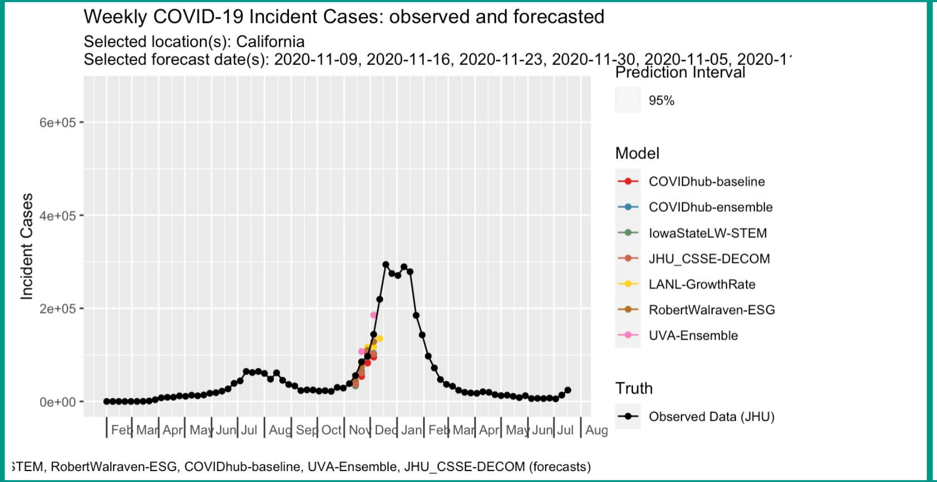
Baseline: median transit mobility value from January 3 - February 6 2020

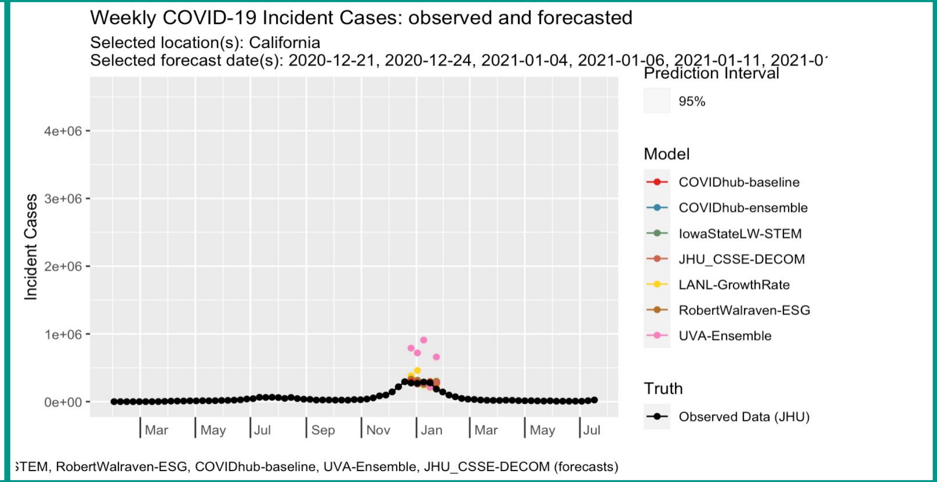
California Truth Data



- Cases increasing before
- Cases decreased after
- Cases increased in between

Forecasts



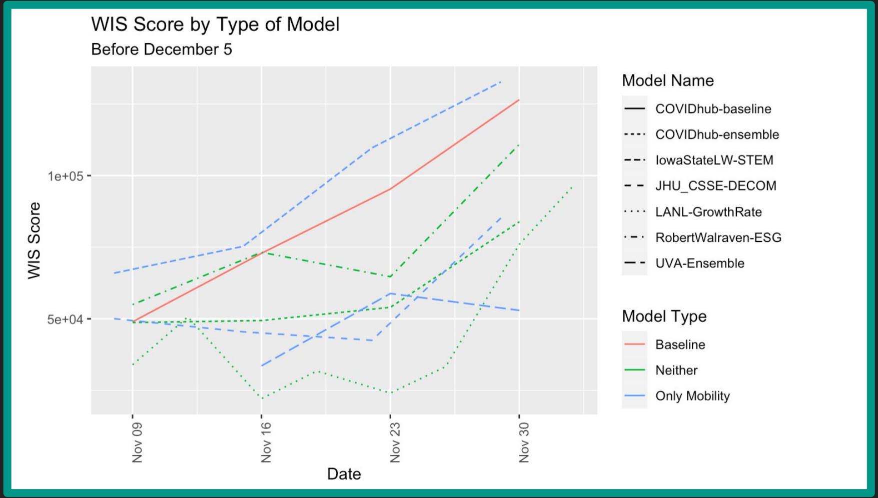


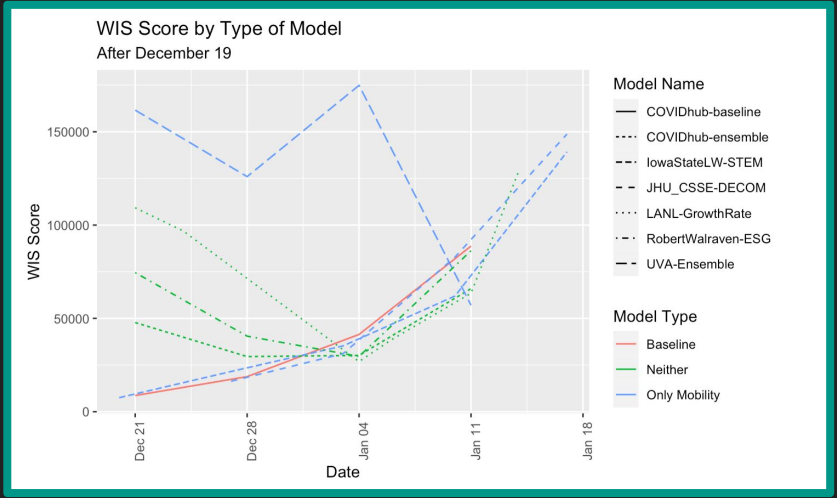
Forecasts 1 month before

Forecasts 1 month after

- Some forecasts seem very off, but this visual is not extremely clear

WIS by Model and Type





- Looked at WIS 1 month before and 1 month after
- Mobility models got worse after
- Models without mobility data had consistent WIS

Results

model	location	wis_before
<chr>	<chr>	<dbl>
COVIDhub-baseline	06	85913.46
COVIDhub-ensemble	06	58995.78
IowaStateLW-STEM	06	95910.37
JHU_CSSE-DECOM	06	55792.20
LANL-GrowthRate	06	46102.31
RobertWalraven-ESG	06	75945.70
UVA-Ensemble	06	48449.70

model	location	wis_after
<chr>	<chr>	<dbl>
COVIDhub-baseline	06	39404.23
COVIDhub-ensemble	06	43353.52
IowaStateLW-STEM	06	53110.55
JHU_CSSE-DECOM	06	69761.35
LANL-GrowthRate	06	77374.51
RobertWalraven-ESG	06	57724.02
UVA-Ensemble	06	129886.03

- Iowa got better but JHU and UVA got worse
- Non-Mobility models got better with the exception of LANL

Results

wis_before_mob <dbl>	wis_before_neither <dbl>
73204.09	52853.59
wis_after_mob <dbl>	**wis_after_neither** <dbl>
93314.89	59625.27

- Average WIS for mobility got much worse
- Average WIS for no mobility got slightly worse

Results

relwis_mob_neither_before <dbl>
1.385035

relwis_mob_neither_after <dbl>
1.565023

relwis_baseline <dbl>
0.4586502

- Relative WIS increased meaning mobility WIS increased
- Baseline performed well

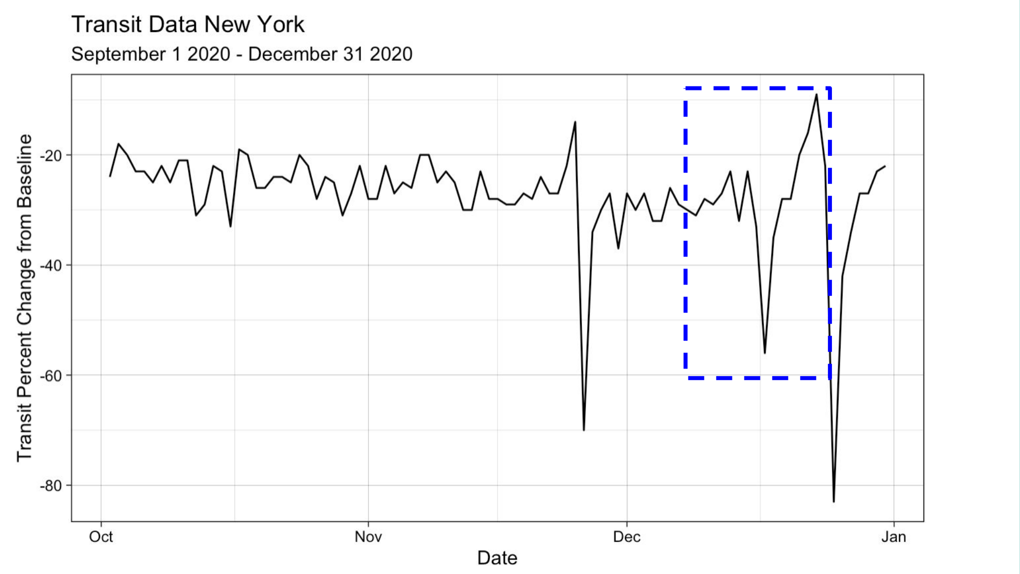
NEW YORK

Time Period Selection

New York

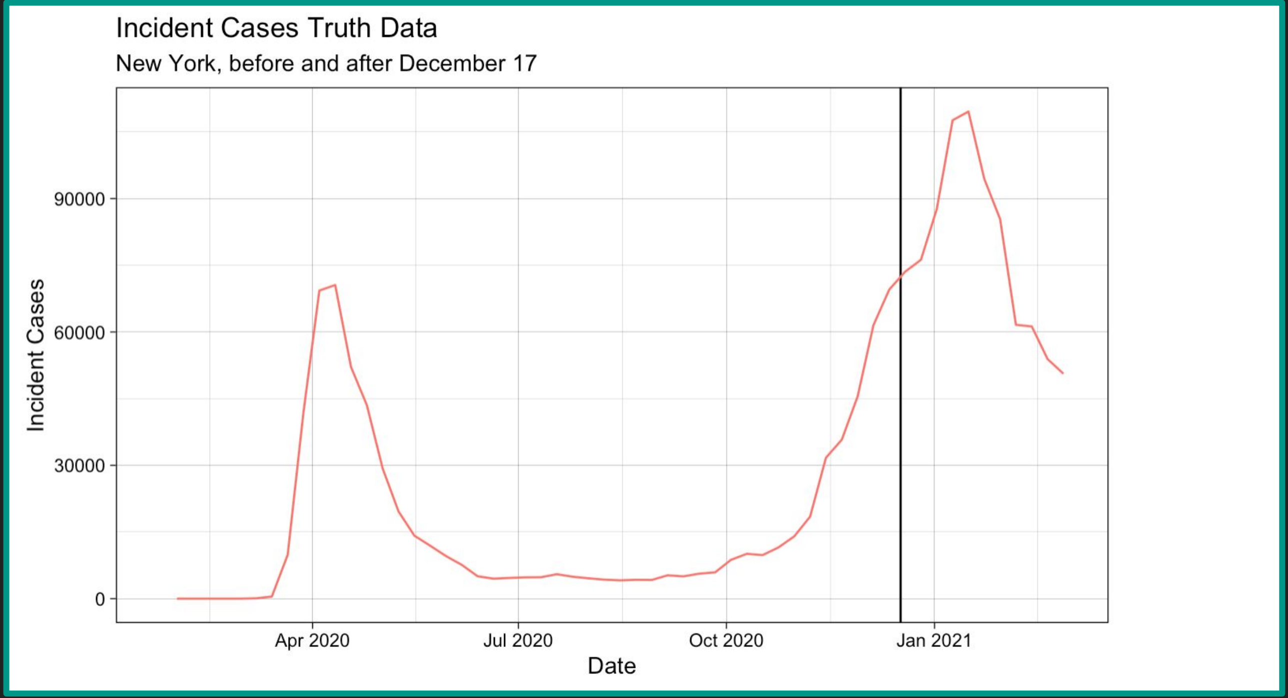
Date: December 17 2020

- Transit Mobility Change



Baseline: median transit mobility value from January 3 - February 6 2020

New York Truth Data



- Cases were increasing before and after

Results

wis_before_mob <dbl>	wis_before_neither <dbl>
11882.57	8234.446

wis_after_mob <dbl>	wis_after_neither <dbl>
19604.35	20117.89

- Average WIS for mobility got slightly worse
- Average WIS for no mobility got much worse

Results

relwis_mob_neither_before <dbl>
1.443032

relwis_mob_neither_after <dbl>
0.9744733

relwis_baseline <dbl>
0.8990317

- Relative WIS decreased meaning mobility WIS decreased and performed better than non-mobility
- Baseline performed well

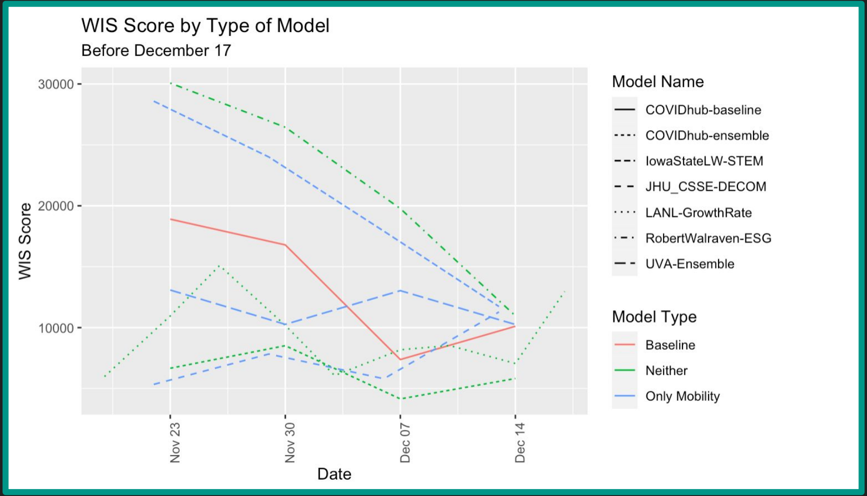
Discussion

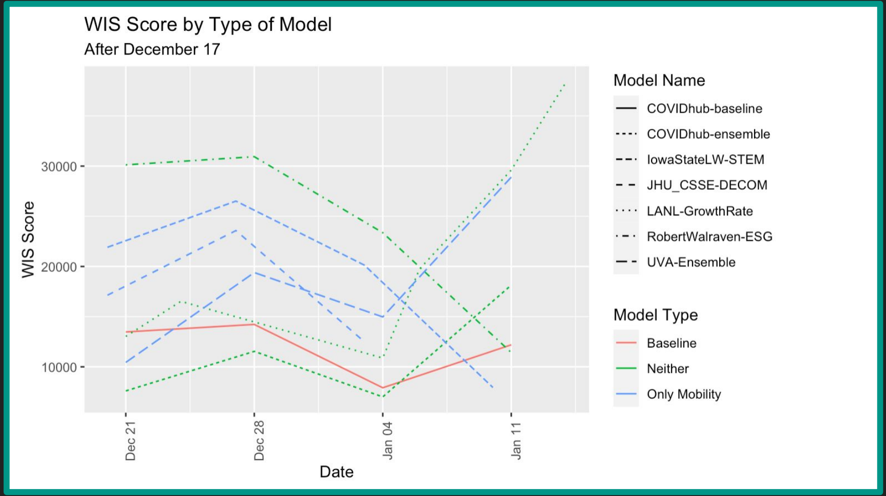
- If models with mobility data performed better, we would have expected the relative WIS after to be less than 1
- In California: the relative WIS after was greater than 1 and greater than the relative WIS before
- In New York: the relative WIS after was less than 1
- Models with mobility data did not perform better than models without mobility data for California, but they did perform better for New York
- 2 Options:
  - Conclude that models with mobility are not more or less likely to perform better
  - OR
  - We need to look into more states and dates to see what the variations are

Appendix

More slides for New York

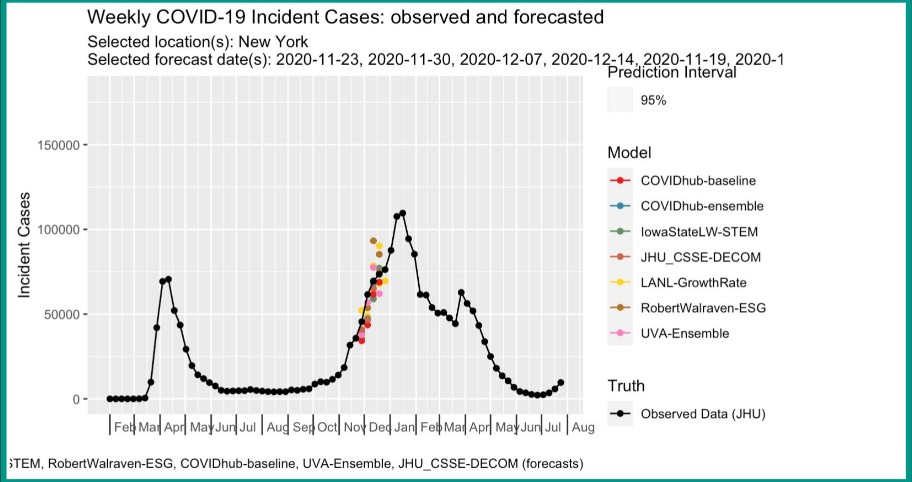
WIS by Model and Type

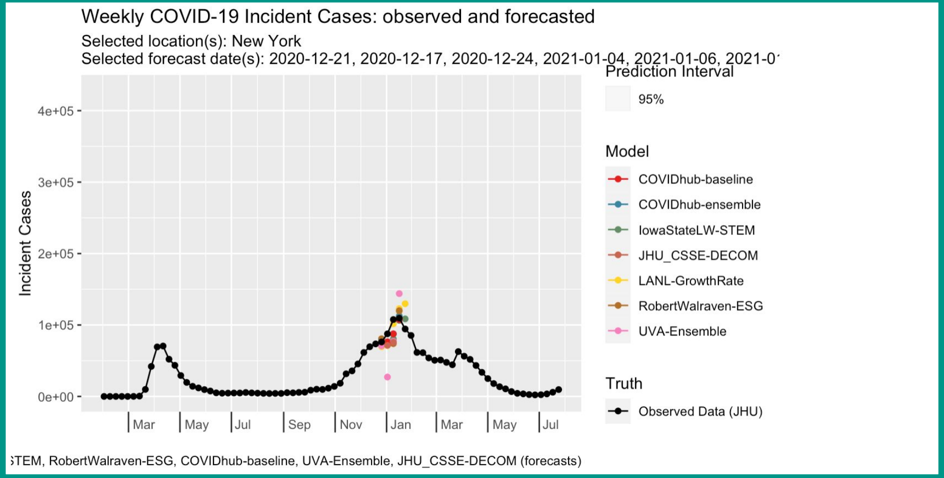




- Looked at WIS 1 month before and 1 month after
- Mobility models got a little better after
  - Below 30000
- Models without mobility data also got a little better
  - LANL got much worse

Forecasts





Forecasts 1 month before

Forecasts 1 month after

- Still relatively close forecasts with a few being off

Results

model	location	wis_before
<chr>	<chr>	<dbl>
COVIDhub-baseline	36	13295.667
COVIDhub-ensemble	36	6288.092
IowaStateLW-STEM	36	20559.537
JHU_CSSE-DECOM	36	7566.242
LANL-GrowthRate	36	9448.235
RobertWalraven-ESG	36	21813.464
UVA-Ensemble	36	11660.976

model	location	wis_after
<chr>	<chr>	<dbl>
COVIDhub-baseline	36	11953.23
COVIDhub-ensemble	36	11072.20
IowaStateLW-STEM	36	19125.26
JHU_CSSE-DECOM	36	17735.72
LANL-GrowthRate	36	21364.05
RobertWalraven-ESG	36	23961.83
UVA-Ensemble	36	18426.98

- Iowa got better but JHU and UVA got worse
- All the non-mobility models got worse