Model Evaluation

Richa Jain and Estee Cramer

Introduction

Many of the models submitting to the COVID-19 Forecast Hub incorporate data about social distancing, mobility, or non-pharmaceutical interventions (NPIs). We look to compare these models with each other and with models that do not incorporate such data and evaluate their performance.

Methods

Step 1: Create Table of Model Characteristics

- Look for which models use social distancing data
- Data used by models (demographic data, hospitalization data)
- Model type (SEIR, Baysian, statistical)
- Look for models that fit criteria but do not include social distancing data

Step 2: Inclusion Criteria

- Locations: Florida, Texas, Massachusetts, New York, California, Michigan
- Time period: November 1 2020 - December 31 2020
- Target: Incident Cases
- Horizons:
- Models:
  - 2 models with both mobility and social distancing
    - CU-select, OliverWynman-Navigator, JHUAPL-Bucky
  - 2 models with neither mobility nor social distancing
    - COVIDhub-baseline, LANL-GrowthRate
  - 2 models with only mobility
    - IowaStateLW-STEM, UVA-Ensemble
  - 1 model with only social distancing +JHU_IDD-CovidSP

Step 3: Evaluation graphs

- After deciding on inclusion criteria, use covidhubUtils to score forecasts and determine which models are best.

Step 4: Linear regression

- Exposure / variable : social distancing, data sources, model type