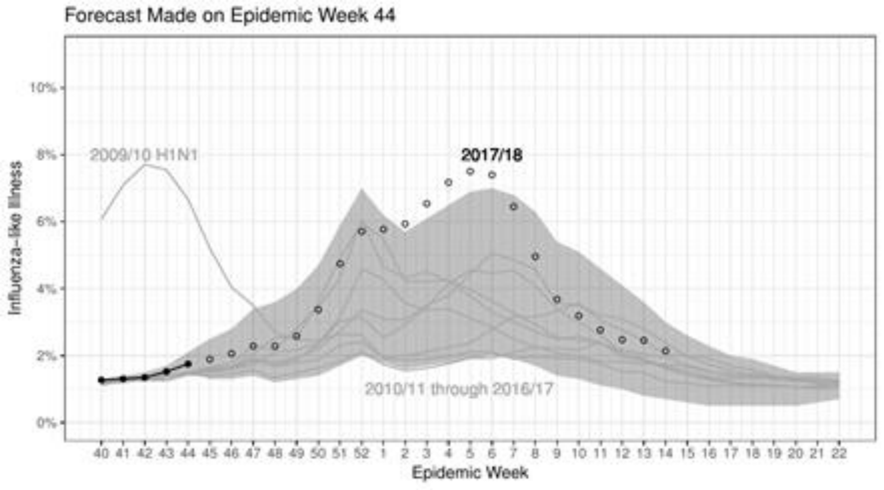
EpiFFORMA: Ensemble Weighting for Forecasting Emerging Epidemic Time Series without Historical Data

Alexander C. Murph, Lauren Beesley, Lauren Castro, Casey Gibson, Sara del Valle, Dave Osthus

January 8th, 2025

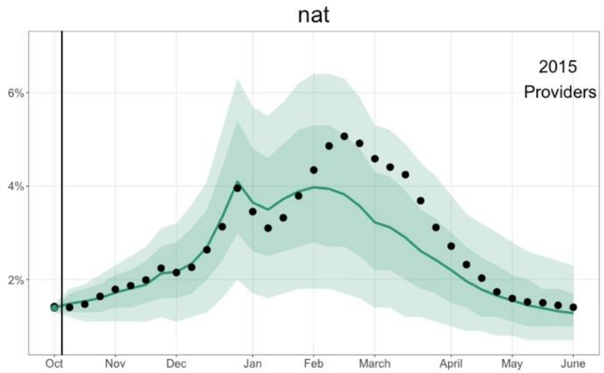


murph@lanl.gov

LA-UR-25-20054

A tale of two outbreaks: Influenza and COVID

- LANL's forecasting model **Dante** won the FluSight Challenge in 2018/19
  - Bayesian hierarchical model
  - Borrowed information across states and seasons
  - Took hours to run



- **However, when COVID hit in late 2019, Dante was all but useless to us. Why?**
  - **Dante** was tailored to forecast a seasonal disease, needing years of historical data
  - **COVID** isn't seasonal, nor were there any historical data at the onset

- LANL's COVID forecasting model had to be made from scratch. **This was stressful!**

GOAL: Develop and test a unified forecasting framework that works for:

1. Diseases with lots of historical data (e.g., flu)
2. Emerging diseases (e.g., COVID in early 2020)
3. Everything in between!

What properties do we want in a forecasting model?

Accurate and **well-calibrated**

Fast to train, faster to forecast

Scalable and **parallelizable**

Able to ingest **large or small** amounts of data

Produce **probabilistic** forecasts (i.e., quantified uncertainties)

Able to borrow information across geographies and/or diseases, but **not require** multiple geographies/ diseases

Seamlessly incorporate **exogenous information** when available (demographic data, mobility data, sequence data, etc.), but **not require** it

Model Ensembling: A Method to Improve Forecasts

$$Ensemble\ Forecast = \sum_k \boldsymbol{w_k}\ Forecast_k$$

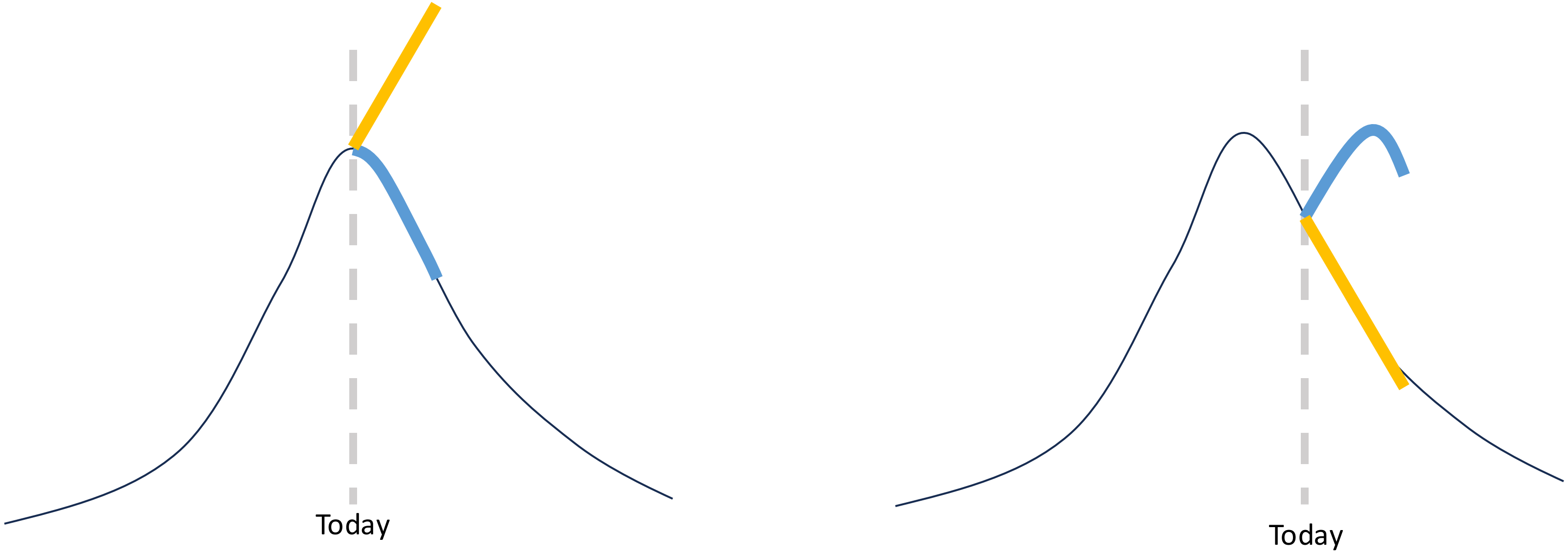
Better Forecast

Okay forecast models

These weights are usually based on past forecast performance

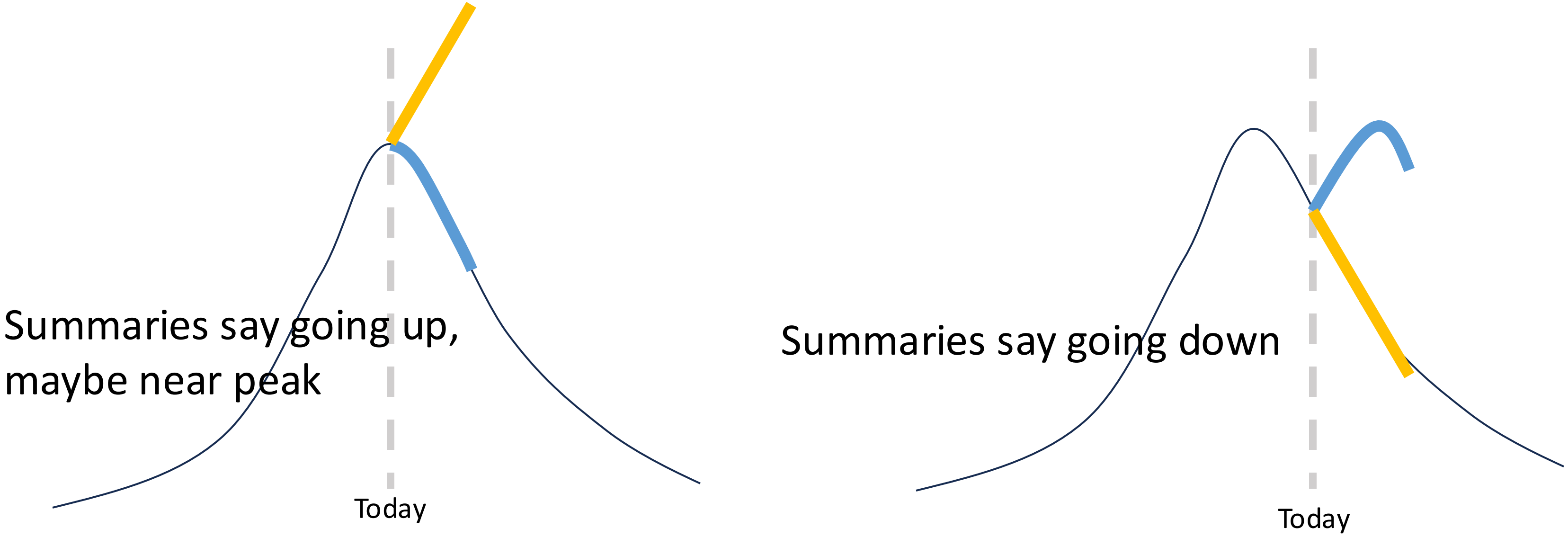
What if we could identify *when* a model performs well?

Model 1: Continue linear trend
Model 2: Reverse last few observations



What if we could identify *when* a model performs well?

Model 1: Continue linear trend
Model 2: Reverse last few observations



Model Ensembling: A Method to Improve Forecasts

$$Ensemble\ Forecast = \sum_{k} \boldsymbol{w_k(x)}\ Forecast_k$$

Problem: How do we specify the weights?

Main Idea: Automated model reweighting based on time series features, x

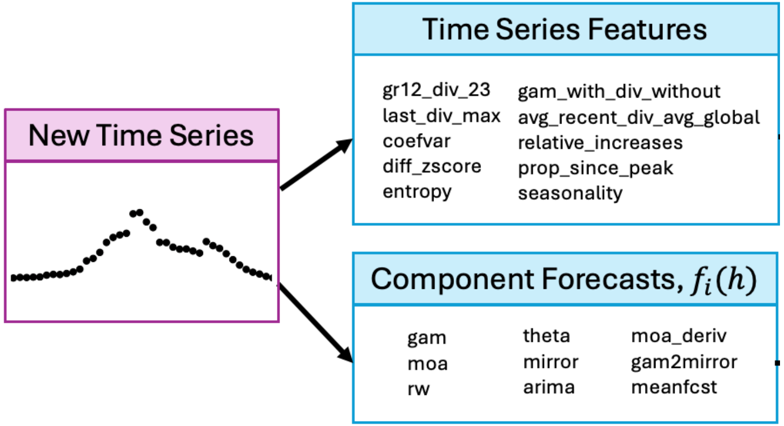
FFORMA: Feature-based Forecast Model Averaging

Pablo Montero-Manso, George Athanasopoulos, Rob J Hyndman, Thiyanga S Talagala

→ **epiFFORMA**

Los Alamos NATIONAL LABORATORY

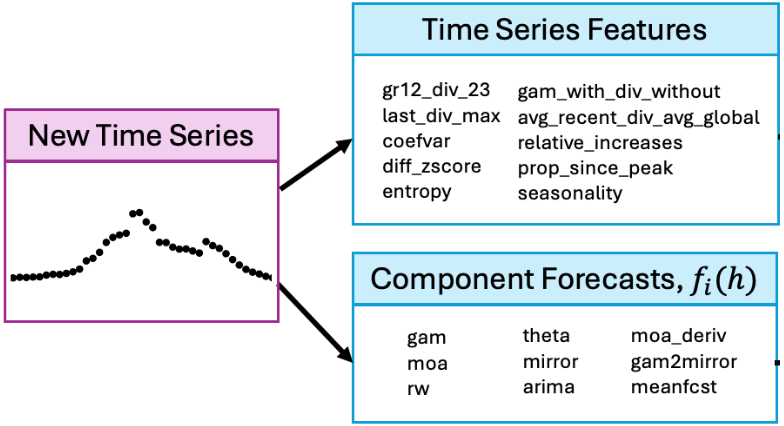
epiFFORMA: data-driven ensembling for infectious disease forecasting



epiFFORMA: data-driven ensembling for infectious disease forecasting

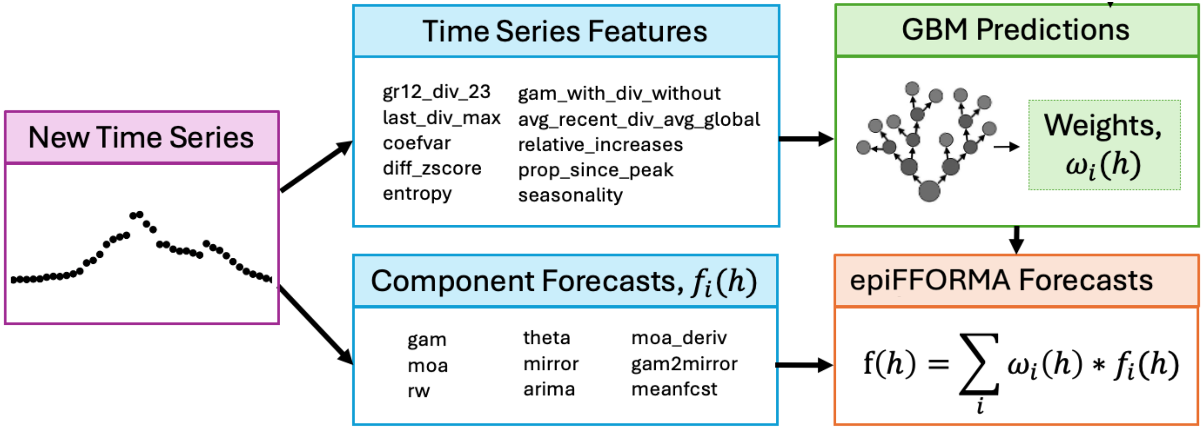
New to epiFFORMA:

Features/components tailored to infectious disease context

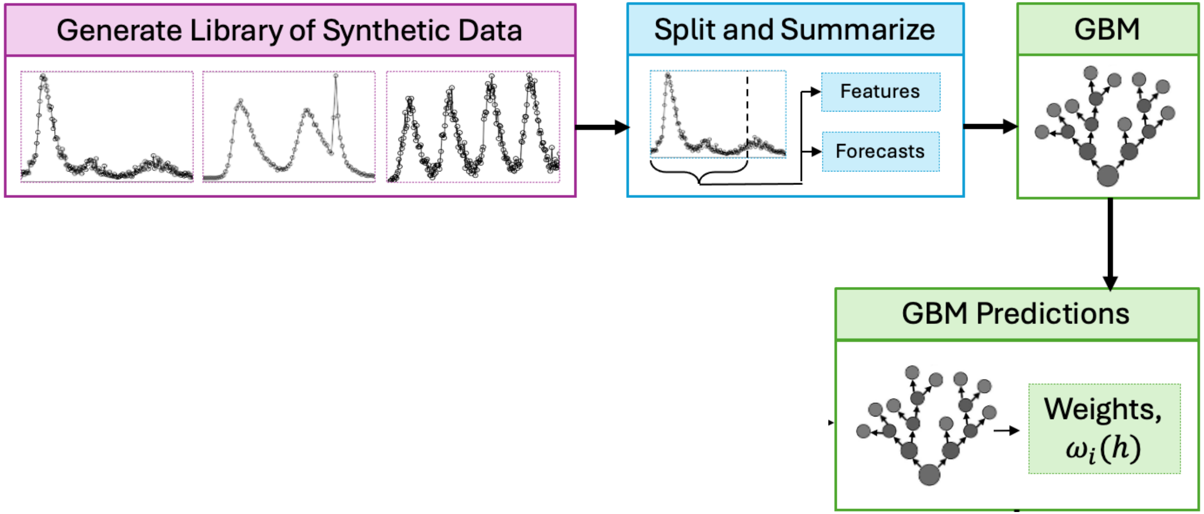


epiFFORMA: data-driven ensembling for infectious disease forecasting

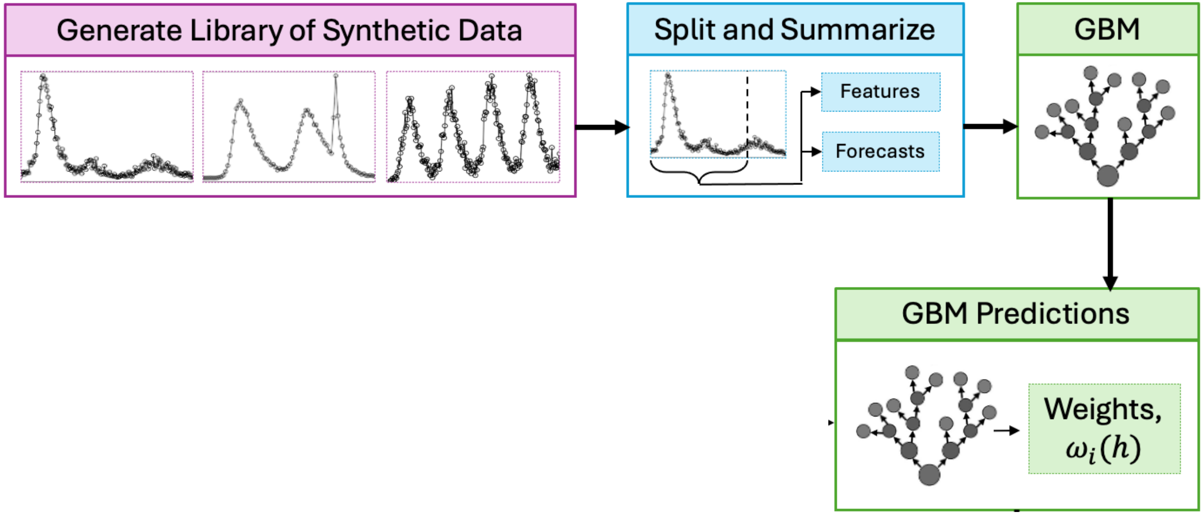
GBM = Gradient Boosted Model



epiFFORMA: data-driven ensembling for infectious disease forecasting



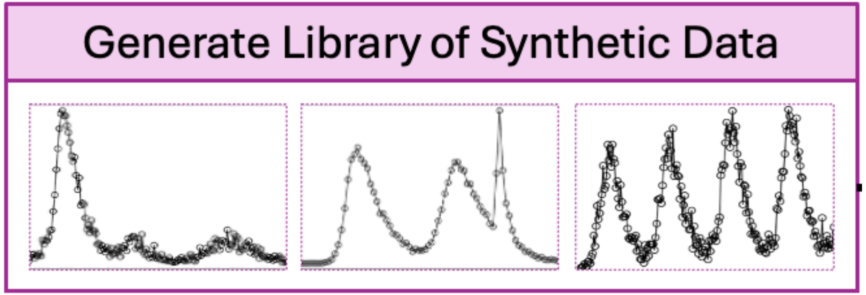
epiFFORMA: data-driven ensembling for infectious disease forecasting



New to epiFFORMA: Use of synthetic data

The Synthetic Data Library

We wish to represent many **possible disease dynamics** without catering to a specific disease.



Generate seasonality components

Generate peaks, troughs, etc.

We mostly used combinations of compartmental models and random effects via sinusoidal functions

Evaluating epiFFORMA

$$Ensemble\ Forecast = \sum_k \boldsymbol{w_k(x)}\ Forecast_k$$

epiFFORMA:	$w_k(x) = GBM$-estimated weight
Equal Weights:	$w_k(x) = 1/K$
Component c:	$w_k(x) = I(k = c)$

Will evaluate using data across **10 different diseases**!!!!

Component Models

rw:
theta:
arima:
gam:
moa:
moa-deriv:
meanfcst:

mirror:

gam2mirror:

From the **forecast** package in R.

Component Models

rw: random walk forecast, with drift = FALSE;
theta:
arima:
gam:
moa:
moa-deriv:
meanfcst:

mirror:

gam2mirror:

From the **forecast** package in R.

Component Models

rw:
theta: equivalent to simple exponential smoothing with drift;
arima:
gam:
moa:
moa-deriv:
meanfcst:

mirror:

gam2mirror:

From the **forecast** package in R.

Component Models

rw:
theta:
arima: autoregressive integrated moving average;
gam:
moa:
moa-deriv:
meanfcst:

mirror:

gam2mirror:

From the **forecast** package in R.

Component Models

rw:
theta:
arima:
gam: generalized additive models;
moa:
moa-deriv:
meanfcst:

mirror:

gam2mirror:

From the **forecast** package in R.

Component Models

rw:
theta:
arima:
gam:
moa: method of analogues, using synthetic data;
moa-deriv:
meanfcst:

mirror:

gam2mirror:

From the **forecast** package in R.

Component Models

rw:
theta:
arima:
gam:
moa:
moa-deriv: the **moa** method on the derivative scale;
meanfcst:

mirror:

gam2mirror:

From the **forecast** package in R.

Component Models

From the **forecast** package in R.

rw:
theta:
arima:
gam:
moa:
moa-deriv:
meanfcst: mean of the time series up through two years before the start of the forecast;
mirror:

gam2mirror:

Component Models

rw:
theta:
arima:
gam:
moa:
moa-deriv:
meanfcst:

From the **forecast** package in R.

mirror: predictions symmetric about the start of a forecast, out through the forecast horizon;
gam2mirror:

Component Models

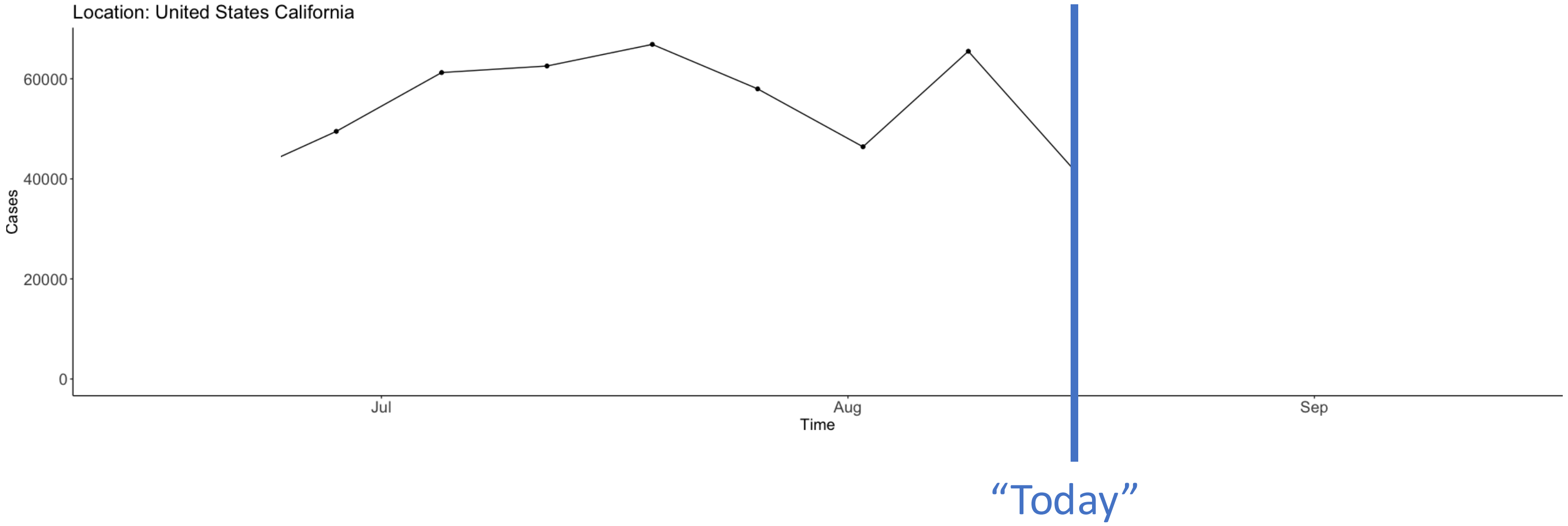
rw:
theta:
arima:
gam:
moa:
moa-deriv:
meanfcst:

From the **forecast** package in R.

mirror:

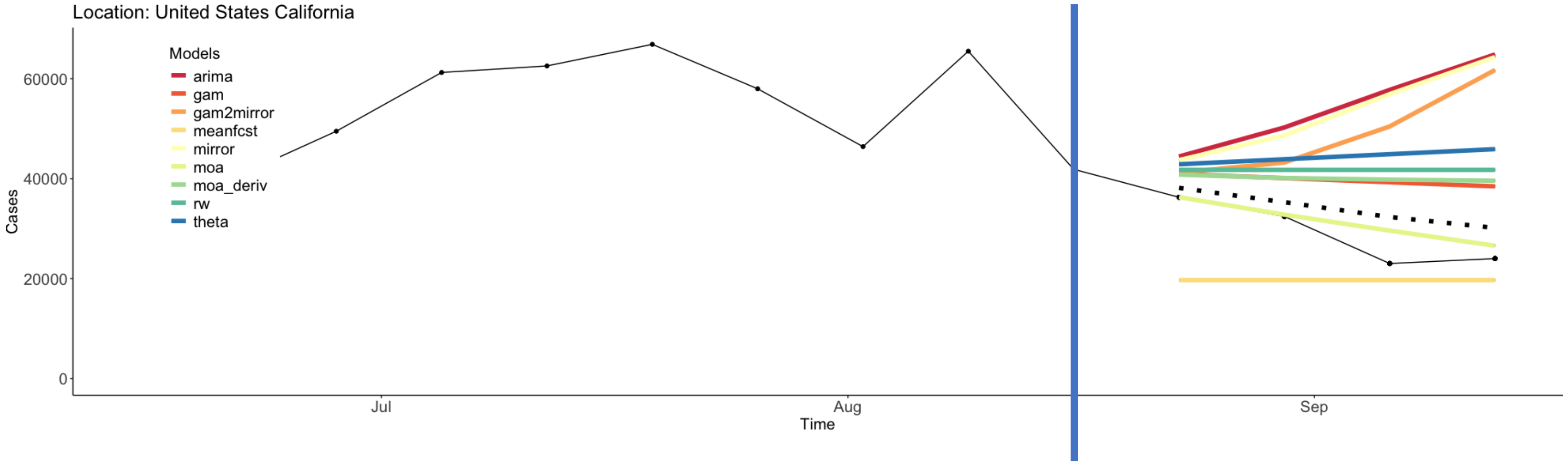
gam2mirror: a linear combination of the **gam** and the **mirror** models.

epiFFORMA Results

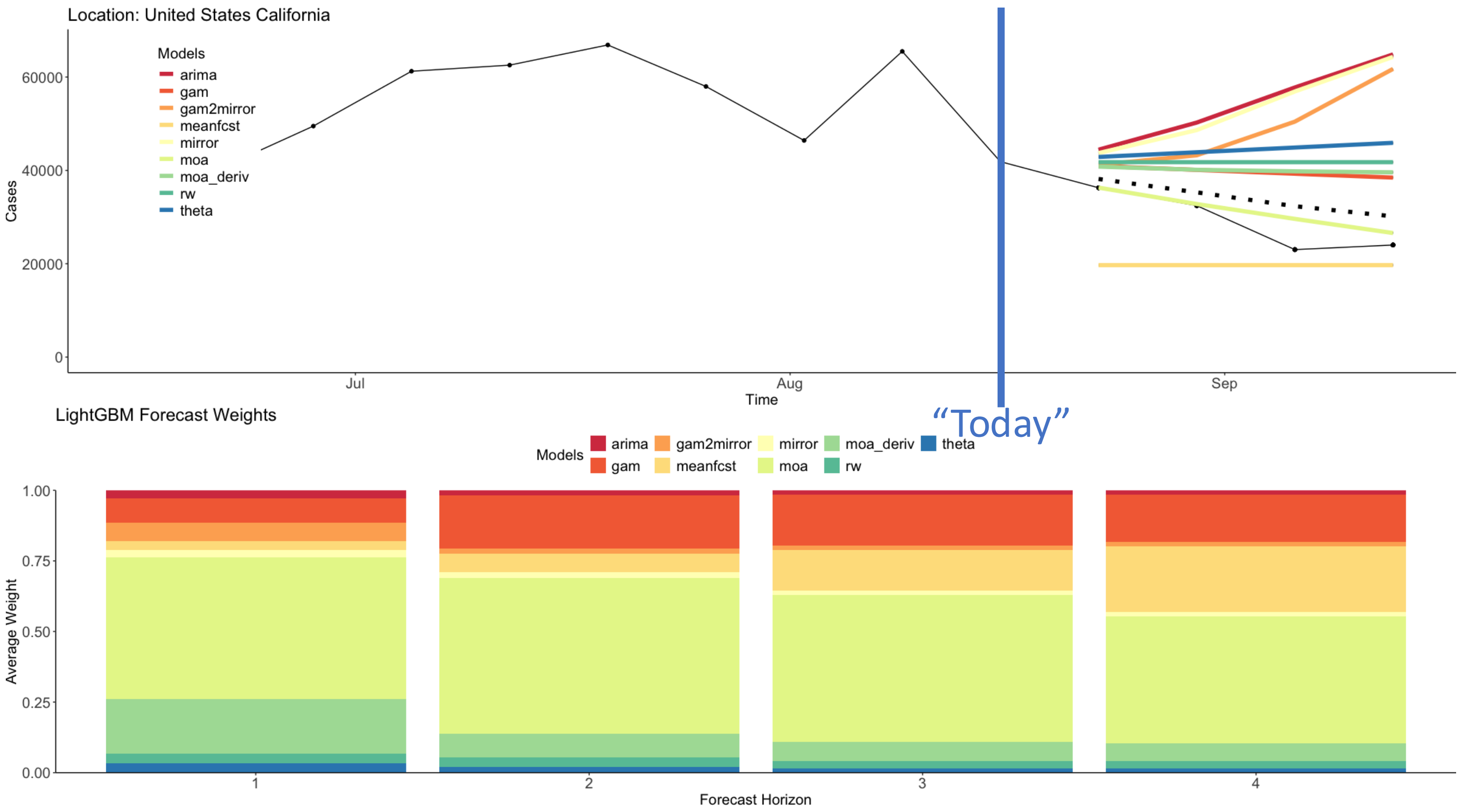


Initial example of epiFFORMA: COVID-19 cases in California.

- **Starting in early September 2020.**
- **Forecast 4 weeks into the future.**


Location: United States California
Models
arima
gam
gam2mirror
meanfcst
mirror
moa
moa_deriv
rw
theta
60000
40000
20000
0
Cases
Jul
Aug
Sep
Time

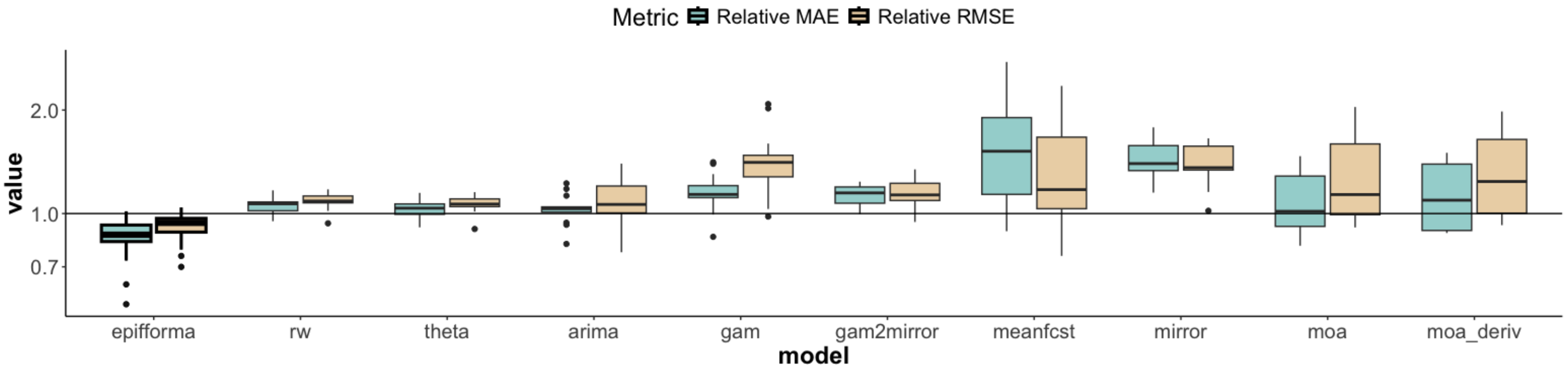
"Today"


Location: United States California
Models
arima
gam
gam2mirror
meanfcst
mirror
moa
moa_deriv
rw
theta
Cases
60000
40000
20000
0
Jul
Aug
Sep
Time
"Today"
LightGBM Forecast Weights
Models
arima
gam
gam2mirror
meanfcst
mirror
moa
moa_deriv
rw
theta
Average Weight
1.00
0.75
0.50
0.25
0.00
1
2
3
4
Forecast Horizon

epiFFORMA Results

(now across 10 diseases)

COVID-19 (US)	Mumps (US)
COVID-19 (Global)	Polio (US)
ILI (US)	Rubella (US)
Diptheria (US)	Smallpox (US)
Measles (US)	Chikungunya (Brazil)

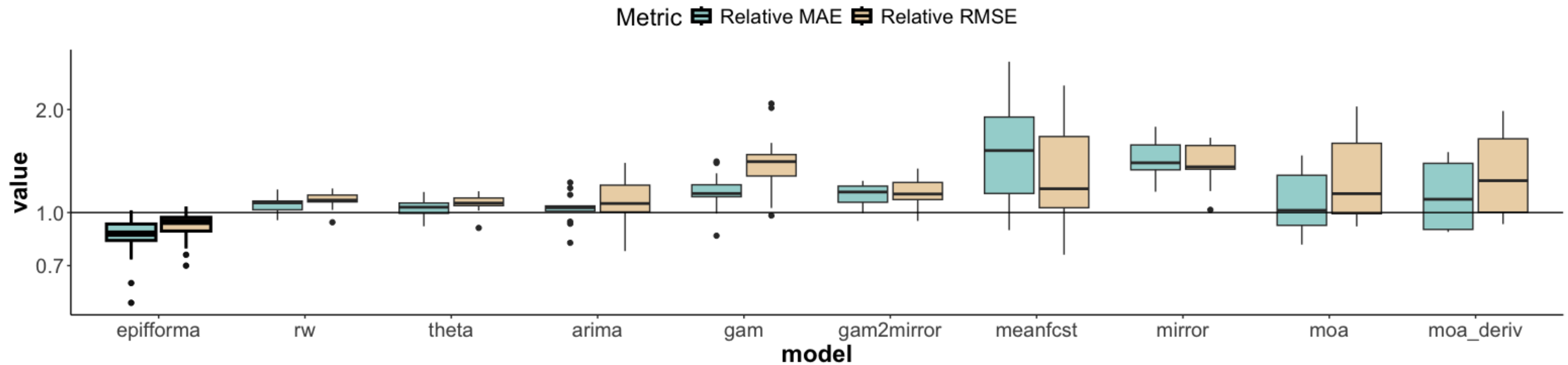


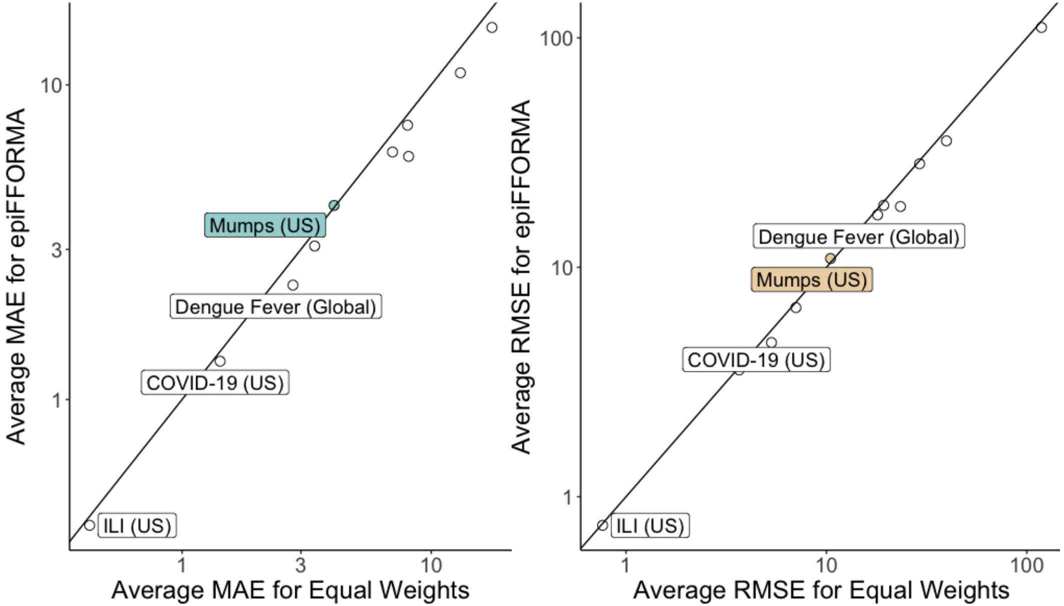
Recall:

epiFFORMA: $w_k(x) = GBM$-estimated weight

Equal Weights: $w_k(x) = 1/K$

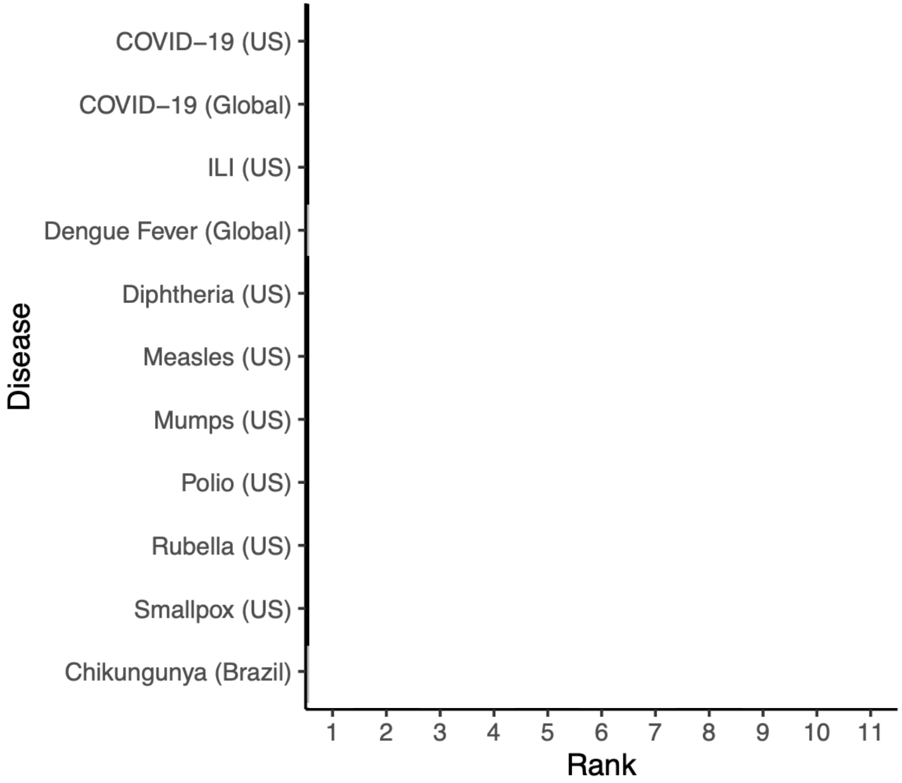
Component c: $w_k(x) = I(k = c)$





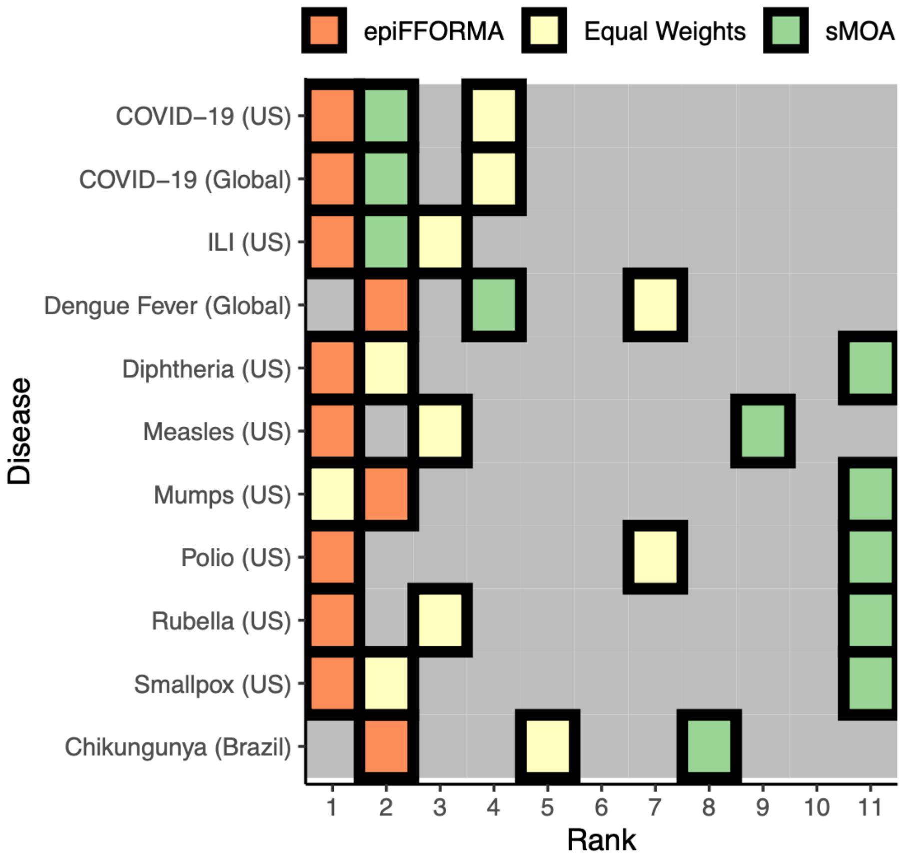
By disease, epiFFORMA out-performs the equal-weights ensemble for every disease except Mumps.

Overall Forecast Rankings (MAE)



Overall Forecast Rankings (MAE)

epiFFORMA outranks other methods



Los Alamos NATIONAL LABORATORY

Uncertainty Quantification for epiFFORMA

$$Ensemble\ Width = \sum_k \boldsymbol{\omega_k(x)}\ Width_k$$

Ensemble
95% Interval Width

Component Model
95% Interval Widths

Uncertainty Quantification for epiFFORMA

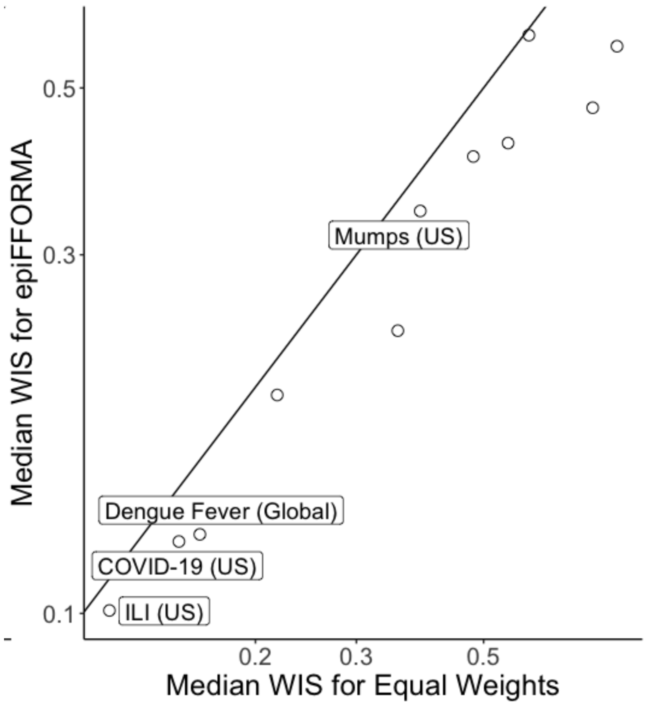
$$Ensemble\ Width = \sum_k \boldsymbol{\omega_k(x)}\ Width_k$$

Ensemble
95% Interval Width

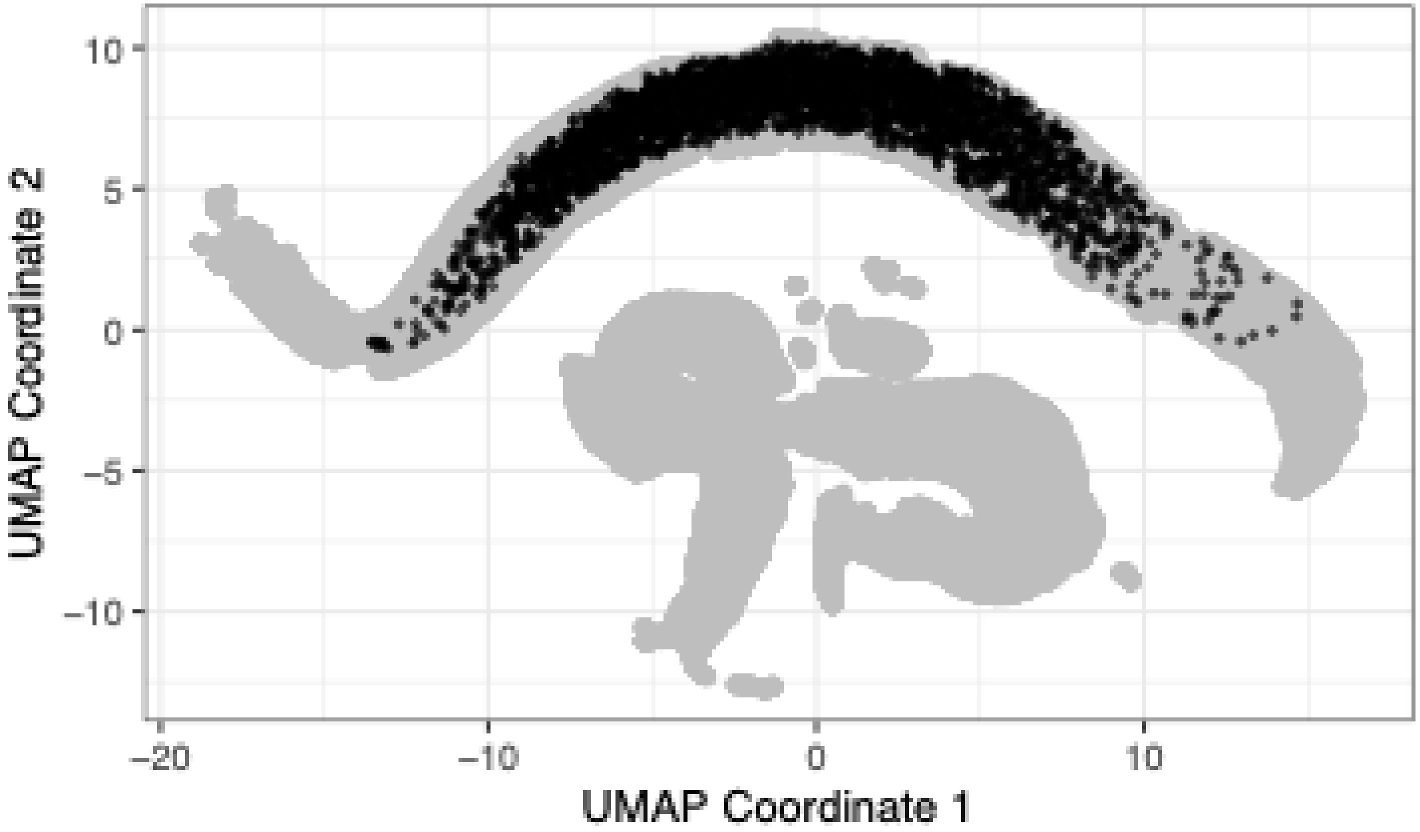
Component Model
95% Interval Widths

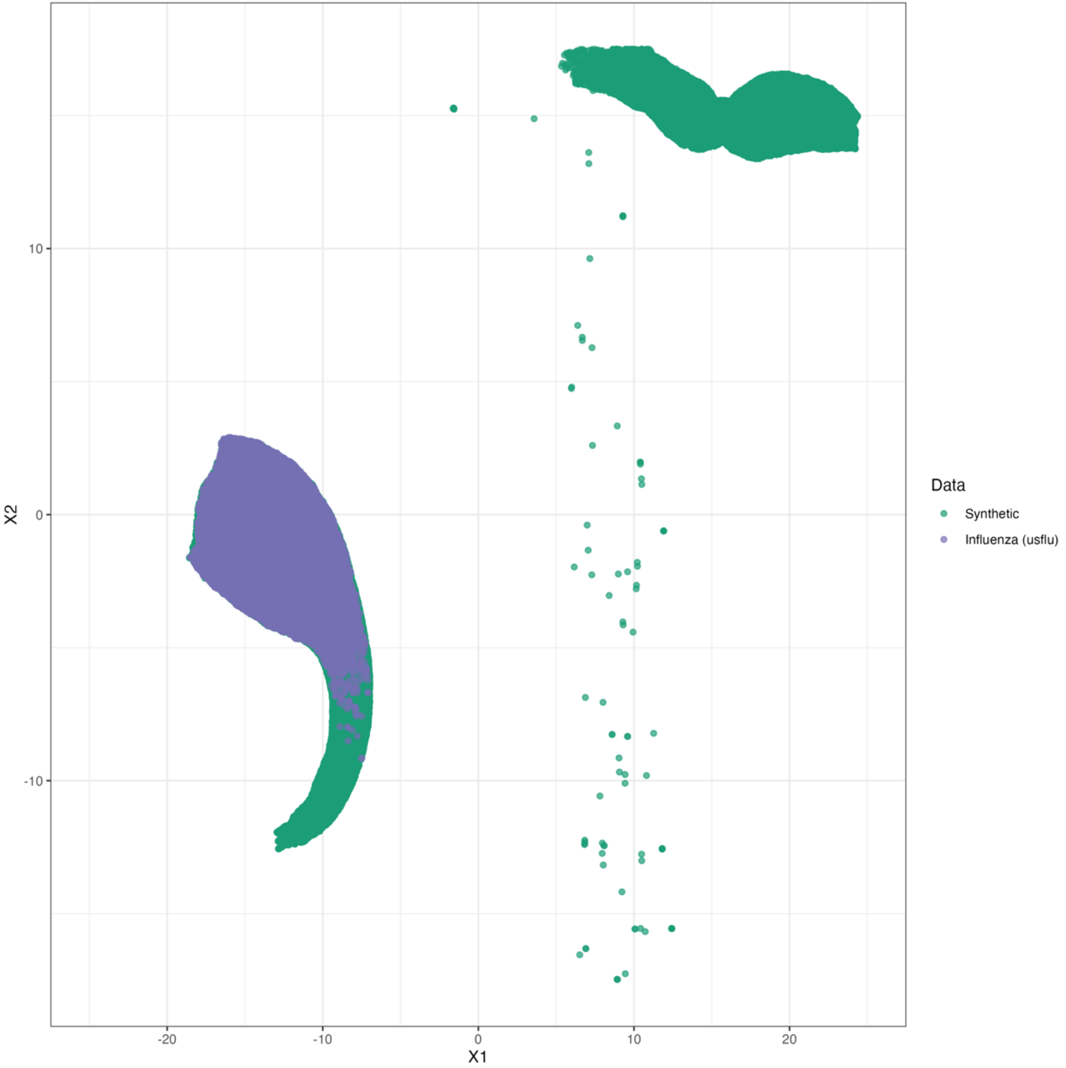
epiFFORMA has smaller WIS than equal weights model

Comparing UQ for epiFFORMA by Disease



Back-up Slides


10
5
0
−5
−10
UMAP Coordinate 2
−20
−10
0
10
UMAP Coordinate 1



X2
10
0
-10
X1
-20
-10
0
10
20
Data
Synthetic
Influenza (usflu)