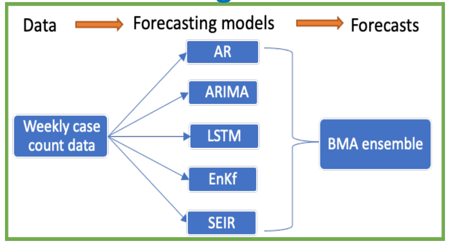
Phase-Informed Bayesian Ensemble Models

Gursharn Kaur
Biocomplexity Institute, UVA

COVID-19 forecasting and Bayesian Model Averaging

Existing Method



- The current Bayesian ensemble combines forecast from *AR*, LSTM, EnKF and *SEIR*.
- The BMA pipeline is computed every week, independently for each county.
- The predictive distribution is assumed to be a *mixture of Gaussians*, with weights given by the posterior probabilities for each model in the BMA.

Goal

- Propose a new Bayesian ensembling method with updated weighting schemes based on different phases of the pandemic.
- Evaluate proposed methods by fitting BMA on past data retrospectively.

Our Approach

- Identify and segment the observed time period into three main phases (waves) in the pandemic as – surge, decline and plateau phase.
- Use phase specific performance of models to obtain the model weights in the BMA ensemble, independently for each county.

Proposed Method and Results

Phase segmentation

Based on the weekly case counts, we classify the trend of the pandemic in one of the three following categories:

- ❖Surge phase
- ❖Decline phase
- ❖Plateau phase

- Break points are estimated recursively with each new week.
- R-package used - ***segmented***.

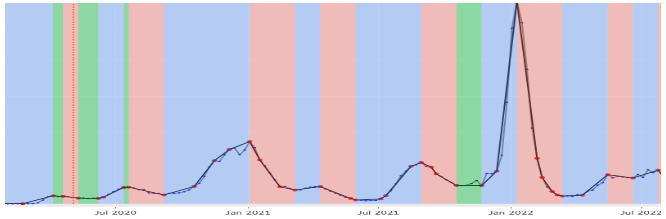
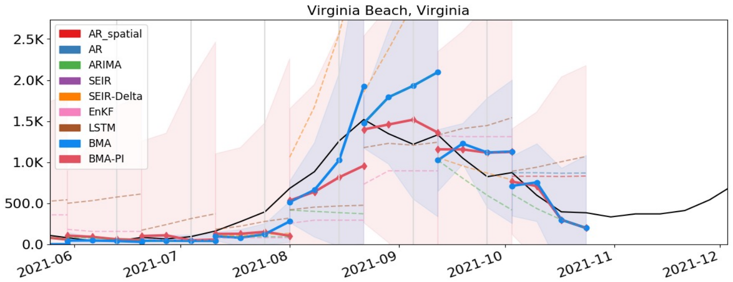


Figure 2: A piece-wise linear fit and phase classification for USA case counts

Phase Informed BMA (PI-BMA)

- Design a BMA ensemble that uses the knowledge of phase segmentation to train the model.
- Phase specific historical performance of individual methods is used for estimating the weights.



Performance evaluation across forecasting weeks

1. R_t-based analysis

- We estimate the $\mathrm{R_t}$ value using the incidence case time-series and a simulation model.
- We observe that different counties experience different $\mathrm{R_t}$ values and different set of methods perform well in different phases.

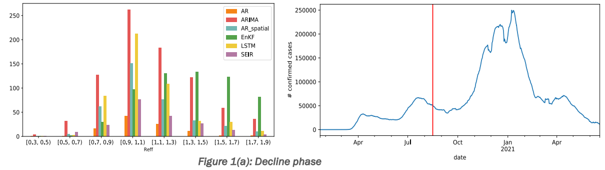


Figure 1(a): Decline phase

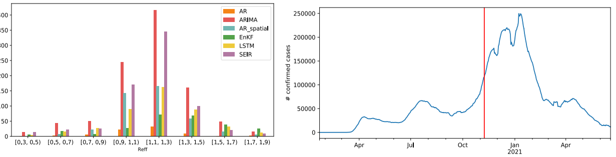


Figure 1(b): Surge phase

2. Ablation analysis

- The payoff set function for a county c at time t:

$$v^{c,t}(S) = \frac{|y^{c,t} - f^{c,t}(S)|}{y^{c,t}} \quad \text{for } S \subseteq N = \{1, \ldots, n\}$$

where $y^{c,t}$ = ground truth and $f^{c,t}(S)$ = forecast obtained with S set of methods in the BMA ensemble.

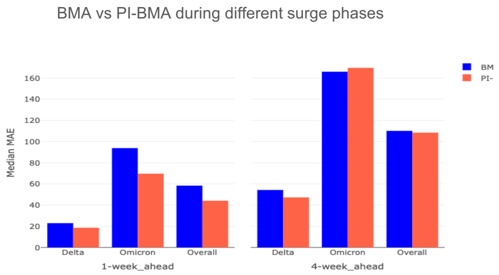
- The influence of a method i is defined as:

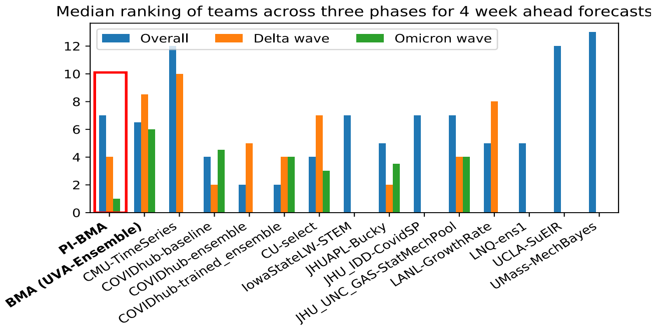
$$\phi^{c,t} = \frac{1}{2^{n-1} - 1} \sum_{S(\neq\emptyset)\subseteq N\setminus\{i\}} v^{c,t}(S \cup i) - v^{c,t}(S)$$

Our findings:

- ➢ARIMA and LSTM get the most significant negative values throughout the observed time period.
- ➢Variable performance of SEIR near surge phase.

Results and retrospective evaluation





Summary and Conclusion

- ➢All models are useful but including every model in the ensemble may reduce forecast performance
- ➢Phase are identified to be important indicator of model specific performance
- ➢Compartmental models are useful during growth and decline phases but tend to over-estimate the future case counts.
- ➢PI-BMA leads to improved performance at a critical phase, when compared to other methods.