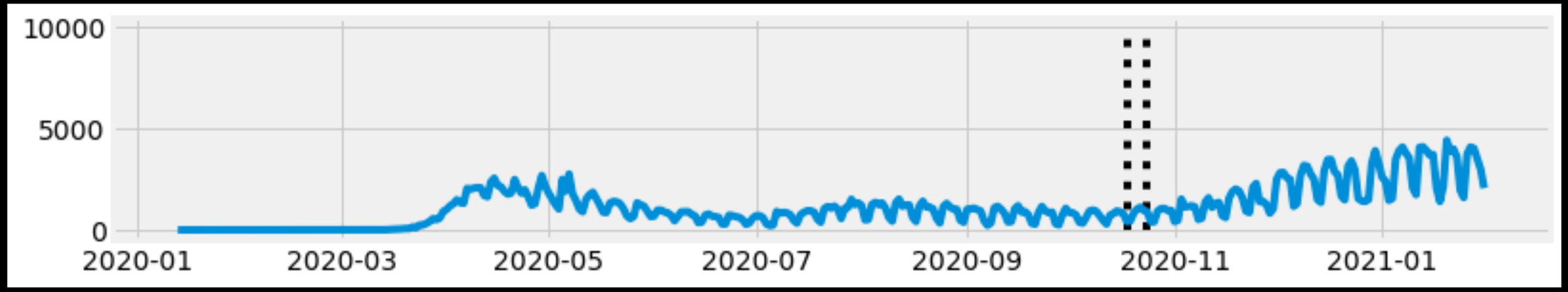
Coronavirus Prediction

Predicting daily numbers of deaths and mortality rates for the US with data sourced from covidtracking.com api

Authors: Ning Chen and Elliot Macy

New Deaths
per day

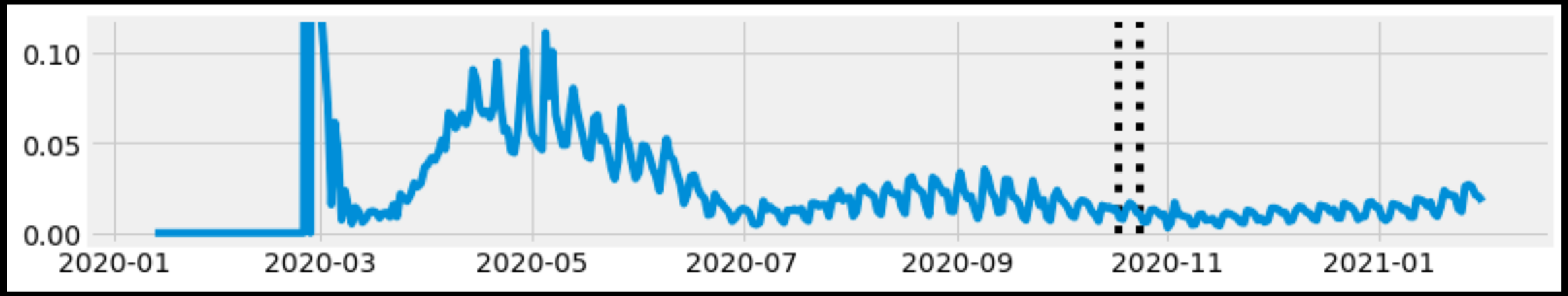
- Weekly seasonality
- Dips on weekends
- Peaks on Mondays
- Reporting?



Mortality Rate

per day

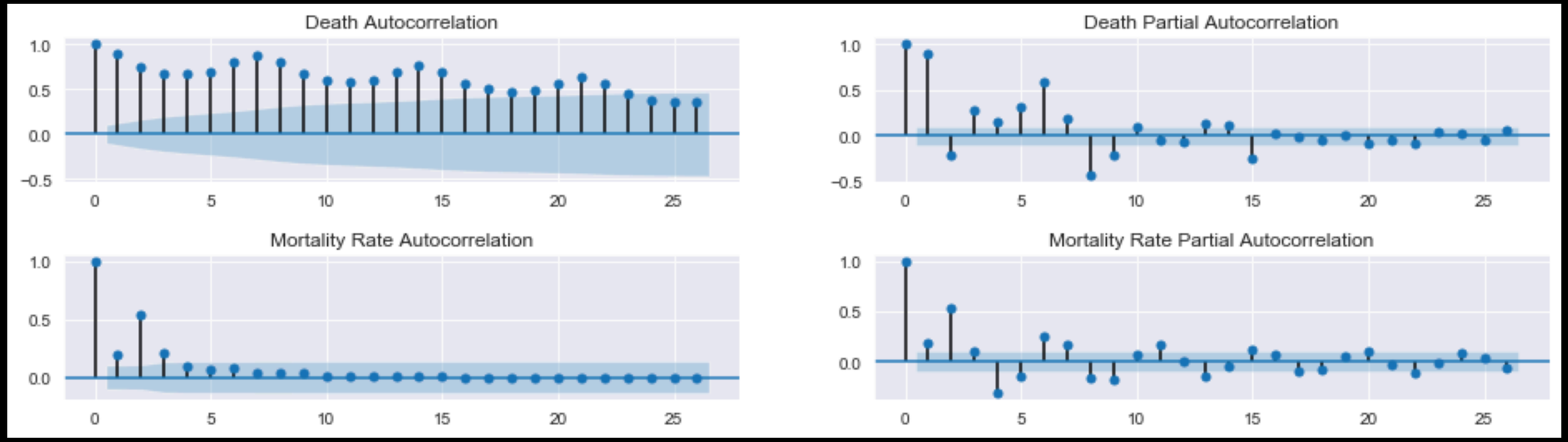
- Weekly seasonality
- Dips on weekends
- Peaks on Mondays
- Reporting?



ACF/PACF

Auto Correlation Formula
/ Partial Auto Correlation Formula

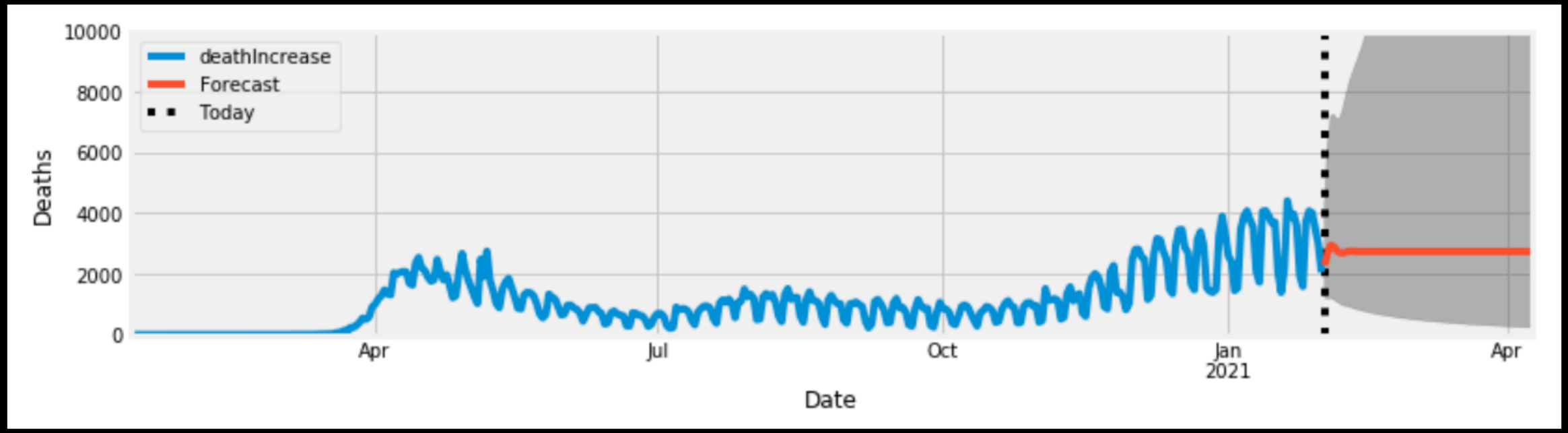
- Weekly seasonality
- Dips on weekends
- Peaks on Mondays
- Reporting?



ARIMA

Auto Regressive Integrated Moving Average

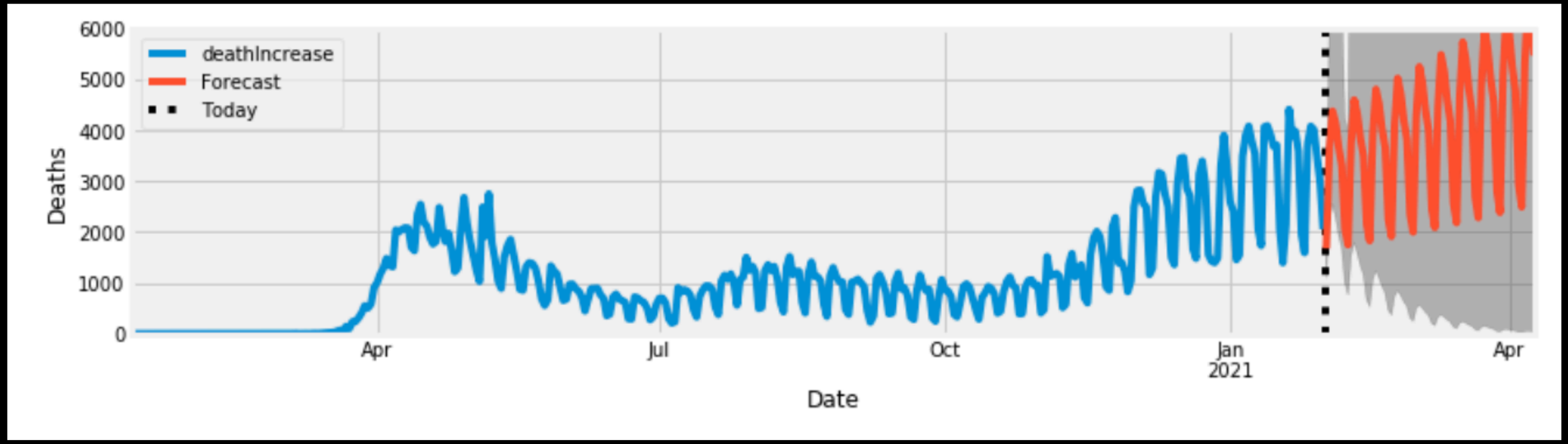
- Non-Dynamic RMSE = 1414.04
- Non-Dynamic MAE = 1174.88
- Dynamic RMSE = 1941.94
- Dynamic MAE = 1599.49
- (Stationarity violated)



SARIMAX

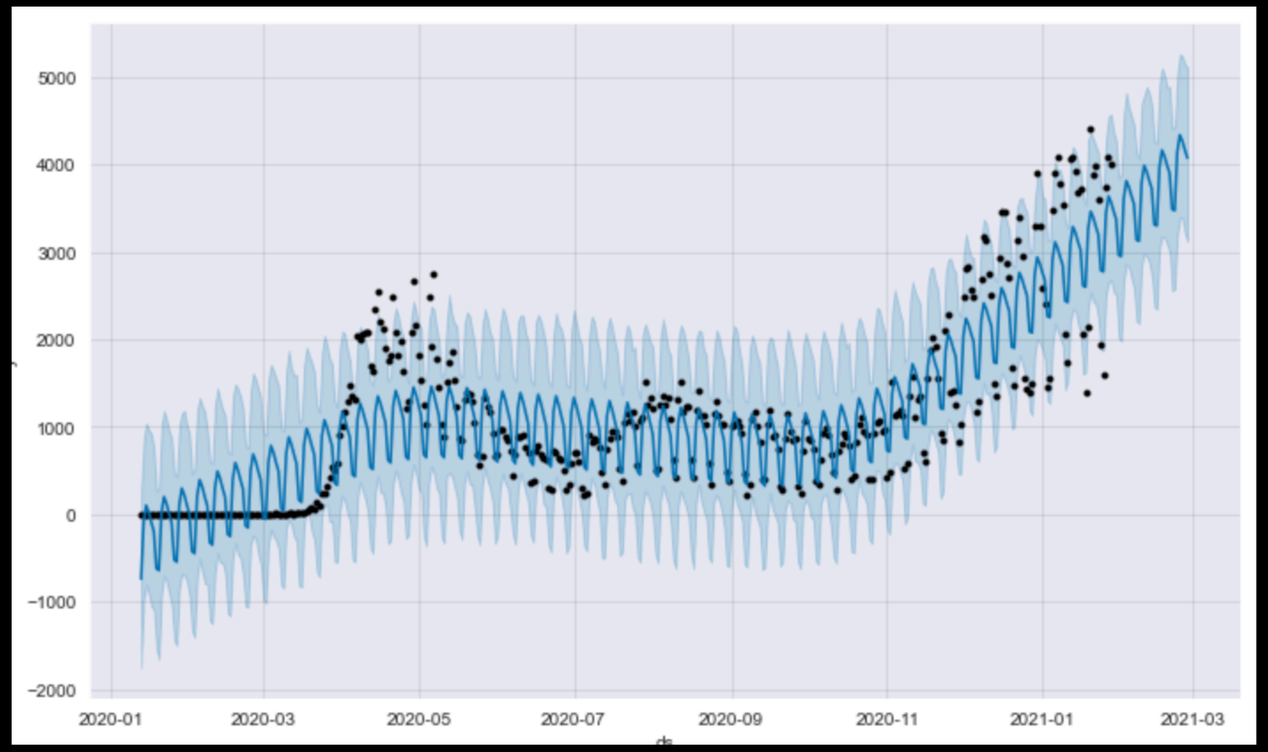
Seasonal Auto Regressive Integrated Moving Average with eXogenous regressors

- Non-Dynamic RMSE = 1464.89
- Non-Dynamic MAE = 1183.11
- Dynamic RMSE = 1743.65
- Dynamic MAE = 1363.64



Facebook Prophet

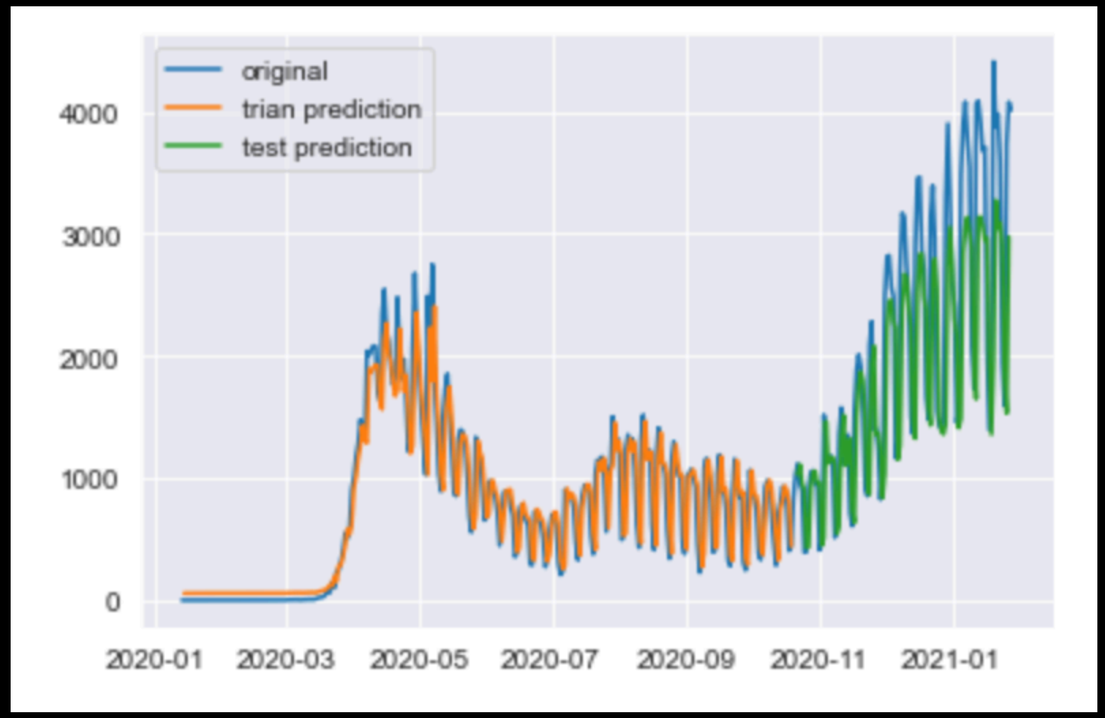
- Train RMSE = 432.03
- Train MAE = 321.32
- Test RMSE = 1929.95
- Test MAE = 1579.82



LSTM

Long Short-Term Memory

- Train RMSE = 275.27
- Train MAE = 788.28
- Test RMSE = 183.45
- Test MAE = 563.24



Model Evaluation
RMSE / MAE

ARIMA:

- Non-Dynamic RMSE = 1414.04
- Non-Dynamic MAE = 1174.88
- Dynamic RMSE = 1941.94
- Dynamic MAE = 1599.49
- (Stationarity violated)

SARIMAX:

- Non-Dynamic RMSE = 1464.89
- Non-Dynamic MAE = 1183.11
- Dynamic RMSE = 1743.65
- Dynamic MAE = 1363.64

FB Prophet:

- Train RMSE = 432.03
- Train MAE = 321.32
- Test RMSE = 1929.95
- Test MAE = 1579.82

LSTM:

- Train RMSE = 275.27
- Train MAE = 788.28
- Test RMSE = 183.45
- Test MAE = 563.24

Conclusions
and Recommendations

1. LSTM Network serves as the best model for coronavirus prediction (smallest RMSE and MAE)
2. SARIMAX with optimized hyperparameters by Gridsearch also works well for predictions
3. Due to many unknown exogenous features, such as transmission rate and the effect of holidays/lockdowns it is difficult to make precise predictions
4. Fewer deaths (and positive cases) on weekends and spikes on Mondays suggest errors in reporting
5. Based on LSTM, new deaths appear to be leveling off, however there is a wide confidence interval, so we recommend being cautiously optimistic