**Team 4Sight2 General methodology:**

We will begin with simple time series models for short-term targets and historical estimates with some corrections for the seasonal targets. Sometime during season, we hope to start using a metapopulation SEIR model of influenza, at the level of counties, connected by commuter and airline matrices.

The following document summarizes the methodology/updates for the corresponding EW submissions. Unless otherwise mentioned, same methodology was followed for both FluSight and StateFluSight.

**EW43 - Submitted Nov 6, 2017**

For Seasonal targets, we obtained the historical estimates. While it is straightforward to do this for peak week and peak percentage, we used historical baseline values for onset ( <https://github.com/craigjmcgowan/cdc-flusight-ensemble/blob/master/baselines/wILI_Baseline.csv> ). We also removed the pandemic season (2009-10) to exclude the erratic peak/onset timing. Given the historical seasonal targets, we obtained the histogram and median, and used it produce binned and point forecasts.

Short term targets were produced using an ARIMA (1,0,3) model trained on all historical data available (excluding the early seasons, before 2002, of National/HHS) plus the current season’s data. Short term binned targets were assigned a probability of 1 at the point forecast produced by ARIMA.

Other details:

When we produced these forecasts, Louisiana’s historical ILI data was unavailable on FluView (later obtained it through the EPI website). So we used Alabama’s seasonal and short-term forecasts for Louisiana. Similarly for Mississippi the current ILI wasn’t available on FluView (but available on EPI website).

**EW44 - Submitted Nov 13, 2017**

Since last submission the following updates were made:

1. Since ARIMA was producing relatively flat forecasts, we switched to polynomial fitting (numpy’s polyfit). For now, a linear fit is done to the most recent four data points, and is used to forecast ahead.
2. We also added a module to update onset predictions by combining seasonal estimates and current season trends. For the current season, for a given region (say US National), we estimate the following variables. Depending on these, we reweight the previously computed binned probabilities with 1-\alpha and add \alpha to the appropriate week. (Point forecasts not being updated though.)
   1. G\_onsetoccur - when/whether onset occurred in Ground truth (\alpha = 1)
   2. G\_basecross - when/whether baseline is crossed in Ground truth (\alpha = 0.75)
   3. P\_onsetoccur - when/whether onset occurred in Prediction (\alpha = 0.5)
   4. P\_basecross - when/whether baseline is crossed in Prediction (\alpha = 0.25)
3. Finally, for StateFluSight, we were able to incorporate the past and current ILI data for Louisiana and also the current ILI data from Mississippi (available from EPI website but not on FluView) into our data.

**EW45 - Submitted Nov 20, 2017**

Minor update from previous version (fixed a bug in (b)).

**EW46 - Submitted Nov 29, 2017**

Since last submission the following updates have been made:

1. The point forecasts for seasonal targets (especially onset) has been made to coincide with the median of the binned forecasts.
2. The historical seasons are weighted by their correlation to current ILI when arriving at seasonal targets forecasts (for both StateFluSight and FluSight).
   1. The correlation is computed as the normalized pearson correlation coefficient. Note that the PCC gives values in [-1,1], so before normalizing across years, it is converted to [0,1] scale: (PCC(curr\_ILI, past\_ILI)+1)/2

**No method updates for EW47, EW48 and EW49**