Bayesian Functional Data Analysis in Influenza Forecasting

Nehemias Ulloa, Jarad Niemi

Iowa State University

ILINet

- U.S. Outpatient Influenza-like Illness Surveillance Network (ILINet) via Centers for Disease Control and Prevention (CDC)
- ILINet consists of over 3,500 enrolled outpatient healthcare providers
  - Total number of patients seen for any reason
  - Number of those patients with influenza-like illness (ILI) by age group

Functional Data Analysis

- Consider the entire season (function) as a data point
- Think of weekly observation as discrete observations of a smooth underlying process

Data + Model

- $y_{r,s}(w)$ - proportion of patients with Influenza-like illness in week w, region r, season s
- $\hat{\mu}(w)$ - estimated weekly means across regions and seasons
- $\theta_{r,s}(w)$ - smooth underlying function of w in region r, season s
- $\epsilon_{r,s}(w)$ - observational error

A functional data model can be written as such:

$$
\begin{aligned}
y_{r,s}(w) - \hat{\mu}(w) &= \theta_{r,s}(w) + \epsilon_{r,s}(w) \\
\epsilon_{r,s}(w) &\overset{ind}{\sim} N(0, \sigma_\epsilon^2)
\end{aligned} \tag{1}
$$

Basis Choices

Possible basis are:

- Polynomial basis
- Fourier basis
- Principal component basis

Functional Principled Component Analysis

The FPCA method (Dauxois and Pousse (1976)) is represented like this:

$$\theta_{r,s}(w) = \sum_{k=1}^{K} \beta_{r,s,k}\hat{\phi}_k(w) \tag{2}$$

where $\hat{\phi}_k(w)$ are the estimated principal component basis

15-16 Influenza Season

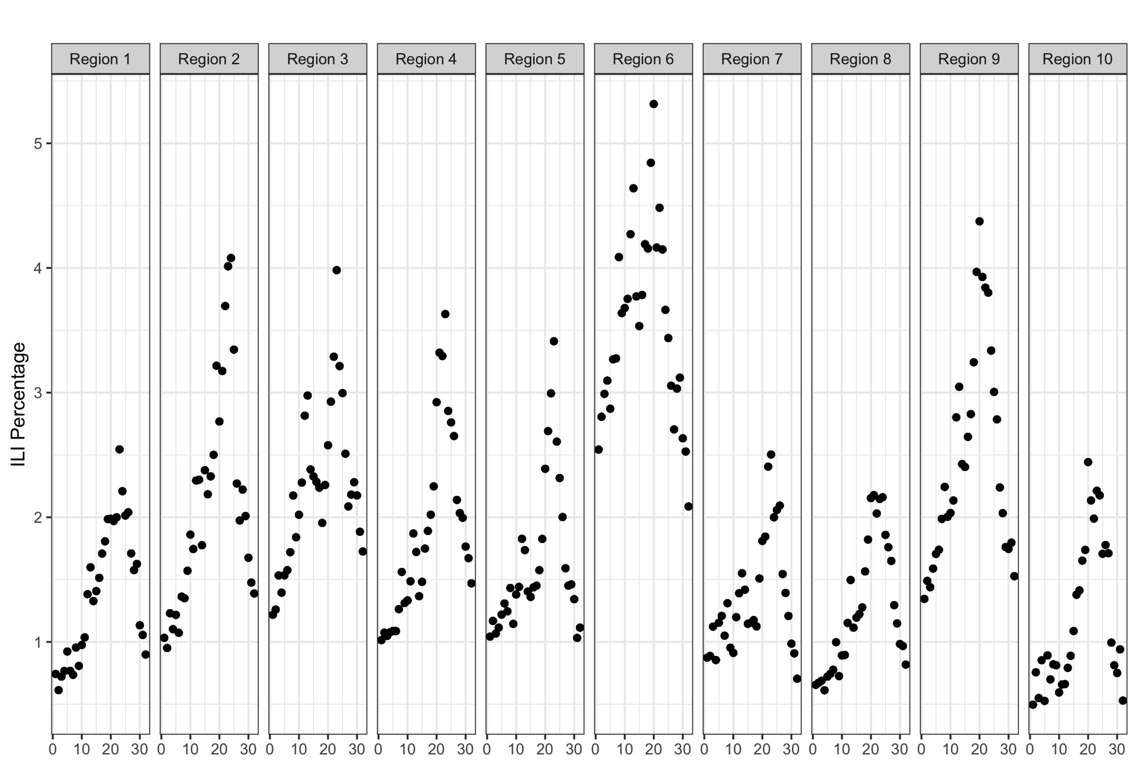

Figure: The 2015-2016 influenza season faceted by region.

Sparse Priors for Basis Selection

Wang et al, (2015):

- Point out the question of how large should K be?
- Amount of variance is accounted for and the arbitrary cutoff value is chosen
- AIC and BIC as well as Leave one out Cross Validation (LOO CV) but they include too many components

Horseshoe Prior

The horseshoe prior (Carvalho et al. (2009)) model is noted by:

$$
\begin{aligned}
y &\sim N(X\beta, \sigma^2 I) \\
\beta_{r,s,i} &\overset{ind}{\sim} \text{HS}(\tau_{r,s})
\end{aligned} \tag{3}
$$

where HS is a scale mixture of normals:

$$
\begin{aligned}
\beta_{r,s,i}|\lambda_{r,s}, \tau_{r,s} &\sim N(0, \lambda_{r,s,i}^2\tau_{r,s}^2) \\
\lambda_{r,s,i} &\overset{ind}{\sim} Ca^+(0,1)
\end{aligned} \tag{4}
$$

Sparse Priors for Basis Selection (cont'd)

Hierarchical Prior

The hierarchical prior (Piironen and Vehtari (2017)) is similar to the horseshoe but has a half-t prior:

$$
\begin{aligned}
y &\sim N(X\beta, \sigma^2 I) \\
\beta_{r,s,i}|\lambda_{r,s}, \tau_{r,s} &\sim N(0, \lambda_{r,s,i}^2\tau_{r,s}^2) \\
\lambda_{r,s,i} &\overset{ind}{\sim} t_4^+(0,1)
\end{aligned} \tag{5}
$$

LASSO Prior

The LASSO prior (Park and Casella (2008)) is a simpler shrinkage prior that the horseshoe type priors:

$$
\begin{aligned}
y &\sim N(X\beta, \sigma^2 I) \\
\beta_{r,s,i}|\lambda_{r,s}, \tau_{r,s} &\sim N(0, \lambda_{r,s,i}^2) \\
\lambda_{r,s,i} &\overset{ind}{\sim} Laplace(a,b)
\end{aligned} \tag{6}
$$

Questions of Interest

- How well do they model the influenza season?
- Do the shrinkage priors naturally choose the number of basis functions?
- How well are the forecasts of a new season?

Model Fit

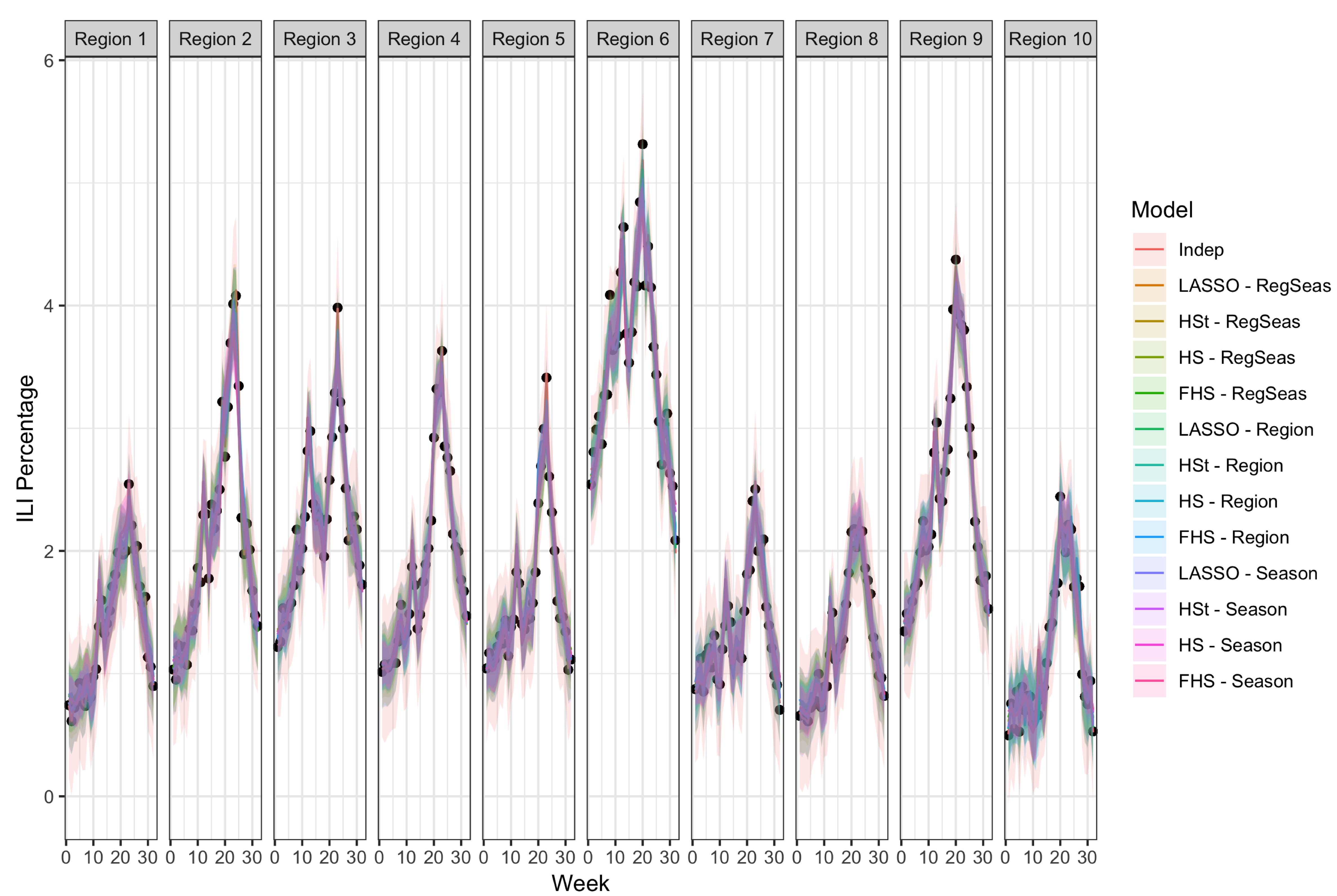

Figure: The 2015-2016 influenza season faceted by region.

Basis Selection

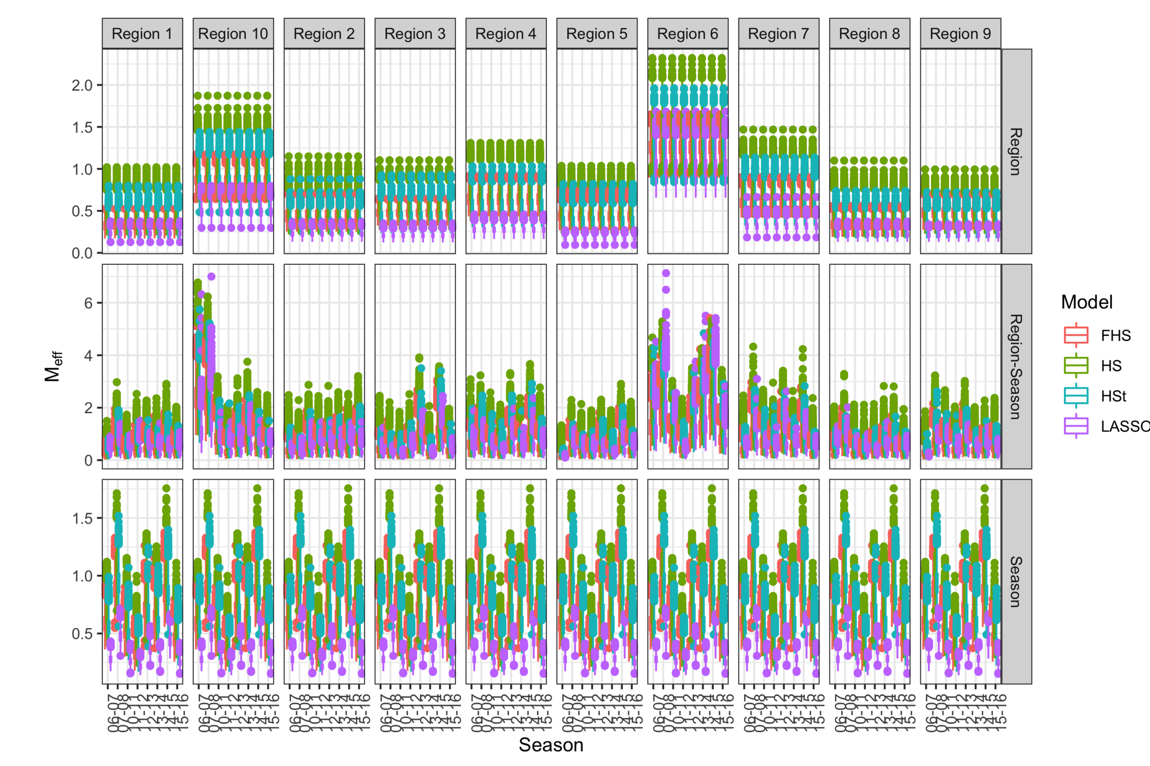

Figure: M_{eff} for the model fits.

Forecasting

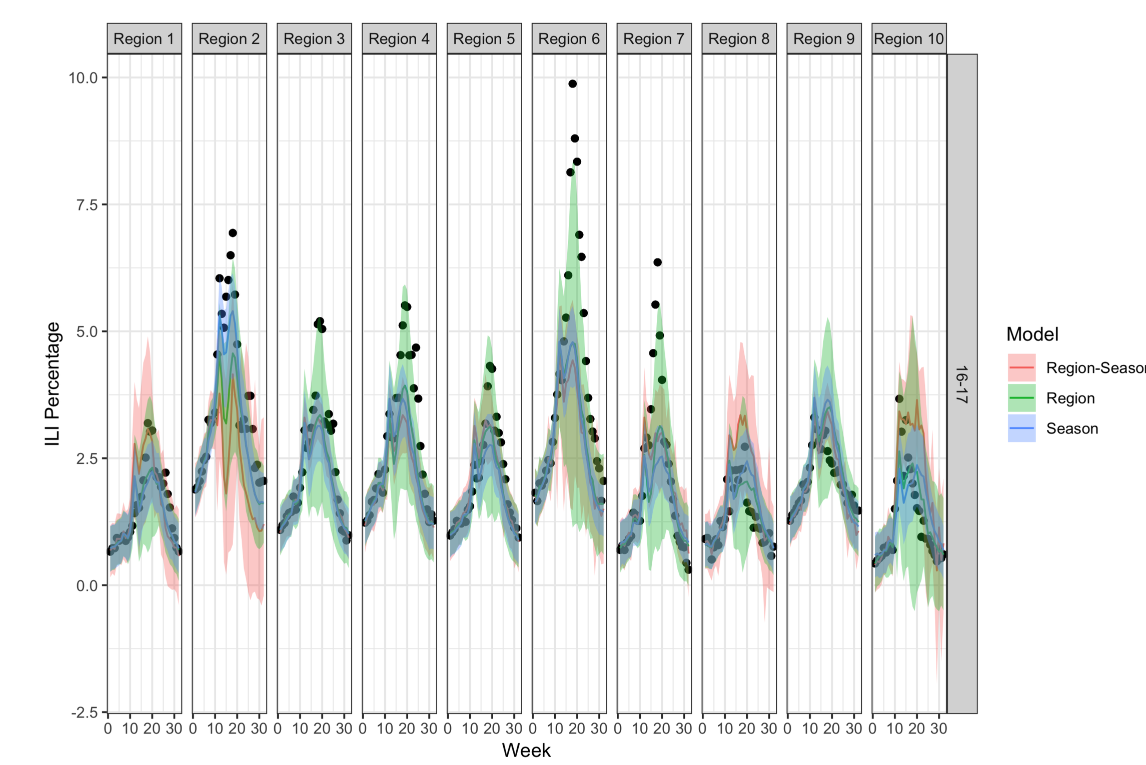

Figure: This plot shows the season forecasts for the HSt model from 10 weeks into the new season. They all capture the timinig of the peak fairly well.

Conclusions

- Bayesian FPCA provides a flexible framework
- Sprinkage priors allow the data to speak and chose min. basis
- Hierarchy structures are good
- Forecasting results are good

References

[1] C. M. Carvalho, N. G. Polson, and J. G. Scott, *Handling sparsity via the horseshoe*, in Proceedings of the Twelth International Conference on Artificial Intelligence and Statistics, vol. 5, 2009, pp. 73–80.

[2] J. Dauxois, A. Pousse, and Y. Romain, *Asymptotic theory for the principal component analysis of a vector random function: Some applications to statistical inference*, Journal of Multivariate Analysis, 12 (1982), pp. 136–154.

[3] T. Park and G. Casella, *The bayesian lasso*, Journal of the American Statistical Association, 103 (2008), pp. 681–686.