

Kernel of Truth

Evan L. Ray, Casey Gibson, Katie House,
Nicholas G. Reich



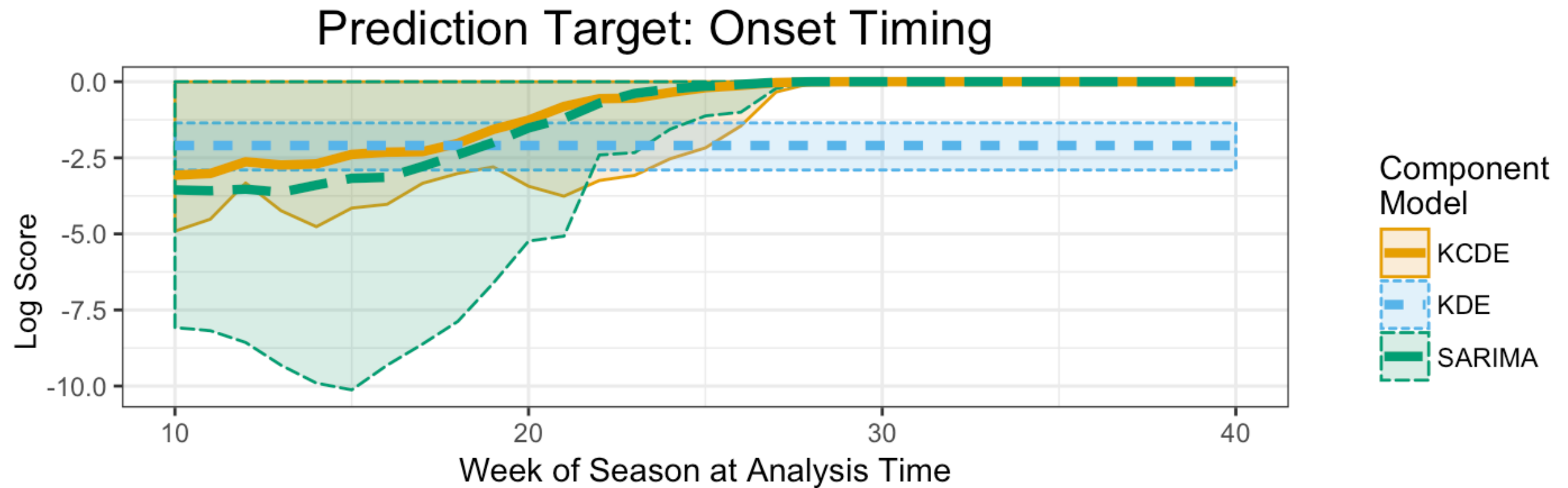
Reich
Lab | AT UMASS
AMHERST



Major Points

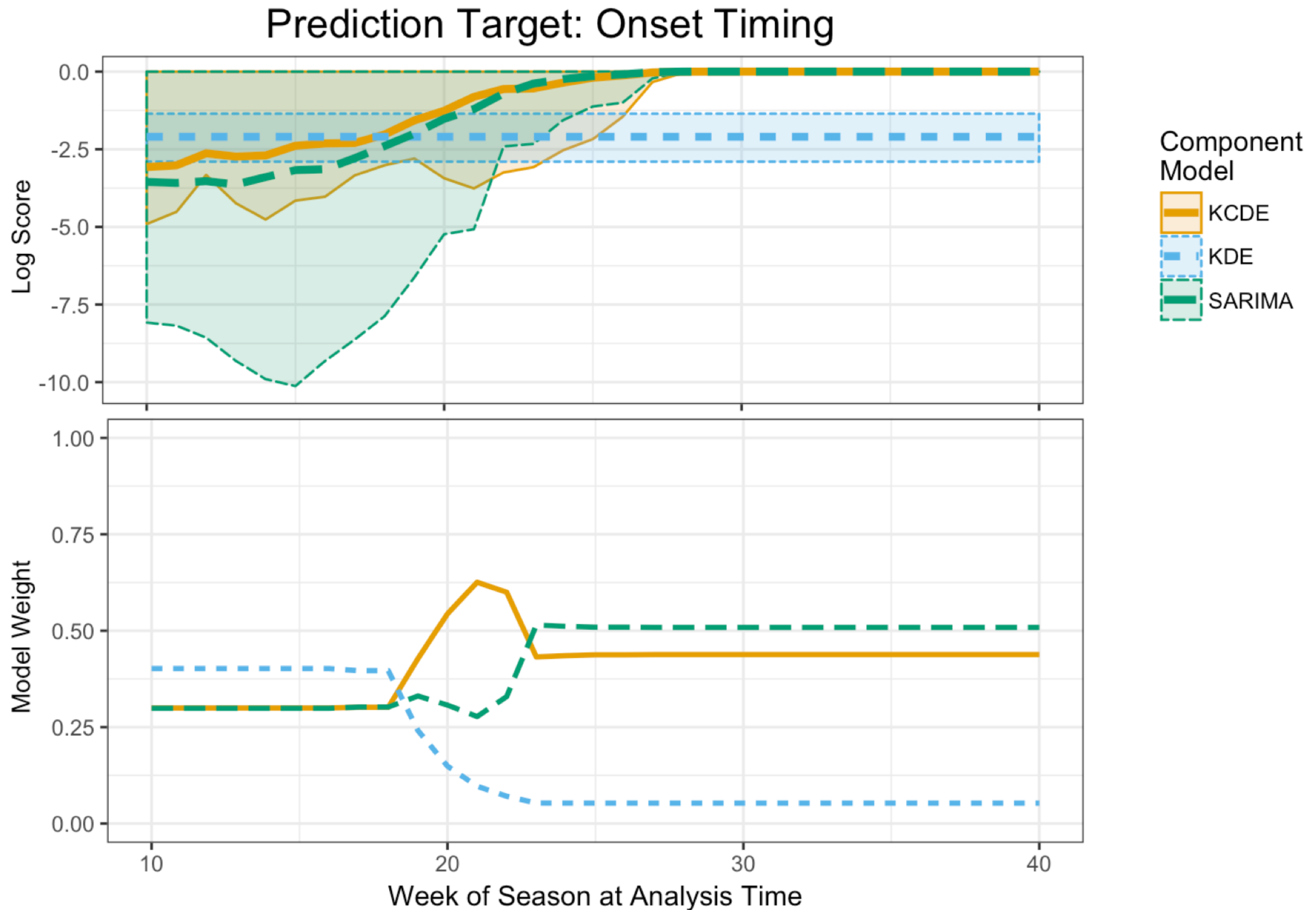
- Overview of our Methods (essentially same as last year)
- SARIMA model did well - implementation available

Component model performance by time of year



Goal: To construct an ensemble that can capitalize on these changing patterns of performance.

Weights Reflect Historical Performance



4 Component Models

- “KDE”
 - KDE estimate of distribution of seasonal targets
 - Seasonal splines for k-week ahead
- KCDE + copulas
 - KCDE gives non-parametric predictive distributions for future incidence in individual future weeks
 - Copulas estimate dependence across seasons
- SARIMA with seasonal differencing forced
- SARIMA without seasonal differencing forced
 - For both SARIMA models, Box-Cox transformation to approximate normality
 - Parameter estimates via R's `forecast::auto.arima`

Performance This Year - Regional

Rank	Team	Score	Skill
1	Delphi-Epicast	-1.0962	0.3341
2	FluSightNetwork	-1.1037	0.3317
3	LANL-DBMplus	-1.1166	0.3274
4	LANL-DBM	-1.1276	0.3238
5	CU Vixen	-1.1301	0.3230
*	UnwghtAvg	-1.1559	0.3148
*	PPFST Crowd	-1.1573	0.3143
6	CU-Network	-1.1690	0.3107
7	CU-SEL	-1.1947	0.3028
8	Delphi-Stat	-1.2040	0.3000
9	CU Puffins	-1.2196	0.2953
10	Kernel of Truth	-1.2422	0.2887
11	PredSci	-1.3191	0.2674
12	KBSI	-1.4761	0.2285
13	02115 emms	-1.4985	0.2235
14	HumNat	-1.6194	0.1980
15	02115-emmsa	-1.6826	0.1859
16	YaleModel	-1.9064	0.1486
*	Hist-Avg	-1.9539	0.1417

**Unofficial Scores for
Component Models**

**SARIMA with
seasonal differencing: -1.13**

**SARIMA without
seasonal differencing: -1.29**

KCDE + copula: -1.44

KDE: -1.91

Performance This Year - State

Rank	Team	Score	Skill
1	Delphi-Stat	-1.2928	0.2745
*	UnwghtAvg	-1.3759	0.2526
2	ColumbiaU	-1.4659	0.2309
3	KBSI	-1.5543	0.2113
4	ColumbiaU1	-1.5726	0.2075
5	Kernel of Truth	-1.7095	0.1810
*	Hist-Avg	-2.2129	0.1094

**SARIMA with
seasonal differencing**

sarimaTD Package for R

- Code for fitting and predicting from SARIMA models
- Based on the `auto.arima` function in the `forecast` package
- Adds extra things:
 - Simpler interface, less decision making
 - Ability to enforce seasonal differencing while still being able to generate predictions
 - Box-Cox transformations to normality
 - Interpolates missing data as needed when generating predictions (not relevant for recent flu seasons)
- Available on GitHub:
 - <https://github.com/reichlab/sarimaTD>
 - Includes vignette demonstrating use!

SARIMA Model (with seasonal differencing)

Coefficient Estimates - Regional

National

ar1	ar2	sar1	sar2
0.418	0.511	-0.677	-0.310

Region 1

ar1	sar1	sar2
0.810	-0.641	-0.346

Region 2

ar1	ar2	sar1
0.702	0.187	-0.443

Region 3

ar1	ar2
0.619	0.223

Region 4

ar1	ar2	ar3	sar1
0.885	0.116	-0.093	-0.525

Region 5

ar1	sar1	intercept
0.896	-0.594	0.031

Region 6

ar1	ar2	ar3	ar4	ar5	sar1
0.708	0.129	0.043	0.076	-0.116	-0.593

Region 7

ar1	ar2	sar1
0.645	0.233	-0.479

Region 8

ar1	sar1
0.798	-0.449

Region 9

ar1	ar2	sar1
0.590	0.279	-0.498

Region 10

ar1	ar2	ar3	sar1
0.634	0.116	0.109	-0.5109935