

STAT 448: Project Proposal

Paul Harmon

November 1, 2017

1 Introduction

I am interested in analyzing historical Elo ratings for NFL teams. FiveThirtyEight uses Elo ratings to quantify the strength of a particular team in any head-to-head matchup in the NFL; teams with higher Elo ratings are expected to beat those with lower Elo scores. Over time, overall NFL team strength has been non-constant, meaning that the best team in 2005 might not be better than the best team in 2015. Elo ratings are simple in terms of the information they require but they generally do well in predicting game outcomes. They are updated after every game so a team that outperforms expectations will gain points after an unexpected win and lose points after a loss.

1.1 Research Questions and Methodology

I'm curious to compare within-season dynamics of teams after controlling for long-term yearly variation in team quality. My basic research question is: Can team behavior be categorized into three or four groups based on whether teams "overachieve" or "underachieve" during a given season?

The method I will use to answer these questions is to compare functional patterns within each year for each team. As such, I will use GAMs to smooth the 16-game Elo curve for each team in each season and then try to classify Team/Year combinations into multiple groups. At the end of the analysis, I want to identify, for instance, that the 2011 Broncos were "overachievers" and the 2016 Minnesota Vikings were "underachievers" based on Elo patterns - Figure 2 in the appendix shows a plot of what I expect these would look like. Although this gets into some more functional analysis stuff than we covered in Mixed Models, this hinges on GAM-smoothed estimates of within-season Elo patterns for clustering.

1.2 The Data

These intensively-sampled functional data are available on FiveThirtyEight's Github repository. I have a dataset that contains the variables of interest: Year, Game, Elo rating per game, and Team. The data contain Elo ratings for all 32 NFL teams throughout the league's history. However, since some of the historical teams in the dataset no longer exist, I am only interested in looking at the data pertaining to teams and games since the Houston Texans joined the league in 2002. Elo ratings are scored from a baseline of 1300; the minimum Elo in the dataset was 1120 and the maximum was 1849. Team Elo means and SD's are shown in the Table 2. There are 32 teams in the NFL and the dataset contains information for 15 seasons, each with 16 regular-season games. There are no missing observations, meaning that $n = 7680$.

Table 1: NFL Elo Data Characteristics

Var	Type	n
Year	Categorical	15
Game	Quantitative (time)	16
Elo	Quantitative	7680
Team	Categorical	32

2 Model

Right now the model would look something like the following:

$$Y_i = \beta_0 + s(WEEK_i) + Year_{i,j} \quad (1)$$

Y refers to the predicted Elo, the main predictor is a smooth function of within-season weeks, and a random intercept term accounts for between season differences. It may be worthwhile to go farther than just a random intercept model - but this is a good starting point. Team scores are normalized at the start of every season with some corrections made to shrink team Elos towards the mean - this is why I think it might be better to estimate within-season means rather than treating each year as continuous. Figure 1 shows the long-term trend with dots representing each game during each season for one NFL team.

3 Final Thoughts and Possible Issues

While one could model linear regressions and look at a long-term trend in Elo, the cyclical nature of the data mean that in the long run, most teams would look pretty flat. Considering the inherent grouping of observations by including year-level random effects is also useful, as is the use of GAMS to smooth out a general trend during each season for each team.

Another challenge involved in this project is that I'm fitting a GAM for each team in each year. I'm not sure if fitting a mixed-effects GAM would achieve quite the same thing, or if that is reasonable given that the data are normalized between years. However, I think that this project seems reasonable. It involves aspects of mixed models, hierarchical data, and includes some forays into functional data analysis that would be interesting as well.

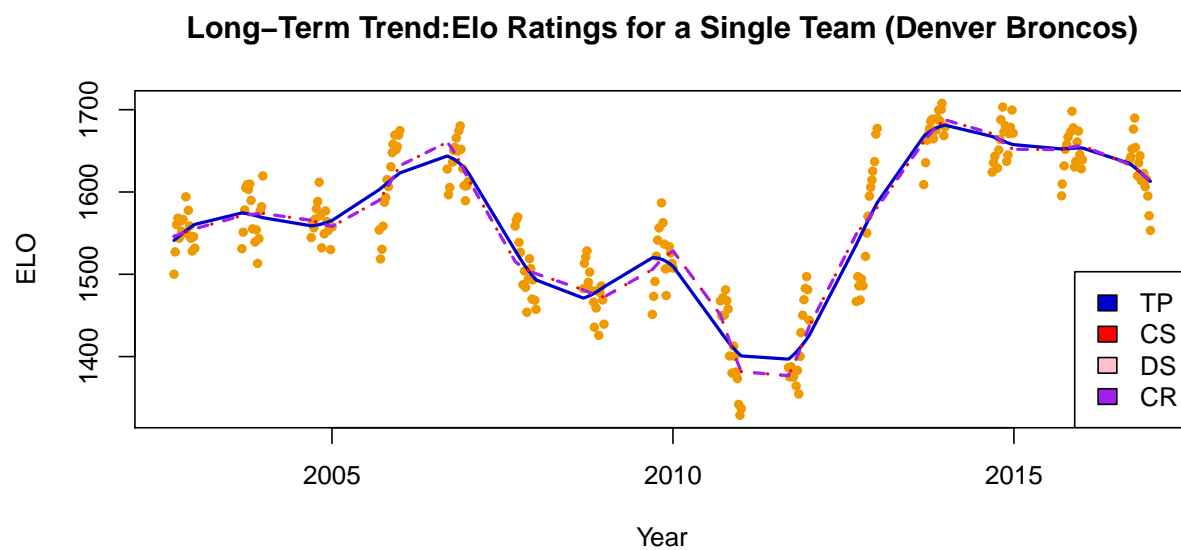


Figure 1: Plot of Long-Term Trend

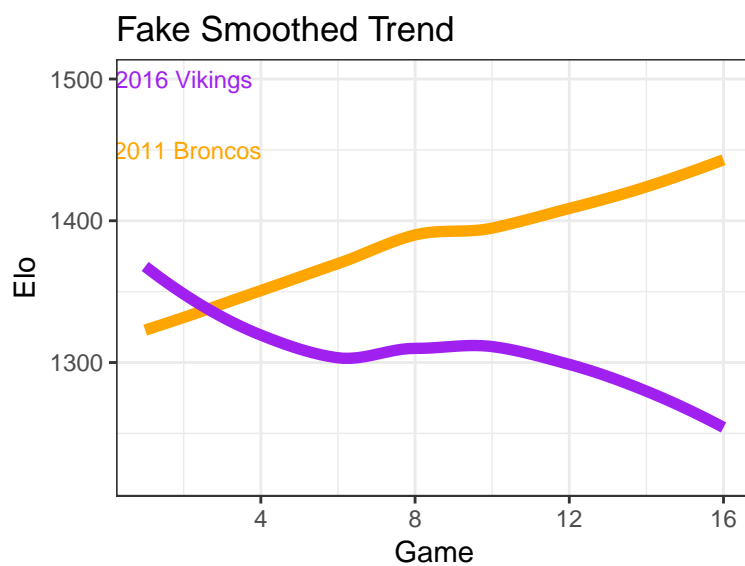


Figure 2: Possible Representative within-season curves

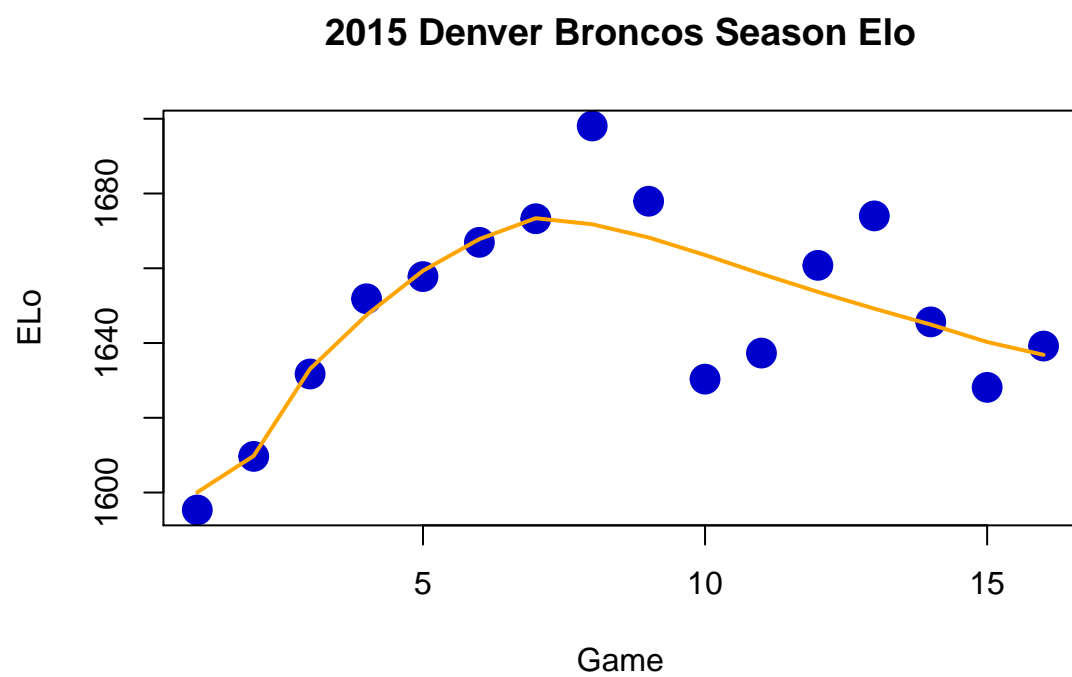


Figure 3: Yearly Smoothed Trends for 2015

Table 2: Mean and SD Elo by Team

	Team	Mean Elo	SD Elo
1	ARI	1466.20	96.53
2	ATL	1514.75	75.29
3	BAL	1556.82	58.80
4	BUF	1475.79	50.49
5	CAR	1503.91	80.86
6	CHI	1494.22	72.65
7	CIN	1500.66	88.70
8	CLE	1411.08	52.59
9	DAL	1520.59	65.72
10	DEN	1562.55	89.35
11	DET	1417.12	85.84
12	GB	1575.36	78.56
13	HOU	1454.69	94.52
14	IND	1583.12	102.55
15	JAX	1437.99	99.73
16	KC	1494.32	98.19
17	LAC	1535.66	81.21
18	LAR	1434.68	88.48
19	MIA	1482.12	59.58
20	MIN	1497.33	61.95
21	NE	1665.05	55.79
22	NO	1525.53	79.28
23	NYG	1523.61	66.22
24	NYJ	1496.84	67.80
25	OAK	1420.34	86.46
26	PHI	1553.37	71.80
27	PIT	1585.42	57.87
28	SEA	1536.45	93.94
29	SF	1474.99	110.91
30	TB	1471.75	83.03
31	TEN	1479.76	94.25
32	WSH	1453.02	54.56