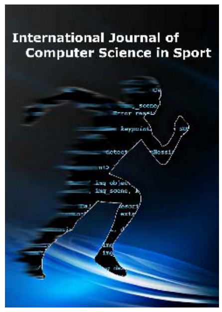
International Journal of Computer Science in Sport

Volume 16, Issue 3, 2017

Journal homepage: http://iacss.org/index.php?id=30

DOI: 10.1515/ijcss-2017-0014



A Logistic Regression/ Markov Chain Model for American College Football

Kolbush, J., Sokol, J.

H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology

Abstract:

Kvam and Sokol developed a successful logistic regression/Markov chain (LRMC) model for ranking college basketball teams part of Division I of the National Colligate Athletic Association (NCAA). In their 2006 publication, they illustrated that the LRMC model is one of the most successful ranking systems in predicting the outcome of the NCAA Division I Basketball Tournament. However, it cannot directly be extended to college football because of the lack of home-and-home matchups that LRMC exploits in performing its Logistic Regression. We present a common-opponents-based approach that allows us to perform a Logistic Regression and thus create a football LRMC (F-LRMC) model. This approach compares the margin of victory of home teams to their winning percentage in games played against common-opponents with the away team. Computational results show that F-LRMC is among the best of the many ranking systems tracked by Massey's College Football Ranking Composite.

KEYWORDS: LOGISTIC REGRESION, MARKOV CHAIN, AMERICAN COLLEGE FOOTBALL, COMMON GAME, MARGIN OF VICTORY

Introduction

College football is a difficult sport to model for a variety of reasons. Teams only play 11-13 games per season, yet there are 128 Football Bowl Subdivision (FBS, formerly NCAA Division I-A) teams to be ranked. Most teams only play 3-4 games outside their conference, making it difficult to compare teams from different conferences. For these reasons, much debate exists over the best way to rank teams, and many different polls and models exist that attempt to answer this question.

The official poll used by the National Collegiate Athletic Association (NCAA) is the College Football Playoff Rankings (CFP). Previously, the Bowl Championship Series (BCS) served as the official ranking system from 1998 to 2013. In addition to this, the Associated Press (AP) Poll and USA Today Coaches Poll are two of the oldest college football ranking systems and are still followed by many (ESPN, n.d.). In addition to polls, many computer models exist in order to provide rankings based on statistical measures. The BCS was notable for incorporating several computer models into its rankings, including successful models by Sagarin, Colley, and Billingsley (Massey, n.d.).

Kvam and Sokol (2006) developed a ranking system using a combination of Logistic Regression and a Markov Chain (LRMC) for college basketball using only “scoreboard data.” That is, for each game, the only information taken into consideration was the names of the winning and losing team, the margin of victory of the winning team, and whether the game was played on the winners home court, the losers court, or a neutral location. Their model is constructed by creating a Markov Chain between all teams in Division I NCAA basketball teams. In order to determine the transition probabilities between the states of the Markov Chain, a Logistic Regression had to be performed. This combination of Markov Chain and Logistic Regression resulted in one of the most accurate College Basketball ranking systems. Their model’s accuracy was evaluated by analyzing the results of the NCAA Division I Basketball Tournament. They found the average rank of the teams that advanced to the later rounds of the tournament was significantly lower than the average rank of the teams when ranked using other models ($p < 0.05$ when compared against AP, Seed, Massey, Sagarin, KG, and Sheridan predication methods). Modified versions of LRMC have been developed for college sports rankings by Brown and Sokol (2010), who used an empirical Bayes approach, and by Maclay (n.d.), who used natural logs of margin of victory rather than the margins themselves. Outside of sports, LRMC models been applied to modeling urban sprawl and population dynamics as demonstrated by Hamdy et al. (2016) and Liu et al. (2015).

In the LRMC for NCAA basketball, determining the transition probabilities between the teams relies on analysis of games where teams play each other twice in the same season, once at each team’s home court. While many of these “home-and-home” matchups occur each season in college basketball, they rarely occur in college football. Therefore, it is not immediately clear how to construct an analogous model for NCAA football. In order to determine the transition probabilities, we present a replacement for these home-and-home games by instead looking at the games played between common opponents that pairs of teams face.

After describing how the model is constructed, we measure its accuracy by counting the number bowl games it predicted correctly for each given year and comparing with other ranking systems. From this analysis we find that our Football LRMC (F-LRMC) model ranks amongst the most accurate predictors of bowl game results.

Methods

Markov Chain

Our base Markov chain model follows that of Kvam and Sokol's (2006). There is one state in the Markov chain for each team in Division I NCAA football. Transitions are made according to the outcome of games played during the season. Let $r_{x(g)}$ be an estimate of the probability that the home team of game g is better than the away team of game g given that the home team won by $x(g)$ points, where $x(g)$ is negative if the home team lost. We then define $r_{x(g)}$ for each game $g = (i,j)$ where i is the visiting team and j is the home team. Then, if N_i is the number of games played by team i, we can define the transition probabilities from state i to be

$$t_{ij} = \frac{1}{N_i}\left[\sum_{g=(i,j)} r_{x(g)} + \sum_{g=(j,i)}\left(1 - r_{x(g)}\right)\right], \qquad \text{for all } j \neq i, \tag{1}$$

and

$$t_{ii} = \frac{1}{N_i}\left[\sum_{g=(\cdot,i)} r_{x(g)} + \sum_{g=(i,\cdot)}\left(1 - r_{x(g)}\right)\right]. \tag{2}$$

We can use these transition probabilities to solve for the Markov chain's stationary distribution. By ordering this stationary distribution in decreasing order, we obtain a ranking of the college football teams.

The difficulty here lies in obtaining values for $r_{x(g)}$ (often denoted as simply r_x). By using methods described in the following section, we can obtain data points that serve as an approximation for r_x. It is expected that r_x should be an increasing function; the more points a team wins a game by, the higher the probability that that team is better than its opponent. For this reason, the data should be fit to an increasing function that approaches 0 as x decreases and approaches 1 as x increases. By performing a logistic regression, the data is fit to such a function.

Logistic Regression

We now discuss approximating the function r_x. The function can be thought of as "given that the home team won by a margin of x, what is the probability that they are better than the away team." In Kvam and Sokol's (2006) LRMC model for college basketball, this function was approximated by a logistic regression analysis of "home-and-home" games, pairs of games where two teams play each other twice within the same season, once at each team's home-court. Of all the teams who won by x points at home, the fraction f_x who beat the same opponent on the road was recorded and a logistic regression was used to smooth the data. In this way, direct comparisons between teams could be made: after the x-point home game, the result between the same two teams was used directly to calculate the estimate of r_x.

However, this direct approach cannot be used with college football, because teams rarely play each other twice during a single college football season; we have found only about 40 instances in the 20 years from 1996-2015. Therefore, in order to approximate r_x for college football, a new estimation approach must be introduced.

A Replacement for r_x

For our estimate, rather than a direct comparison, we use an indirect approach using common opponents. Let G_x be the set games in a season which the home team won by x points. For each $g = (i,j) \in G_x$, the common opponents are the set of teams $C(g)$ that played against both home team j and away team i during the season. Let $G_i(g)$ be the set of games played between team i and each $k \in C(g)$; likewise, let $G_j(g)$ be the set of games played between team j and each $k \in C(g)$. Define $\mu(G_j(g))$ to be the number of games team j won in $G_j(g)$ and $\nu(G_i(g))$ to be the number of games team i lost in $G_i(g)$. From here we arrive at an estimate for r_x defined as

$$r_x \approx \frac{\sum_{g \in G_x}\left(\mu\left(G_j(g)\right) + \nu\left(G_i(g)\right)\right)}{2\sum_{g \in G_x}|\{C(g)\}|} \tag{3}$$

In simple terms, given the winning percentages p_i and p_j of teams i and j against their common opponents,

$$\frac{p_j + (1 - p_i)}{2} \tag{4}$$

is an estimate of the probability that i is better then j. To find r_x, we take this collective estimate over all games with an x-point win for the home team.

Similar approaches based on results from common opponents have been used by other models. Knottenbelt et al. (2012) presents a stochastic model designed to predict the result of tennis matches. In order to establish the advantage one tennis player has over his opponent, they compare the proportion of points won between the two players and their common opponents.

The values r_x generated by (3) for the 2011-2014 seasons are plotted in Figure 1. As in Kvam and Sokol's (2006) model, we use logistic regression to smooth the data; the curve shown in Figure 1 is the logistic regression function for 2011-2014.

Because college football styles of play change over time, we use a rolling four-season window to compile data. So, for example, for the 2015 season we build our estimate of r_x using data from the 2011 to 2014 seasons. Multiple training set sizes were analyzed and the four-season window provided the strongest results. Too small of a training set resulted in outliers in the data having too great of an impact. Conversely, a large window of seasons resulted in outdated game results shaping the regression.

The logistic regression model finds values (a, b) such that $r_x = \frac{e^{(ax+b)}}{1+e^{ax+b}}$ is best fit to the data. We constructed our model for 15 seasons of college football, 2002 – 2016. For each of these seasons, we used the previous four seasons as the training set. The values for a and b for each season are shown in Table 1. The margin of victory was shown to be statistically significant as a predictor of r_x for all seasons ($p < 0.001$, Wald Test).

Typically, a logistic regression is performed to classify binary data. However, this regression is used to establish the transition probabilities in the Markov chain, thus the logistic regression is not used for direct classification so no cutoff value is needed. The result of the logistic regression can be used to answer the question "given that team j beat team i by x points, what are the odds that team j is better than team i." Rather than simply comparing two teams, the Markov chain allows us to use this result to compare all the teams and obtain a full set of rankings.

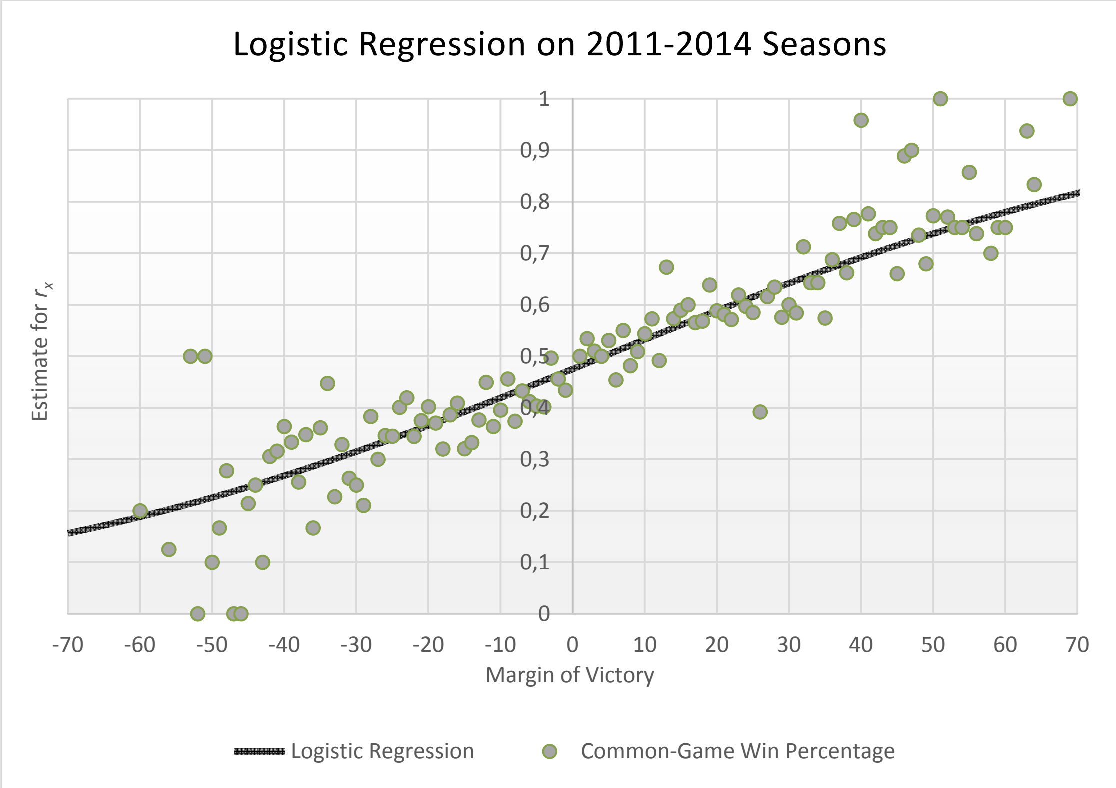


Figure 1: The common-game winning percentage and the logistic regression through the data for the 2011-2014 regular seasons

Results

We measure the accuracy of our model by training it during the regular season in order to predict the outcomes of the NCAA postseason bowl games. We can conclude that those models who are most accurately predict bowl games provide the most accurate rankings of the teams. Bowl games are an ideal measure for determining the accuracy of a ranking system for three main reasons. First, they take place at the end of the season, allowing a full season's worth of data to be taken into consideration. Second, each of the bowl games are played between teams of approximately similar strength, allowing for there to be disagreement between models on who the expected winner should be. Third, all the bowl games are played on neutral fields (i.e., neither team is playing at its home stadium), meaning that the higher-ranked team should be favored to win.

This last assumption is often not true in games not played in a neutral location, because the advantage of playing at home might outweigh a small difference in team strength. Not all models offer means to predict non-neutral-site games; many, including the F-LRMC, have the exclusive function as ranking systems. However, neutral-site games do not require a separate predictive model as we can rather infer the predictions directly from the rankings. We also make the assumption that the field is truly neutral, despite some bowl games' locations being slightly favored towards one team.

Table 1: The logistic regression parameters for each four-season range, along with the standard error for each of the values.

Logistic Function Parameters by Season			
Season	**Training set**	***(a,b)***	**Std. error for *(a,b)***
2016	2012-2015	(.02375, -.09720)	(.00058, .01172)
2015	2011-2014	(.02269, -.09786)	(.00058, .01171)
2014	2010-2013	(.02223, -.09632)	(.00056, .01163)
2013	2009-2012	(.02207, -.09680)	(.00057, .01151)
2012	2008-2011	(.02130, -.08730)	(.00056, .01149)
2011	2007 -2010	(.02067, -.08427)	(.00056, .01139)
2010	2006-2009	(.02091, -.09022)	(.00058, .01138)
2009	2005-2008	(.01978, -.08089)	(.00058, .01154)
2008	2004-2007	(.02012, -.09919)	(.00061, .01181)
2007	2003-2006	(.02051, -.11819)	(.00060, .01197)
2006	2002-2005	(.02117, -.12753)	(.00060, .01218)
2005	2001-2004	(.02164, -.13514)	(.00059, .01201)
2004	2000-2003	(.02205, -.12919)	(.00058, .01208)
2003	1999-2002	(.02177, -.12064)	(.00059, .01224)
2002	1998-2001	(.02254, -.12064)	(.00060, .01229)

Comparison with other Computer Models

Massey (n.d.) maintains an archive of dozens of college football rankings for each week of the season. By looking at the rankings of various models before the bowl games, we can compare the predictive accuracies of various models. We consider a system to have predicted a game correctly if it has the winning team ranked higher than the losing team, and we count the total number of bowl games that each model predicted correctly. Trono (2012) provided this comparison for models appearing on Massey's archive during the 2002-2011 seasons. Using game data from Forman (n.d.), we expanded the results of Trono (2012) to include the 2012-2016 seasons, as well as to include our Football LRMC (F-LRMC). The results are included in Table 2. As shown in Table 2, the F-LRMC has a bowl game prediction accuracy of 60.79%, which ties it for third amongst all models that appear in Massey's composite for the 2002-2016 seasons.

Additionally, we tested for statistical significance between F-LRMC and each other model using McNemar's test. Those models with a p-value less then .05 appear in bold in Table 2. It should be noted that Kambour (KAM) and PerformanZ (PFZ), the two models with better results then the F-LRMC, fail to be statistically significant over the F-LRMC (with p-values of .1282 and .1933, respectively).

Kambour and PerformanZ are the two models that outperform the F-LRMC, with bowl game accuracies of 63.17 and 62.57, respectively. Each of these models generate their rankings very differently then the F-LRMC. Kambour's (2003) model is based on the idea that teams that are historically good tend to stay good. While the F-LRMC only looks at that season's data to generate the model, Kambour takes into account previous seasons' data. Furthermore, the PerformanZ model, constructed by Beck (2002), is centered around the idea that in game statistics, not game results, are the strongest indication of who the best teams are. So while the LRMC uses only scoreboard data, PerformanZ accounts for several other statistics such as measures of a team's run and pass offense and defense. These differences between models influence the requirements of their implementation. Compared to F-LRMC, more seasons' data is needed for Kambour's rankings and additional statistics are needed to implement PerformanZ. Furthermore, the inclusion of certain measures, such as margin of victory, within a model is heavily debated. Some analysts may not want to include previous seasons' data, as Kambour's models does, for they believe that the rankings should only reflect a team's performance for the current season. Likewise, the inclusion of a statistic such as pass offense may be biased against teams who are effective at running an offense with little passing.

The only models included in Table 2 are those that appear in Massey's (n.d.) composite the final week before the bowl games occur every year. If a model's site doesn't publish rankings for that week they are not included in the composite. Several models appear in all but one or two years. We can estimate how many games a model will correctly predict in a year when its rankings were not published by looking at the average of the percent difference between the number of bowl games that model correctly predicts and the average number of bowl games predicted by all models. Table 3 compares the F-LRMC with all models that missed only one or two years and estimates the number of bowl games that model would have predicted for the year(s) that are missing. The F-LRMC ranks near the top when compared to these models as well. The F-LRMC was outperformed by three models, CPA, ARGH (ARG), and Kislanko Isof (KLK). All three of these did not show statistical significance over the LRMC (with p-values of .1945, .3773, and .4122, respectively).

Table 2: The number of bowl games correctly predicted by all models that appear in Massey's composite (n.d.) for the 2002 – 2016 seasons. Models that appear in bold signify a statistically significant difference from F-LRMC .

Number of Correctly Predicted Bowl Games, by Model, by Year

System name	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	Total	Correct %	p-value
No. of Games	28	28	28	28	32	32	34	34	35	35	35	35	39	41	41	505		
KAM	13	21	16	15	19	28	23	22	23	23	22	21	21	28	24	319	63.17	
PFZ	15	18	15	19	20	23	21	22	21	23	25	17	21	32	24	316	62.57	
F-LRMC	15	20	18	13	22	23	21	18	25	23	22	22	19	27	19	307	60.79	
BSS	11	18	20	18	18	19	24	14	22	21	22	22	22	32	24	307	60.79	0.5355
MOR	15	17	15	15	22	17	24	18	24	21	25	20	22	26	23	304	60.20	0.4196
BORN	14	20	14	17	21	20	22	18	22	19	25	19	19	29	23	302	59.80	0.3323
COF	14	20	18	18	22	21	18	18	22	22	20	19	20	30	19	301	59.60	0.3069
WLK	14	18	16	14	23	21	15	20	21	23	22	21	21	28	24	301	59.60	0.2906
PIG	13	19	16	17	17	17	22	18	22	23	24	19	25	24	24	300	59.41	0.2863
SAG[1]	14	19	15	15	20	21	17	15	21	24	22	20	22	30	24	299	59.21	0.2111
HOW	16	18	15	16	23	20	16	19	23	23	20	16	22	28	22	297	58.81	0.1841
MAR	14	21	16	17	19	20	15	20	22	24	23	15	19	27	23	295	58.42	0.1231
MAS[2]	14	19	15	15	22	18	16	18	18	22	22	21	24	28	20	292	57.82	0.0795
DOL	15	21	17	17	22	21	17	17	18	23	17	19	22	25	20	291	57.62	0.0871
ASH	14	17	15	14	22	19	20	16	20	22	19	21	24	28	20	291	57.62	**0.0447**
SOL	15	17	16	15	19	22	15	15	21	23	23	20	23	26	21	291	57.62	0.0817
Avg.	14	18	14	15	23	20	16	16	21	21	21	21	21	28	22	291	57.62	0.6884
MRK	15	18	16	16	20	17	15	19	20	21	20	18	24	25	23	287	56.83	**0.0131**
BIH	15	19	15	14	20	18	18	16	22	23	18	23	22	24	19	286	56.63	**0.0225**
RTH	14	18	15	15	19	22	18	15	20	23	20	19	23	25	20	286	56.63	**0.0241**
SEL	14	20	16	14	18	21	18	16	19	22	20	20	24	24	20	286	56.63	**0.0230**
BIL[2]	14	20	17	17	20	16	15	19	22	18	17	20	20	27	23	285	56.44	**0.0430**
COL[2]	16	14	18	16	21	21	16	13	20	22	16	16	24	31	18	282	55.84	**0.0245**
WEL	15	15	16	16	22	22	15	15	19	22	16	16	25	28	19	281	55.64	**0.0198**
WIL	15	18	16	15	20	21	15	17	21	22	17	18	22	24	18	279	55.25	**0.0092**
MJS	15	15	16	17	21	19	15	14	22	21	17	16	25	27	19	279	55.25	**0.0146**
DES	15	17	17	17	18	17	12	20	21	24	19	17	17	25	22	278	55.05	**0.0013**
AND[1]	15	14	16	14	21	21	16	18	20	22	16	20	21	24	19	277	54.85	**0.0063**
WOL[2]	15	17	12	16	21	18	14	16	19	21	18	23	23	24	19	276	54.65	**0.004**
WOB	16	18	14	15	21	18	16	15	20	22	15	20	22	24	19	275	54.46	**0.003**
CSL	16	17	15	16	21	17	15	16	22	20	16	14	24	24	19	272	53.86	**0.0053**

1- Specifies that model was used in calculations of the BCS Standings from 1998-2013
2- Specifies that model was used in calculations of the BCS Standings from 2004-2013

Table 3: The number of bowl games correctly predicted by all models appearing in Massey's composite (n.d.) for 2002-2016 seasons, sans one or two years. Entries that appear in italics were estimated and rounded to the nearest whole number. Models that appear in bold signify a statistically significant difference from F-LRMC

Number of Correctly Predicted Bowl Games, by Models with Missing Data, by Year

System Name	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	Total	Correct %	p – value
No. of Games	28	28	28	28	32	32	34	34	35	35	35	35	39	41	41	505		
CPA	15	18	15	20	*21*	23	24	20	22	24	23	20	20	30	*22*	316.6	62.69	
ARG	15	19	16	16	24	*19*	15	21	23	22	23	19	21	32	22	307.34	60.86	
KLK	*15*	*18*	18	15	20	24	19	23	22	24	23	20	19	27	21	307.28	60.85	
F-LRMC	15	20	18	13	22	23	21	18	25	23	22	22	19	27	19	307	60.79	
DP	*14*	19	15	14	21	22	19	21	23	22	24	19	20	26	25	304.48	60.29	0.4538
LAZ	*15*	20	16	16	21	20	18	17	22	*22*	22	19	24	29	21	301.54	59.71	0.3688
DUN	14	17	*16*	17	*20*	18	24	16	24	24	20	17	21	29	23	300.01	59.41	0.4142
DWI	15	18	19	21	21	18	17	*18*	23	21	21	19	17	*28*	23	298.61	59.13	0.2317
DOK	*14*	19	16	14	20	19	15	17	23	23	23	19	21	31	24	298.39	59.09	0.1981
CGV	*14*	19	18	16	19	17	15	18	22	25	20	16	25	29	22	295.3	58.48	0.2095
BO	*14*	18	13	17	20	22	19	17	22	20	22	20	21	29	*21*	295.07	58.43	0.0827
KEE	*14*	*17*	18	14	21	20	19	18	21	22	22	17	21	25	22	291.52	57.73	0.1493
MAU	13	19	17	15	20	17	17	20	21	22	20	21	21	*26*	*20*	289.67	57.36	**0.0252**
CPR	14	18	17	18	*20*	18	16	19	20	20	19	24	19	27	*21*	289.45	57.32	0.0711
RUD	14	18	15	15	21	19	*17*	16	22	23	20	21	23	24	21	288.55	57.14	0.0851
MCK	16	21	22	16	20	20	15	14	19	21	19	*19*	*21*	25	18	286.03	56.64	**0.0322**
MAA	*14*	15	15	17	23	19	17	15	20	23	17	21	23	26	*21*	285.52	56.54	**0.0234**
JNK	*14*	*17*	16	15	21	19	16	18	21	21	20	16	22	30	18	283.97	56.23	**0.0464**
CMV	*14*	*17*	19	17	19	19	18	14	17	25	21	21	18	23	21	283.34	56.11	**0.0322**
GBE	*14*	*17*	16	15	23	19	13	16	20	23	19	15	26	28	19	283.04	56.05	0.0502
MEA	16	16	16	16	22	18	17	15	19	21	18	20	22	26	*20*	282.36	55.91	**0.0110**
SE	14	18	13	16	20	20	15	17	21	23	16	20	23	*26*	*20*	282.16	55.87	**0.0083**
D1A	*14*	*17*	16	14	22	20	14	17	21	24	16	16	22	24	21	277.64	54.98	**0.0157**
SOR	14	18	15	15	20	20	17	15	19	22	18	*18*	21	22	21	275.25	54.50	**0.0039**

Comparison against Polls

There are two major college football ranking polls that have been used for many years, the Associated Press (AP) Poll and the USA Today Coaches Poll. Each of these polls ranks only the top 25 teams each week. We report the accuracy of our model compared to these polls as before, but we can only take into account games that include a ranked team in the poll. Thus, to compare F-LRMC against the AP Poll we compared the number of bowls each of the ranking systems got correct, only in those games where at least one team was ranked in the AP Poll's top 25. The same approach was used for the Coaches Poll. The results are shown in Table 4, using poll data taken from the 2002 to 2016 seasons (NCAA College Football Polls – ESPN, n.d.). F-LRMC correctly predicted more bowl games the either of the Polls, but the difference was not statistically significant.

We can furthermore compare the results from the BCS and CFP rankings to those from the F-LRMC. Prior to 2003, the BCS rankings would only include the top 15 teams. So for purposes of uniformity, we will compare the results from 2003 to 2013. As with the AP and Coaches Poll, the F-LRMC outperformed the BCS rankings, but not enough for statistical significance. The BCS was replaced in 2014 with the new College Football Playoff rankings (Selection Committee Protocol, 2015). We only have three years of results from the CFP but currently the F-LRMC has predicted 25 bowl games correctly while the CFP has predicted 26. These results are likewise included in Table 4.

Table 4: The number of bowl games correctly predicted by the major polls and the number of bowl games F-LRMC correctly predicted, only considering games in which at least one team appeared in the poll's top-25.

Number of Correctly Predicted Bowl Games, by F-LRMC and Polls, by Year

System name	'02	'03	'04	'05	'06	'07	'08	'09	'10	'11	'12	'13	'14	'15	'16	Total	Correct %
No. of Games	17	15	15	16	17	16	14	16	16	15	16	16	16	17	17	238	
F-LRMC[1]	11	9	10	6	12	10	8	9	11	12	11	9	7	11	8	142	59.66
AP	8	11	13	9	12	7	4	8	10	13	10	9	8	11	7	139	58.40
No. of Games	17	15	14	16	17	16	14	16	17	16	15	16	16	17	17	238	
F-LRMC[1]	11	9	10	6	12	10	8	7	11	13	10	9	7	11	8	141	59.24
Coaches	8	10	12	9	12	6	4	7	10	13	10	9	8	11	8	136	57.14
No. of Games		17	14	15	17	17	15	16	17	16	16	16				176	
F-LRMC[1]		11	10	6	12	10	8	8	11	14	11	9				110	62.50
BCS		13	9	8	13	7	4	6	11	12	10	9				102	57.95
No. of Games													16	15	18	49	
F-LRMC[1]													7	10	8	25	51.02
CFP													9	8	9	26	53.06

[1] – Only considering bowl games including a team ranked by the respective poll in their top 25

Discussion

The Importance of Margin of Victory

The F-LRMC relies heavily on margin of victory in its construction. A version of the F-LRMC can be constructed that does not consider margin of victory, but only considers whether the game was won or lost by the home team. This version is much less accurate, only predicting 53.66% of bowl games between 1998 and 2016, while the original F-LRMC had a prediction accuracy of 60.79%.

However, the inclusion of margin of victory is a highly debated topic in college football. In 2002, the BCS changed its policy to no longer consider margin of victory in its rankings (Palm, 2013). As consequence of this, several of the computer models that were used in the BCS were either removed or changed so that they no longer considered margin of victory. Furthermore, the CFP have also indicated that they do not consider margin of victory in its rankings (Collegefootballplayoff.com, 2012). The motivation behind this non-inclusion is to prevent teams from running up the score during games. While this is fine reason for the CFP to not consider margin of victory, the F-LRMC has shown that its inclusion creates a far more accurate model.

Conclusion

We have presented a method to create a logistic regression/Markov chain model for ranking college football teams. The main difficulty in creating such a model was the lack of home-and-home games that were exploited by Kvam and Sokol (2012) in their development of an LRMC model for college basketball. We overcame this difficulty by examining the common opponents that teams play in a given season. Similar approaches to the F-LRMC may be applied in other sports that lack home-and-home games.

Computational testing shows that our new football LRMC (F-LRMC) model is, like the original LRMC, among the best ranking systems in college football for predicting postseason bowl games.

References

BCS computer rankings. (2012). Retrieved December 28, 2016, from http://www.bcsfootball.org/news/story?id=4765872

Beck, T. (2002). About the PerformanZ ratings. Retrieved December 28, 2016, from http://tbeck.freeshell.org/fb/descript.txt

Brown, M. and J. Sokol (2010). An Improved LRMC Method for NCAA Basketball Prediction. *Journal of Quantitative Analysis in Sports*, *6* (3), Article 4.

Forman, S. (n.d.) Sports Reference. Retrieved from http://www.sports-reference.com/cfb/

Hamdy, O., Shichen, Z., Osman, T., Salheen, M. A., & Eid, Y. Y. (2016). Applying a Hybrid Model of Markov Chain and Logistic Regression to Identify Future Urban Sprawl in Abouelreesh, Aswan: A Case Study. *Geosciences*, *6*(4), 1-17. doi:10.3390/geosciences6040043

Kambour, E. (2003). PPT. Edward Kambour. Retrieved from http://www.kambour.net/football.ppt

Knottenbelt, W. J., Spanias, D., & Madurska, A. M. (2012). A common-opponent stochastic model for predicting the outcome of professional tennis matches. *Computers & Mathematics with Applications*, *64*(12), 3820-3827.

Kvam, P. and J.S. Sokol (2006). A logistic regression/Markov chain model for NCAA basketball. *Naval Research Logistics*, *53*, 788-803.

Liu, Y., Dai, L., & Xiong, H. (2015). Simulation of urban expansion patterns by integrating auto-logistic regression, Markov chain and cellular automata models. *Journal Of Environmental Planning & Management*, *58*(6), 1113-1136. doi:10.1080/09640568.2014.916612

Massey, K. (n.d.). Massey Ratings. Retrieved from http://www.masseyratings.com/

Maclay, L.A. (n.d.). Retrieved February 14, 2017 from https://bracketology.engr.wisc.edu/ncaa-bb-rankings/

NCAA College Football Polls - ESPN. (n.d.). Retrieved November 10, 2016, from http://www.espn.com/college-football/rankings

New Formula for Football Championship Announced. (1998). Retrieved April 26, 2012, from http://www.umterps.com/sports/m-footbl/spec-rel/061098aaa.html

Palm, J. (2013). Sagarin changes formula, finally removes 'Margin of Victory.' Retrieved April 26, 2017, from http://www.cbssports.com/college-football/news/sagarin-changes-formula-finally-removes-margin-of-victory/

Selection Committee Protocol. (2015). Retrieved January 10, 2017, from http://www.collegefootballplayoff.com/selection-committee-protocol

Trono, J. (2012). Bowl Game Predictions. Retrieved October 3, 2016, from http://academics.smcvt.edu/jtrono/OAF_BCS/Compare.html