When Perfect Isn't Good Enough: Retrodictive Rankings in College Football

R. Drew Pasteur

Abstract. *Mathematical ranking systems, such as those used in college football, can be classified in two broad categories. Predictive methods seek to forecast outcomes of future games, while retrodictive rankings aim to most closely match the results of contests already played. Ideally, a retrodictive method would order teams such that each team is ranked ahead of all teams it defeated, and behind all the teams to which it lost. However, this is generally impossible at the end of the season, as any ranking will "violate" the results of some games, by having the loser ranked above the winner. For any given set of game results, there is a minimum possible number of violations, and we call a ranking that induces the minimal number a perfect ranking; computing such rankings is an NP-complete problem. Jay Coleman, an operations research professor at the University of North Florida, developed a model called MinV to identify perfect rankings. For the 2008 season, each of the six computer ranking systems used in college football's Bowl Championship Series induces at least 80% more violations than MinV. However, perfect rankings are not unique, raising the question of which perfect ranking is the best ranking. If all perfect rankings agreed on the top teams, this might not be a major concern, but in the 2008 season, there were five teams that could have been at the top of a perfect ranking, at season's end. Because of clustered scheduling, it is possible to move groups of teams up or down, and sometimes even whole conferences, while maintaining a perfect ranking. Under MinV, a highly-ranked team may be unaffected by a loss to a far-inferior opponent, contrary to logic. The latter portion of this paper details multiple examples of these issues.*

The highest division of collegiate football, the Football Bowl Subdivision (FBS) of NCAA Division I, has the distinction of being the only NCAA team sport which does not determine an official national champion. Based on long-standing tradition, top teams each compete in a season-ending bowl game, and a national champion is unofficially chosen by polls of coaches and media members. Several major conferences have historically sent their champions annually to a particular bowl game. More often than not, the consensus top two teams, after the regular season, have not faced one another in a bowl game, sometimes leading to disagreements over which team was most deserving of the national title. Under this system, it was not uncommon for two major polls to anoint different national champions at season's end.

Since the formation of the Bowl Championship Series (BCS) before the 1999 season, the top two teams are guaranteed to face another in a *de facto* national championship game. The difficulty comes in determining which two teams ought to play in that game. Currently, the teams are selected by a formula that weights equally the results of the Harris Interactive College Football Poll [14], the USA Today Coaches' Poll [21], and a consensus of six independent computational rankings [1, 5, 10, 16, 19, 23]. Prior to 2004, other factors were explicitly included in the computation – losses, strength-of-schedule, and "quality" wins [6]. However, these were already implicitly included in each of the component rankings, so their additional weighting was deemed unnecessary.

The six BCS "computer rankings" use a variety of algorithms to rank teams, but have in common that they use only the game scores and schedules to determine their rankings. Beginning with the 2002 season, computer ranking systems may not use margin of victory as a component factor [6]. While only these six computer rankings are used for BCS selections, many others exist. The Prediction Tracker [4] analyzes the accuracy of dozens of computer ranking systems, and David Wilson's website [22] lists and categorizes an even larger number.

The methodologies used in these computational rankings vary widely. Richard Billingsley's rankings [5] are recursive, meaning that each week's rankings are computed using only the current week's games and the previous week's rankings. Some include margin-of-victory and/or home-field advantage, while others do not. While many involve advanced mathematical techniques, including maximum likelihood estimation (in Kenneth Massey's rankings [17]) or the solution of large linear systems (in the Colley Matrix rankings [11]), there are also some which use quite simple formulae. An example is the Ratings Percentage Index (RPI), which rates teams using a weighted average of winning percentage, opponents' winning percentage, and opponents' opponents winning percentage. While RPI is more commonly associated with college basketball, it is also widely published for college football.

There is no single measure of which ranking system is the best, due to a continuum of philosophies. At one end lie predicting ranking systems, those which attempt to most accurately forecast the outcome of upcoming games, based upon currently available information. In these systems, teams are rated and ranked according to their perceived ability to win future games. Over the last decade, no computer rating system has consistently outperformed the Las Vegas oddsmakers. (See "most accurate predictor" in [3]). The bookmakers have a strong financial interest in accurate predictions, and ostensibly use all available information in setting the betting line, including injuries, detailed game statistics, and any relevant intangibles. At the opposite end of the spectrum from the predictive systems are retrodictive ranking systems, whose goal is to determine a ranking that accurately describes what has already occurred during the current season. Such rankings can be described as "fairer" for determining which teams have performed best during the current season. Many rating are based on philosophies which are at neither extreme, attempting to model past games and predict future ones with a single set of ratings.

If margins-of-victory and home-field advantage are removed, then evaluating the results of retrodictive ranking methods becomes simple. The best method is the one whose rankings contradict the known game results least often. In other words, an ideal retrodictive method would determine a ranking with the property that each team is ranked above the teams they defeated, and below the teams to which they lost. We will see that this is likely impossible, so the goal becomes to rank teams in a way that minimizes inconsistencies (known as *ranking violations* or *reversals*), occurrences in hindsight of a lower-ranked team beating a higher-ranked one. This is a variation of the *linear ordering problem* from graph theory, which involves ordering a group of related objects such that some function is minimized, in this case, minimizing the number of violations. The linear ordering problem is an example of an NP-complete problem ("Directed Optimal Linear Arrangement," denoted [GT43] in [12]), part of a class of computationally difficult problems for which no fast (polynomial-time), efficient, general method is known.

When the final 2008 rankings for the six BCS computer rankings are evaluated for retrodictive quality, the violation percentages are all similar, as shown in Table 1. However, we will find that there is substantial room for retrodictive improvement in all of these methods. While having zero violations would be ideal, there are reasons why some are typically inevitable.

In a three-team round-robin (where every team plays each of the others), it would not be uncommon for each to win one game and lose one game. This outcome is called a *cyclic triad* [9] or a *loop* [20]. Such occurrences are particularly noteworthy when the teams involved share a conference or divisional championship as a result, as was the case in the Big 12 South Division in 2008. Oklahoma beat Texas Tech, Texas Tech beat Texas, and Texas beat Oklahoma. Each team won all of their other conference games, so all three had identical conference records (7-1), divisional records (4-1), and head-to-head records (1-1). The determination of which team advanced to the conference championship game had implications for the national championship. In a three-team loop, any ordering of the three teams will induce at least one violation, and some will cause two violations. While any of the three teams could be ranked highest relative to the others, the choice of the first team determines the order of the other two. To avoid a second ranking violation, the middle team must be the one which lost to the top team, so the lowest-ranked team is the one which defeated the top team. (Thus, the win of the lowest-ranked team over the highest-ranked team is the violated result.)

There is another situation in which ranking violations are inevitable, the true "upset," in which a weaker team defeats a stronger one. A recent example of an upset is Mississippi's 2008 win over eventual national champion Florida. Mississippi had a record of 8-4, with losses to Wake Forest (8-5), Vanderbilt (7-6), South Carolina (6-6), and Alabama (12-2), while Florida (12-1) defeated three highly-rated teams and lost only to the Rebels. It is not reasonable to rank Mississippi ahead of the Gators, and doing so would cause multiple violations, as Florida beat three of the teams that defeated Mississippi. (See Figure 1.) To avoid this, we accept the one violation induced by Mississippi's upset of Florida, and henceforth largely overlook this game.

It is noteworthy that some unexpected results, especially those occurring early in the season, are no longer viewed as upsets in hindsight. Unheralded Alabama's season-opening victory against a preseason-top-ten Clemson team became less surprising with each passing week, as Alabama won its first twelve games while the Tigers finished 7-6. Although Clemson was the higher-ranked team at the time of the game, making the result an upset in the traditional sense, we do not consider it an upset, because Alabama was ranked above Clemson by season's end. Upsets can be viewed as loops involving three or more teams. If Team A beats Team B, which beats Team C, which beats another, etc., and the last team beats Team A, then a loop is created. Appropriately resolving intertwined loops is a focal point of the GridMap rankings [20].

Finally, a small number of necessary violations occur based on a pair of teams meeting twice during a season, with each team winning once. While college football teams do not

BCS Rating System	Violations	Percentage
Sagarin	125	17.4%
Billingsley	126	17.6%
Massey	131	18.3%
Anderson/Hester	133	18.5%
Wolfe	133	18.5%
Colley	138	19.2%

Table 1: Ranking violation comparison among the six BCS computer rating systems

schedule two games against the same opponent during a given season, post-season contests (conference championships or bowl games) can produce rematches of regular-season games. One notable example was the Sugar Bowl following the 1996 season, in which Florida beat rival Florida State to win the national championship, after losing to the Seminoles in the last game of the regular season. When a pair of teams split two games, there will be exactly one violation produced, regardless of how the teams are ranked relative to one another.

Jay Coleman, an operations research professor at the University of North Florida, developed an integer programming model in [9] called *MinV*, which determines the minimum possible number of ranking violations for a given set of data, then attempts to find a ranking which satisfies this minimum. The methodology used is beyond the scope of this paper, but the algorithm is described in [9]. We will consider a ranking to be retrodictively *perfect* if it results in the minimum number of violations. While the problem of finding a perfect ranking is very computationally intensive, as previously discussed, it is made slightly easier by the fact that perfect rankings are not unique. In practice, there are typically large numbers of such rankings, however, they are still difficult to find among the $120! \bullet 10^{200}$ permutations of the 120 teams in major-college football.

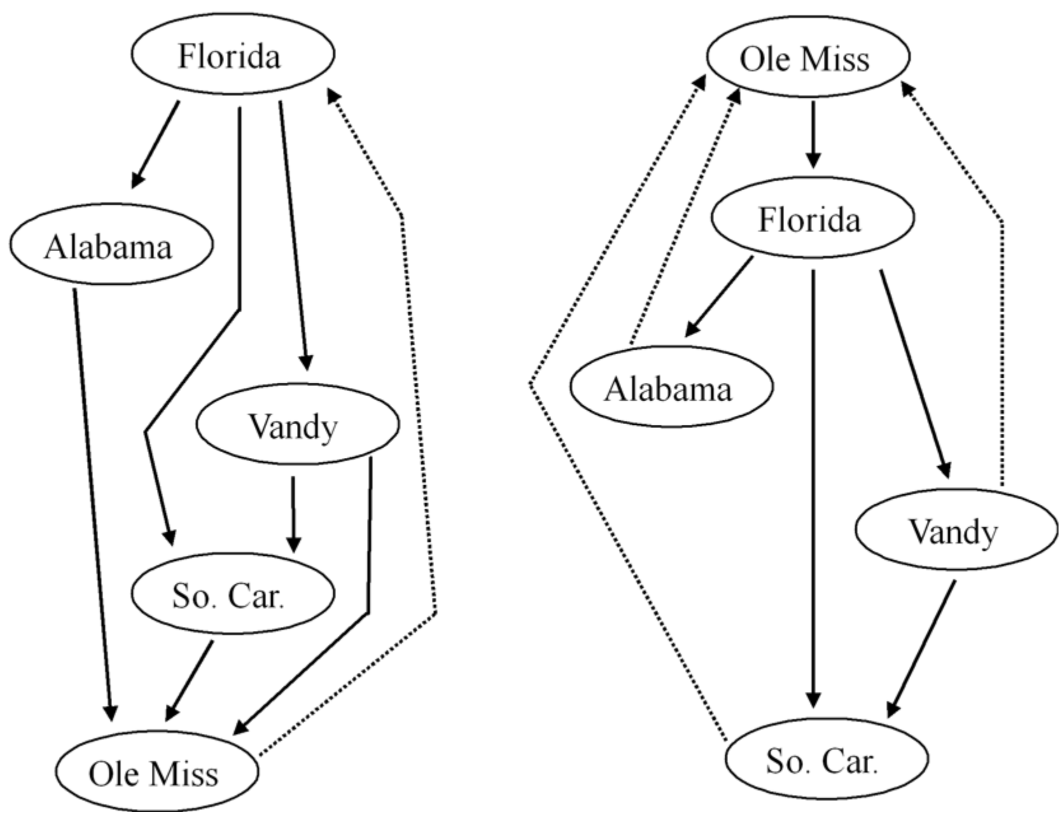

Figure 1: Two sets of relative rankings among teams, differing how they handle the defeat of Florida by Mississippi (Ole Miss). Teams nearest the top are ranked highest, and arrows indicate a victory by one team over another. Solid arrows indicate games in which the winner is ranked higher, while dashed arrows denote violations. In the ranking on the left, there is only one violation. In the ranking on the right, placing Mississippi above Florida leads to three violations.

For each season since 2004, Coleman has published a post-season ranking [8] which induces the minimum possible number of violations. For the 2008 season, there were 717 games contested between FBS teams, and the minimum number of violations was 69, so a perfect ranking correctly reflects the outcome of 90.4% of the games played during the season. By this intuitive measure of the retrodictive quality of a ranking system, Coleman's MinV rankings are far superior to other ranking systems, such as the BCS rankings in Table 1. Even the best of those six by this measure, the Sagarin ratings, has 81% more violations than Coleman's ranking. Of the 100+ rankings compared by Massey [15], only three others come within 25% of the minimum number of violations. Two of those three ranking systems are direct extensions of Coleman's MinV model; the other, GridMap [20], is independent, but uses similar principles.

Coleman's model is clearly outstanding at reflecting the outcome of games during the current season, so this raises the question of why nothing comparable to MinV is used in the BCS rankings, as a part of that system. The answer lies in the non-uniqueness of perfect rankings. Because there are many possible perfect rankings, we would need further criteria to determine which one should be used. If the differences among perfect rankings are relatively minor, at the level of the top teams, then we might conclude that any such ranking is acceptable and could set arbitrary criteria to achieve uniqueness.

To determine the degree of variation among retrodictively perfect rankings, we consider the results of the 2008 season, including the bowl games. Because only teams within the FBS are being ranked, games against non-FBS teams are not included in teams' records. To create new perfect rankings, we start with Coleman's final 2008 MinV ranking [8], then move teams in such a way that no additional net violations are generated. Departing from Coleman's practice in the MinV rankings, we will not allow teams to be tied for a particular place.

There are five teams which could be #1 in a perfect ranking: Utah, Florida, USC, Texas, and Texas Tech, as shown in Table 2. Any team which is undefeated (and untied) will have a claim to the top spot, but in 2008, Utah was the only such team. We find that four other teams could also appear atop a perfect ranking. Florida's only defeat was the previously-mentioned upset by Mississippi, so there is no team which must be ranked ahead of Florida. Oklahoma, Texas, and Texas Tech were part of a cyclic triad, so any of these could be ranked ahead of the other two. Having no other losses, Texas could then be ranked #1. Texas Tech had one other loss, to Mississippi. Even two wins over highly-ranked teams would not allow Mississippi to be placed above either Texas Tech or Florida, thus Texas Tech could land the top spot. However, Oklahoma's loss to Florida forced the Sooners to be ranked below Florida. Finally, Southern Cal could also be #1, with a record blemished only by an upset loss to Oregon State.

It is interesting to note the wide range of potential rankings for some teams. There is no team which appears in the top five of every perfect ranking, and only three teams – Florida, USC, and Oklahoma – are even guaranteed a spot in the top ten. Despite being the only undefeated team, Utah can be ranked as low as #13. Texas could fall as far as #21, and Alabama (ranked 6th by consensus of the BCS computer rankings, see Appendix A) could be altogether removed from the top 25. It is equally possible to move teams higher than their results seem to warrant. For example, there is a perfect ranking in which the top four teams (in descending order) are Florida, Florida State (7-4), Virginia Tech (9-4), and Maryland (7-5). The latter three, all from the relatively weak Atlantic Coast Conference, finished the season ranked 18th, 16th, and 35th, respectively, by consensus. The much stronger Big 12 Conference provides another example of how whole conferences can be over-ranked or under-ranked, due to the strong interconnectedness of teams within a conference. Table 3 lists the top 25 for two perfect rankings, with very different results for the Big 12. In the ranking on the left, eight of the top sixteen teams are from the Big 12, including 4-7 Colorado at #16. In the ranking on the right, the Big 12 has just one of the top nineteen teams. Neither is an accurate representation of the conference's performance.

Just as there are can be questionable outcomes at the top of a perfect ranking, the same is true at the bottom. Traditional power Michigan had a miserable season by its standards, winning just three games. However, two of those victories came against bowl-bound teams, Wisconsin and Minnesota. Because these upset wins are largely ignored, the Wolverines can be ranked as low as #119 out of 120 FBS teams, even though there were 28 other teams which won three or fewer games. Middle Tennessee State won five games, including a victory over a Maryland team that eventually won a bowl game, yet MTSU can be ranked as low as #117. On the other hand, Washington, one of only three winless FBS teams, can be ranked in the upper half at #57. For comparison, the consensus rankings of Michigan, MTSU, and Washington are # 94, #97, and #115, respectively. While less relevant for applications involving the BCS, these results still demonstrate weaknesses of a ranking system based solely on minimizing violations.

In general, this system penalizes a defeat to a far-inferior team less than a loss to a similarly-ranked or slightly inferior opponent. Counter-intuitively, a contending team with one defeat is affected less by losing to a weak team than by losing to a strong one. On the other side of the coin, upsetting a far-superior team may not result in a ranking improvement, while beating a slightly-superior team will generally be rewarded. At season's end, Mississippi was ranked in the top 20 of both major polls and all six BCS computer rankings, largely on the strength of twin upsets of Florida and Texas Tech, but minimizing violations requires that Ole Miss be ranked below three mediocre teams to which they lost. These principles are in opposition to those of most other ranking methods, in which every upset carries weight, a loss to a weak opponent is more damaging than a loss to a strong one, and defeating a top team is highly rewarded.

1) Utah (12-0)	1) Florida (12-1)	1) Southern Cal (12-1)	1) Texas (12-1)	1) Texas Tech (9-2)
2) Florida (12-1)	2) Southern Cal (12-1)	2) Florida (12-1)	2) Florida (12-1)	2) Florida (12-1)
3) Southern Cal (12-1)	3) Texas (12-1)	3) Texas (12-1)	3) Southern Cal (12-1)	3) Southern Cal (12-1)
4) Texas (12-1)	4) Oklahoma (11-2)	4) Oklahoma (11-2)	4) Oklahoma (11-2)	4) Texas (12-1)
5) Oklahoma (11-2)	5) Utah (12-0)	5) Utah (12-0)	5) Utah (12-0)	5) Oklahoma (11-2)
6) Alabama (12-2)	6) Alabama (12-2)	6) Alabama (12-2)	6) Alabama (12-2)	6) Utah (12-0)
7) TCU (10-2)	7) TCU (10-2)	7) TCU (10-2)	7) TCU (10-2)	7) Alabama (12-2)
8) Penn State (10-2)	8) Penn State (10-2)	8) Penn State (10-2)	8) Penn State (10-2)	8) TCU (10-2)
9) Texas Tech (9-2)	9) Texas Tech (9-2)	9) Texas Tech (9-2)	9) Texas Tech (9-2)	9) Penn State (10-2)
10) Boise State (11-1)	10) Boise State (11-1)	10) Boise State (11-1)	10) Boise State (11-1)	10) Boise State (11-1)

Table 2: Top-10 listings headed by each of the five possible #1 teams. In every case, teams #11-120 could remain unchanged from Coleman's MinV ratings.

The existence of multiple perfect rankings was an issue well-known to Coleman at the time of his 2005 paper [9]. Logically, the next step in extending violation-minimizing method is to add additional criteria to narrow down the possible outcomes. Beginning with the 2006 season, Coleman has used victory margins to determine which on-field results should be violated when there is a choice (for example, in the Texas – Texas Tech – Oklahoma cyclic triad). In the games among those three teams, the smallest margin of victory was Texas Tech's six-point win over Texas, so this result is violated. This is considered the preferred option, as compared to violating Texas' ten-point win over Oklahoma, or Oklahoma's 44-point win over Texas Tech. Thus, Texas is rated above Oklahoma, which is rated above Texas Tech. While the use of margin-of-victory in rankings is contentious (and prohibited in the BCS computer rankings), Coleman's choice makes logical sense. For the 2008 season, Coleman determined that, with the minimum 69 violations, the smallest possible sum of the margins of violated games is 573 points [8]. This step does not make the ranking unique (of the five possible #1 teams, only Texas Tech is eliminated from potentially topping the ranking). To further narrow the possibilities toward uniqueness, the MinV rankings use the Sagarin rankings [19] as a final criterion, attempting to

1) Texas Tech (9-2, # 8)	1) Florida (12-1, #2)
2) Texas (12-1, # 3)	2) Southern Cal (12-1, # 4)
3) Florida (12-1, #2)	3) Utah (12-0, #1)
4) Oklahoma (11-2, # 5)	4) Alabama (12-2, # 6)
5) Southern Cal (12-1, # 4)	5) Penn State (10-2, # 11)
6) Utah (12-0, #1)	6) Florida State (7-4, # 18)
7) TCU (10-2, # 7)	7) Virginia Tech (9-4, # 16)
8) Boise State (11-1, # 10)	**8) Oklahoma (11-2, # 5)**
9) Oregon (10-3, # 12)	9) TCU (10-2 # 7)
10) Oklahoma State (8-4, # 22)	10) Boise State (11-1, # 10)
11) Missouri (9-4, # 19)	11) Oregon (10-3, # 12)
12) Florida State (7-4, # 18)	12) Oregon State (9-4, # 15)
13) Virginia Tech (9-4, # 16)	13) Cincinnati (10-3, # 17)
14) Nebraska (9-4, # 20)	14) Pitt (9-4, # 24)
15) Kansas (7-5, # 31)	15) Arizona (8-5, # 38)
16) Colorado (4-7, # 60)	16) Maryland (7-5, # 35)
17) Alabama (12-2, # 6)	17) California (9-4, # 21)
18) Penn State (10-2, # 11)	18) Brigham Young (9-3, # 28)
19) Oregon State (9-4, # 15)	19) Air Force (7-5, # 49)
20) Ohio State (9-3, # 13)	**20) Texas Tech (9-2, # 8)**
21) Arizona (8-5, # 38)	**21) Texas (12-1, # 3)**
22) Maryland (7-5, # 35)	22) Ohio State (9-3, # 13)
23) California (9-4, # 21)	**23) Oklahoma State (8-4, # 22)**
24) Cincinnati (10-3, # 17)	**24) Missouri (9-4, # 19)**
25) Pitt (9-4, # 24)	**25) Nebraska (9-4, # 20)**

Table 3: Two perfect rankings, with the one on the left favorable for the Big 12 conference, and the one on the right unfavorable for the Big 12. In both rankings, Big 12 teams are listed in bold. Listed after each team's record is its consensus rank among the six BCS computer rankings.

most closely match those rankings while maintaining the minimum number of violations and total margin of violations. This is a rather unsatisfactory solution, as it is preferable for ranking systems to each stand on their own, independent of those produced by other algorithms. Because a version of the Sagarin ratings is already used in the BCS formula, adding the MinV ratings to the BCS would overweight Sagarin's ratings.

GridMap, also designed to minimize violations, takes another approach to resolving loops. It utilizes a different algorithm (which is not fully disclosed), focusing on intertwined loops, in which one or more games are a part of multiple loops. One notable feature is that each violation is listed, along with the sequence of games in the associated loop (since every violation involves a loop). In Massey's 2008 comparison of ranking systems, GridMap finished second only to MinV in violation percentage [15]. The approach appears to be less computationally intensive than Coleman's method, and no external ranking is used. However, the result is generally not a retrodictively perfect ranking, and many unbroken ties are included in the final results. In 2008, GridMap had Utah at #1, followed by a three-way tie between Florida, Texas, and Southern Cal, and then a four-way tie at #5 between Oklahoma, Alabama, Penn State, and Ball State. It is notable that Ball State, which achieved an 11-2 record against a weak schedule, was not ranked in the final top 25 by MinV, either major poll, or any of the BCS computer rankings. Also, winless Washington (consensus #115) is ranked ahead of 10-3 Tulsa (consensus #40) by the GridMap algorithm. These results seem to indicate that the method used by GridMap is susceptible to questionable outcomes similar to early versions of MinV (not accompanied by the Sagarin ratings). Such outcomes occur, at least in part, because GridMap's method ignores any violated game result for ranking purposes, so (as with MinV) upset losses to weak teams are generally not penalized at all, while losses to strong teams are penalized.

The rankings resulting from Coleman's work may not be suitable for inclusion in the BCS rankings at this time, yet they remain useful because they are based on sound logical principles, and they provide a contrarian view to that of other ranking systems. The huge retrodictive advantage of MinV, as compared to previous methods, shows that there is a great deal of room for improvement among other types of computational methods in accurately reflecting prior on-field results. Perhaps new ranking systems will emerge, combining near-perfect retrodiction with a high degree of predictive accuracy, in a way that meaningfully accounts for upsets. Standing on its own, the MinV algorithm is a significant advance, using the power of high-speed computers to achieve the highest possible retrodictive success.

Acknowledgments

Thanks to Jay Coleman and Warren Repole for providing historical game scores.

Sources Cited

1. Anderson, Jeff. "The Anderson & Hester College Football Computer Rankings" Anderson Sports. <http://www.andersonsports.com/football/ACF_frnk.html>

2. "AP Football Poll Archive." AP Poll Archive. <http://www.appollarchive.com/football/index.cfm>

3. Beck, Todd. "College Football Ratings PT Awards." The Prediction Tracker. Accessed on 16 July 2009. <http://tbeck.freeshell.org/fb/awards2008.html>

4. Beck, Todd. "Computer Rating System Prediction Results for College Football." The Prediction Tracker. Accessed on 16 July 2009. <http://www.thepredictiontracker.com/ncaaresults.php?type=1&year=08>

5. Billingsley, Richard. "Billingsley Report on Major College Football." College Football Research Center. <http://www.cfrc.com>

6. "BCS Chronology." 25 June 2009. Bowl Championship Series. Accessed on 11 August 2009. <http://www.bcsfootball.org/bcsfb/history>

7. "BCS Computer Rankings." 11 Nov. 2008. Bowl Championship Series. Accessed on 21 July 2009. <http://www.bcsfootball.org/bcsfb/rankings>

8. Coleman, Jay. "Minimum Violations (MinV) College Football Ranking. 10 Jan. 2009. University of North Florida. Accessed on 8 July 2009. <http://www.unf.edu/~jcoleman/minv.htm>

9. Coleman, Jay. "Minimizing Game Score Violations in College Football Rankings." Interfaces 35:6 (2005), pp. 483-496.

10. Colley, Wesley N. "Colley's Bias Free Matrix Rankings." Colley Matrix. <http://www.colleyrankings.com>

11. Colley, Wesley N. "Colley's Bias Free College Football Ranking Method." 30 May 2002. Colley Matrix. Accessed on 25 Aug. 2009. <http://www.colleyrankings.com/method.html>

12. Garey, Michael R. and Johnson, David S. Computers and Intractability: A Guide to the Theory of NP-Completeness. New York: W. H. Freeman and Co., 1979.

13. "GridMap post-Bowls" GridMap of College Football. Accessed on 15 July 2009. <http://www.gridmaponline.com>

14. "Harris Interactive College Football Poll." Harris Interactive. <http://www.harrisinteractive.com/news/bcspoll.asp>

15. Massey, Kenneth. "College Football Ranking Comparison." 23 Mar. 2009. Massey Ratings. Accessed on 22 July 2009. <http://www.mratings.com/cf/compare.htm>

16. Massey, Kenneth. "College Football Rankings." Massey Ratings. <http://www.masseyratings.com/rate.php?lg=cf>

17. Massey, Kenneth. "Massey Ratings Description." 15 Aug. 2000. Massey Ratings. Accessed on 28 July 2009. <http://www.mratings.com/theory/massey.htm

18. Repole, Warren. "Historical Scores." 31 May 2009. The Sunshine Forecast. Accessed on 23 June 2009. <http://www.repole.com/sun4cast/scores2009.html>

19. Sagarin, Jeff. "Jeff Sagarin NCAA Football Ratings." 9 Jan. 2009. USA Today. Accessed on 22 July 2009. <http://www.usatoday.com/sports/sagarin/fbt08.htm>

20. "Threads, Loops, and Upsets." GridMap of College Football. Accessed on 15 July 15 2009. <http://www.gridmaponline.com/#/threads-loops-and-upsets/4525582180>

21. "USA Today Coaches' Poll." USATODAY.com. <http://www.usatoday.com/sports/college/football/usatpoll.htm>

22. Wilson, David L. "American College Football - Rankings." Nutshell Sports Ratings. Accessed on 22 July 2009. <http://www.nutshellsports.com/wilson/popup.html>

23. Wolfe, Peter. "Ratings of all NCAA and NAIA Teams." 2008 College Football. <http://prwolfe.bol.ucla.edu/cfootball/home.htm>

R. DREW PASTEUR is an Assistant Professor of Mathematics at The College of Wooster. Before pursuing graduate study at North Carolina State University, he was a math teacher and head athletic trainer at Fuquay-Varina High School, in Fuquay-Varina, NC. It was during this time that Drew became interested in mathematical ranking algorithms and developed the first version of his Fantastic 50 high school football rankings. His other research area is in mathematical biology, applying dynamical systems to human physiology. Outside of mathematics, Drew enjoys singing, running, and spending time with his wife and son.

311 Taylor Hall, 1189 Beall Ave., Wooster, OH, 44691. rpasteur@wooster.edu

Appendix: Compilation of various rankings

Retrodictive Rankings: MinV [8], GridMap [13]
BCS Computer Rankings: Sagarin [19], Anderson/Hester [1], Billingsley [5], Colley [10], Massey [16], Wolfe [23]
Major Polls: Associated Press [2], USA Today Coaches' [21]

Consensus rankings were determined by dropping the highest and lowest rank for each team among the six BCS computer rankings, then averaging the remaining four. Ties were broken by using the two dropped rankings.

Consensus ranking among BCS computer rankings	Retrodictive		BCS computer rankings						Polls	
	MinV	GMap	Sag	And	Bil	Col	Mas	Wol	AP	USA
1) Utah (12-0)	5	1	1	1	3	3	1	1	2	4
2) Florida (12-1)	1	2-T	2	2	1	1	2	2	1	1
3) Texas (12-1)	3	2-T	3	3	4	2	4	3	4	3
4) USC (12-1)	2	2-T	5	5	2	4	5	4	3	2
5) Oklahoma (11-2)	4	5-T	4	4	5	5	3	5	5	5
6) Alabama (12-2)	6	5-T	6	6	6	6	6	7	6	6
7) TCU (10-2)	7	9-T	7	7	8	11	8	6	7	7
8) Texas Tech (9-2)	9	9-T	9	8	7	7	7	9	12	12
9) Georgia (9-3)	29-T	17-T	8	11	12	10	9	8	13	10
10) Boise State (11-1)	10	12-T	15	9	9	9	10	12	11	13
11) Penn State (10-2)	8	5-T	13	10	10	8	11	11	8	8
12) Oregon (10-3)	11	14-T	11	13	14	14	12	10	10	9
13) Ohio State (9-3)	13	9-T	17	12	11	12	17	15	9	11
14) Mississippi (8-4)	59-T	66-T	10	20	13	19	13	13	14	15
15) Oregon State (9-4)	12	17-T	12	16	19	16	14	14	18	19
16) Virginia Tech (9-4)	15	12-T	14	18	20	15	15	16	15	14
17) Cincinnati (10-3)	22-T	14-T	19	14	22	13	19	17	17	17
18) Florida State (7-4)	14	21-T	16	21	26	20	16	19	21	23
19) Missouri (9-4)	17	21-T	22	15	21	18	22	20	19	16
20) Nebraska (9-4)	18	25-T	21	19	25	22	20	22		
21) California (9-4)	21	28-T	18	24	24	21	23	18		25
22) Oklahoma State (8-4)	16	17-T	24	22	15	25	21	21	16	18
23) Georgia Tech (7-4)	28	14-T	23	26	18	24	18	28	22	22
24) Pittsburgh (9-4)	24-T	21-T	27	17	29	17	28	25		
25) Michigan State (9-4)	33	31-T	31	23	23	23	31	27	24	24
26) West Virginia (8-4)	26	33-T	25	27	28	27	29	24	23	
27) LSU (7-5)	64	75-T	20	34	17	38	30	23		
28) BYU (9-3)	32	25-T	34	25	16	26	26	31	25	21
29) Wake Forest (8-5)	52-T	43-T	26	28	47	31	25	29		
30) Boston College (8-5)	31	17-T	28	29	38	29	24	32		

Consensus ranking	**MinV**	**GMap**	**Sag**	**And**	**Bil**	**Col**	**Mas**	**Wol**	**AP**	**USA**
31) Kansas (7-5)	22-T	28-T	32	30	34	34	32	30		
32) North Carolina (7-5)	27	40	33	31	41	33	27	37		
33) Iowa (8-4)	35	35	37	32	32	32	38	33	20	20
34) Rice (10-3)	59-T	66-T	39	33	39	30	40	26		
35) Maryland (7-5)	20	25-T	30	37	35	37	35	35		
36) Vanderbilt (7-6)	57-T	62-T	29	38	36	43	37	34		
37) Rutgers (7-5)	45	55-T	38	40	54	40	39	39		
38) Arizona (8-5)	19	21-T	35	48	27	48	44	36		
39) Connecticut (7-5)	47-T	59-T	42	39	43	39	43	38		
40) Tulsa (10-3)	68	80-T	50	42	30	35	48	40		
41) Northwestern (8-4)	34	33-T	48	36	33	36	46	47		
42) Miami-FL (6-6)	50-T	41-T	36	44	60	45	34	42		
43) South Carolina (6-6)	36	36-T	40	41	45	44	36	43		
44) Ball State (11-2)	69	5-T	62	35	31	28	54	51		
45) Clemson (5-6)	29-T	28-T	41	46	40	46	33	48		
46) South Florida (7-5)	46	57-T	45	45	37	41	45	45		
47) East Carolina (9-5)	67	75-T	46	47	50	42	50	41		
48) Kentucky (6-6)	59-T	66-T	44	50	49	52	47	44		
49) Air Force (7-5)	39-T	51	57	43	51	47	49	53		
50) Houston (7-5)	62-T	71-T	52	54	48	51	58	46		
51) Virginia (4-7)	57-T	62-T	43	58	52	57	41	54		
52) Colorado State (6-6)	41-T	31-T	53	49	77	53	52	49		
53) Notre Dame (7-6)	37	41-T	51	52	59	54	56	52		
54) Navy (7-5)	70	49-T	54	51	62	50	53	56		
55) NC State (5-7)	47-T	38-T	47	55	58	56	42	57		
56) Arkansas (4-7)	62-T	71-T	49	57	57	62	51	50		
57) Western Michigan (8-4)	77	71-T	64	53	42	49	63	55		
58) Wisconsin (6-6)	41-T	36-T	60	56	44	55	60	60		
59) Stanford (5-7)	54-T	59-T	56	67	53	67	59	58		
60) Colorado (4-7)	24-T	52	58	59	66	68	57	59		
61) Auburn (4-7)	65	80-T	59	64	55	71	62	61		
62) Tennessee (5-7)	54-T	59-T	61	61	63	70	61	64		
63) Louisiana Tech (7-5)	41-T	49-T	66	63	61	61	68	65		
64) Nevada (6-6)	39-T	46-T	65	60	68	59	65	66		
65) Minnesota (6-6)	73	38-T	71	62	64	63	71	69		
66) Arizona State (4-7)	50-T	55-T	63	74	46	73	64	74		
67) Duke (3-8)	54-T	46-T	55	78	75	76	55	70		
68) Hawaii (6-7)	38	43-T	74	68	69	65	67	78		
69) Buffalo (8-6)	78-T	75-T	73	66	86	60	76	63		
70) Southern Miss (7-6)	66	99	67	75	71	69	77	62		
71) Fresno State (7-6)	44	53-T	75	69	83	66	70	71		
72) Louisville (4-7)	89	95-T	69	76	70	75	72	72		
73) Troy (7-5)	72	100	78	71	74	64	80	67		
74) Mississippi State (3-8)	81	85-T	68	83	67	86	69	76		
75) Illinois (4-7)	78-T	75-T	77	73	65	74	75	75		
76) Baylor (3-8)	71	62-T	70	72	78	79	66	80		
77) Central Michigan (7-5)	76	46-T	86	65	80	58	82	73		
78) Kansas State (4-7)	90	97-T	76	77	76	80	74	77		
79) UCLA (4-8)	52-T	57-T	72	84	56	81	73	81		
80) Florida Atlantic (7-6)	75	101	80	80	79	72	86	68		

Consensus ranking	MinV	GMap	Sag	And	Bil	Col	Mas	Wol	AP	USA
81) UNLV (5-7)	47-T	53-T	79	70	89	78	78	83		
82) Texas A&M (4-8)	105-T	108-T	81	79	81	85	79	88		
83) San Jose St (5-6)	91	62-T	83	81	82	77	84	82		
84) New Mexico (4-8)	80	85-T	82	82	88	89	81	84		
85) Purdue (3-8)	74	43-T	85	85	73	84	83	87		
86) UTEP (5-7)	96-T	95-T	87	86	87	83	87	79		
87) Arkansas State (5-6)	103-T	106-T	90	90	91	87	95	85		
88) Memphis (5-7)	101-T	97-T	89	89	97	88	88	92		
89) Bowling Green (6-6)	83	80-T	92	87	99	82	92	90		
90) Syracuse (2-9)	88	93-T	84	91	100	94	85	91		
91) Louisiana-Lafayette (6-6)	98-T	102-T	91	93	90	90	94	86		
92) Marshall (3-8)	100	85-T	88	95	95	92	89	89		
93) Wyoming (3-8)	85	88	95	88	84	96	90	99		
94) Michigan (3-9)	105-T	108-T	94	94	72	98	91	101		
95) Northern Illinois (5-7)	82	75-T	100	92	101	91	96	96		
96) Temple (5-7)	84	80-T	99	96	102	93	97	94		
97) MTSU (5-7)	107	108-T	97	100	94	99	98	93		
98) Akron (5-7)	87	91-T	101	98	106	95	100	95		
99) Florida International (5-7)	101-T	104-T	102	99	98	97	103	97		
100) Utah State (3-9)	92-T	66-T	98	97	110	100	93	102		
101) UCF (3-8)	98-T	104-T	93	102	107	101	99	100		
102) UAB (3-8)	96-T	102-T	96	103	104	103	101	98		
103) Indiana (2-9)	109	80-T	104	101	85	102	102	105		
104) Ohio U. (3-8)	86	89-T	110	104	103	104	108	103		
105) Washington State (1-11)	112	66-T	103	109	92	108	106	112		
106) Iowa State (1-10)	108	111-T	106	105	96	113	107	111		
107) Louisiana-Monroe (3-8)	115	116-T	109	108	108	105	110	104		
108) New Mexico State (2-9)	95	93-T	108	107	111	107	109	107		
109) Kent State (3-8)	111	113-T	112	110	105	106	111	106		
110) Army (3-8)	110	111-T	105	111	114	109	104	109		
111) San Diego State (2-9)	92-T	89-T	107	106	113	111	105	113		
112) Toledo (3-9)	103-T	106-T	111	113	109	112	112	108		
113) Eastern Michigan (2-9)	116	113-T	115	112	115	110	113	110		
114) Tulane (2-10)	114	113-T	114	114	116	114	114	114		
115) Washington (0-12)	113	71-T	113	116	93	115	116	119		
116) Idaho (1-10)	94	91-T	116	115	118	116	115	115		
117) Miami-OH (1-10)	118	116-T	118	118	112	119	117	116		
118) Southern Methodist (0-11)	117	116-T	117	117	117	118	118	118		
119) North Texas (1-11)	119	119	119	119	119	120	119	117		
120) Western Kentucky (0-10)	120	120	120	120*	120	117	120	120		

Notes:

- The AP and USA Today polls involve voting only for the top 25 teams, while computer rankings generally give a complete ranking of the 120 FBS teams.
- The final Anderson/Hester rankings listed only 119 teams, omitting Western Kentucky, so WKU was assumed to be ranked #120.
- The final Harris Interactive Poll voting occurs prior to the bowl games, so we have not included it.
- Only games against FBS teams are included in the teams' records.