Predicting Sports Rankings using Linear Algebra and Matrix Methods

Matthew Nay, Haley Samuelsen, and Jordan Sell

I. Introduction

RANKING sports teams is an important yet difficult task. Ranking teams aims to put teams in numerical order such that the "best" team is first and the "worst" team is last. Major League Baseball (MLB) has historically used winning percentage to rank teams during the regular season and consequently determine which teams make post-season. However, it is unclear if winning percentage is the best ranking method for predicting post-season outcomes. In our paper, we aim to look at two known matrix methods for calculating sports rankings – Massey's Method and Colley's Method – to see if either of these performs better than winning percentage in terms of aligning end of regular season rankings with post-season results given their more sophisticated formulation. Specifically, for each year spanning 2012-2021 (except 2020 because the the COVID-19 pandemic changed the MLB post-season format), we calculate winning percentage rankings and produce Massey and Colley rankings for the MLB regular season and compare them to the actual MLB post-season outcomes to analyze which method most closely matched reality.

II. Background

Ranking is a subject that has a long history and extends far beyond just sports. Biologists rank species of animals when creating food chains, the United Nations ranks countries using their Human Development Index (HDI), *US News* ranks US colleges and universities annually, and Netflix ranks movies and TV shows based on their ratings [1]. Ranking has garnered heightened attention in recent years largely because of today's data collection capabilities and increasing obsession with evaluation [1].

For many professional sports leagues, win-loss record, or synonymously winning percentage, has been widely used as the measure to rate and rank teams. We define a *rating* as a numerical score resulting from a particular method and a *ranking* as a rank-ordered list of teams. Ratings allow for rankings. While winning percentage is simple and intuitive, it leaves out several important factors that could play a big difference in determining which teams are truly the best. For instance, strength of schedule, point differential, and home field advantage, among other factors, should also plausibly be considered when ranking teams. This is especially relevant for college sports where it is not possible to play every team and there is a wide range of skill level across the board. To remedy this, polling experts have long been tasked with ranking teams on a holistic level given their experience around the sport. However, as with any human craft, polling introduces inherent human subjectivity.

Instead, linear algebra and matrix methods can be used to produce more objective, verifiable, and repeatable results. As early as 1929, sports ranking hobbyists began using mathematical methods to hand calculate sports rankings, specifically for college football [2]. In 1984, the National Collegiate Athletic Association (NCAA) first officially used a computer ranking system developed by Jeff Sangarin to help determine the field of 64 for that year's NCAA basketball tournament [2]. There are a variety of computer methods used to rank sports teams, each with their own strengths and weaknesses. This paper discusses Massey's Method and Colley's Method, two of the more well-known methods. However, there are other notable methods, including Keener's Method, Elo's System, the Markov Method, and the Offense-Defense Rating Model. See reference [1] for more information on each of the methods listed.

The use of linear algebra and matrix methods in ranking sports teams today varies widely depending on level, league, and sport. For example, the college football bowl committee takes two human polls and the standings of each of six computer polls to determine who gets to play in the post-season [3]. However, NCAA division III baseball and softball, the sports in which we play at MIT, only use human polls to determine rankings. On the other hand, the MLB only uses a formula, namely winning percentage, to determine rankings and avoids human involvement altogether.

With the many methods available to generate rankings, one might ask which is the best. While there can be several ways to define what makes a ranking method the "best," perhaps the most natural and applicable way would be to actually measure how well that ranking method matches where teams end up placing at the end of the season. In this sense, we are measuring how well the rankings predict end of season outcomes.

Using MLB regular season data for the years 2012-2021 (excluding 2020 because there was a different post-season format due to the COVID-19 pandemic), we will compare the MLB method of ranking teams (which is simply ordering teams by winning percentage) with the Massey and Colley rankings and see which was most accurate in predicting the results of the post-season. Before the analyses can occur, information about Massey's Method and Colley's Method will be provided. For both, we offer a derivation of the main formulas used in each method, highlight key properties, and walk through examples.

III. Massey's Method

A. Formulation

Massey's Method, also known as the Point Spread Method, was developed in 1997 by Kenneth Massey as a way to rank college football teams [4]. Massey believed it was important to take into account more than just a team's win-loss record when determining the rating of each team. He considered other factors, like the number of games played by each team and the point differential for each game. Massey's Method follows this general idea: the difference in ratings for each pair of teams should equal the difference between the score of their match/game/competition [5]. One equation can summarize Massey's Method:

$$r_i - r_j = y_k, \tag{1}$$

where for every game k, r_i and r_j are the rankings of teams i and j, respectively, and y_k is the margin of victory for the game. In an ideal scenario, the difference in the ratings of teams i and j ($r_i - r_j$) is an accurate predictor of the margin of victory between the two teams in a head-to-head match [6]. The next section explores the matrix formulation of Massey's Method, which is an expansion of Equation 1.

B. Matrix Method

All rating systems have the same goal: to assign a rating to each of the n teams in a league/conference/division that has played m games so far in a season. Looking back at Equation 1, while the individual ratings r_i's of each team are unknown, the value of y_k is known. Thus, we can create an equation of this form for every game k. This results in a system of m linear equations with n unknowns, which can be written in matrix form as

$$\mathbf{Xr} = \mathbf{y}. \tag{2}$$

The coefficient matrix $\mathbf{X}_{m\times n}$ is defined to be a sparse matrix: the matrix is filled with mostly zeros, with the exception of a 1 in location i and a -1 in location j, indicating that team i beat team j in their matchup. The vector $\mathbf{r}_{n\times 1}$ is the vector of unknown ratings and the vector $\mathbf{y}_{m\times 1}$ is the margin of victory for each matchup. Typically, $m >> n$, which causes some issues as the linear system is inconsistent and highly over-determined. However, a least squares solution can be obtained using the normal equations $\mathbf{X^T X r} = \mathbf{X^T y}$ [6].

When working through this formulation, Massey discovered that the best way to work with the normal equation was to set the coefficient matrix $\mathbf{M} = \mathbf{X^T X}$. The benefit to this is that $\mathbf{M}$ does not need to be computed due to a few of its characteristics. First, the diagonal entries $\mathbf{M}_{ii}$ are the total number of games played by team i. Then, for $i \neq j$, the $\mathbf{M}_{ij}$ entry is the negation of the number of games played by team i against team j. Looking at the right-hand side of the normal equations, $\mathbf{X^T y}$ also does not need to be calculated. This can be found by accumulating point differentials for each team, where the i^{th} entry is equal to the sum of the point differentials from every game played by team i. Similar to before, we can define $\mathbf{p} = \mathbf{X^T y}$. The last step is to adjust both $\mathbf{M}$ and $\mathbf{p}$ to fix the issue of the rank of $\mathbf{M}$. As $\mathbf{M}$ is currently defined, there is no unique solution to the system of equations due to the rank of $\mathbf{M}$. Massey fixed this by replacing the last row of $\mathbf{M}$ with all 1's and the last entry of $\mathbf{p}$ with a 0. Any row of $\mathbf{M}$ would be adjusted in this way, but the convention Massey set in [4] was to alter the last row. Finally, Massey's Method can be described using the following formula:

$$\mathbf{Mr} = \mathbf{p}, \tag{3}$$

where $\mathbf{M}$ is the $n \times n$ **Massey matrix** as described above, $\mathbf{r}$ is the $n \times 1$ vector of unknown ratings, and $\mathbf{p}$ is the $n \times 1$ vector consisting of cumulative point differentials [5]. Solving for $\mathbf{r}$ gives the **Massey rating** for each team. Placing these values in order gives the **Massey ranking**.

C. Properties

Massey's Method has a few notable properties that are key to keep in mind when using the system [5], [7].

1) **Biased:** Because this entire rating system is based on point differentials, it can be heavily skewed by teams defeating weaker opponents by a large amounts. This is not necessarily a strong indicator of team quality, yet may significantly affect the Massey Ratings.
2) **Games are all weighted equally (in terms of when they were played):** Neither this method nor Colley's Method (seen later in this paper) take into account the point during the season that a game was played, which is crucial. Teams can improve as the season go on, or a team may lose their star player due to injury halfway through the season. All games are weighted equally, no matter when they occur during the season.

D. Example

In this section, we will walk through an example using data from the 2012 Division III college football season [14]. Five teams are examined: Johns Hopkins, Franklin & Marshall, Gettysburg, Dickinson and McDaniel. Let J represent Johns Hopkins, F represent Franklin & Marshall, G represent Gettysburg, D represent Dickinson, and and M represent McDaniel. The following is a table of matchups between the teams and the scores:

Game Number	Team 1	Team 2	Score
1	G	M	35-3
2	F	D	36-28
3	F	M	35-10
4	J	D	49-0
5	J	G	49-35
6	D	M	38-31
7	G	D	13-23
8	J	F	12-14
9	J	M	49-7
10	F	G	31-38

Using Equation 2, where the columns of $\mathbf{X}$ correspond to teams $D, F, G, J,$ and M, respectively, and the rows correspond to games 1-10, the problem turns into:

$$\begin{bmatrix} 0 & 0 & 1 & 0 & -1 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 0 & -1 \\ 1 & 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & -1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} r_D \\ r_F \\ r_G \\ r_J \\ r_M \end{bmatrix} = \begin{bmatrix} 32 \\ 8 \\ 25 \\ 49 \\ 14 \\ 7 \\ 10 \\ 2 \\ 42 \\ 7 \end{bmatrix} \tag{4}$$

Now, using the steps to transform Equation 2 into Equation 3, the problem now becomes what we see below, where each column/row corresponds to teams D, F, G, J, and M, respectively:

$$\begin{bmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 & -1 \\ -1 & -1 & 4 & -1 & -1 \\ -1 & -1 & -1 & 4 & -1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} r_D \\ r_F \\ r_G \\ r_J \\ r_M \end{bmatrix} = \begin{bmatrix} -40 \\ 32 \\ 15 \\ 103 \\ 0 \end{bmatrix} \tag{5}$$

Solving for r, we can see that the Massey ratings for these teams are

Team	Massey Rating (r)
D	-8
F	6.4
G	3
J	20.6
M	-22

As we can see, Johns Hopkins has the highest Massey rating while McDaniel has the lowest Massey rating. This follows our general intuition when we look at things like the win-loss record for each team and the point differential.

This is just one example of the Massey method being used in a college sports rankings. The next section examines Colley's Method, and the two methods will be used later in the paper to examine the rankings for the MLB regular season in comparison to the method the MLB currently uses to rank its teams.

IV. Colley's Method

A. Formulation

Colley's Method began as a side project for astrophysicist Dr. Wesley Colley in 2001 [9]. Colley wanted to develop a method for ranking teams that preserved the simplicity of using winning percentage but addressed the the problematic fact that defeating a weak team would affect the rankings in the same way that defeating a strong team would, also known as teams having a different strength of schedule. The *winning percentage*, r, for team i is calculated as

$$r_i = \frac{w_i}{t_i}, \tag{6}$$

where w_i is the number of wins and t_i is the total number of games played by team i. To address the strength of schedule problem, Colley's Method solves a system of linear equations using an approach that adjust's a team's winning percentage to factor in their strength of schedule. In particular, the traditional winning percentage formula shown in Equation 6 becomes

$$r_i = \frac{1 + w_i}{2 + t_i}. \tag{7}$$

While at first glance, the adjustments may seem arbitrary, we will show that these modifications in fact make intuitive sense and elegantly incorporate each team's strength of schedule through LaPlace's rule of succession [10].

Firstly, perhaps the most intuitive argument that Colley set forth for the use of Equation 7 is that with traditional winning percentage from Equation 6, each team starts out with a pre-season rating of $\frac{0}{0}$ which does not carry any real value. Additionally, a team that loses its first game would have a winning percentage of 0%, whereas a team that wins their first game would have a winning percentage of 100%, so it appears that the latter team is infinitely better than the former team. Naturally, this is probably not the case. Using Colley's version of winning percentage in Equation 7, each team starts out with a pre-season rating of $\frac{1+0}{2+0} = \frac{1}{2}$ and then as the season progresses, the rating fluctuates around this. For example, a team that loses its first game would have a rating of $\frac{1}{3}$, and a team that wins their first game would have a rating of $\frac{2}{3}$. This seems to be a more sensible range compared to that produced by the traditional winning percentage.

Now the question is how does Equation 7 account for each team's strength of schedule. Essentially, the rating of team i should somehow be connected to the ratings of each of its opponents. Well, as just mentioned, each team starts out with a pre-season rating of $\frac{1}{2}$ and the ratings move above and below $\frac{1}{2}$ as the season progresses. Since each team's win is another team's loss, the ratings are interdependent. Another way we can see this is to decompose Equation 7 [11]. We introduce l_i as the number of losses team i has and O_i as the set of opponents team i has faced. We can then decompose Equation 7 into

$$\begin{aligned} w_i &= \frac{w_i - l_i}{2} + \frac{w_i + l_i}{2} \\ &= \frac{w_i - l_i}{2} + \frac{t_i}{2} \\ &= \frac{w_i - l_i}{2} + \sum_{j=1}^{t_i} \frac{1}{2}. \end{aligned}$$

As mentioned, all teams start with $r_j = \frac{1}{2}$, so the summation $\sum_{j=1}^{t_i} \frac{1}{2}$ is initially equal to $\sum_{j \in O_i} r_j$. As the season progresses, the summation $\sum_{j=1}^{t_i} \frac{1}{2}$ is not exactly equal but hovers around $\sum_{j \in O_i} r_j$ because all ratings fluctuate around $\frac{1}{2}$. Thus, we can approximate w_i as

$$w_i \approx \frac{w_i - l_i}{2} + \sum_{j \in O_i} r_j. \tag{8}$$

We can assume equality and insert Equation 8 into Equation 7 to reveal interdependence:

$$r_i = \frac{1 + \frac{w_i - l_i}{2} + \sum_{j \in O_i} r_j}{2 + t_i}. \tag{9}$$

We can now see that the r_i's depend on the r_j's, and this reveals how Colley's method incorporates strength of opponents into a team's ratings. Thus, we have reasoned that Colley's modification on the traditional winning percentage formula achieves its goal of incorporating a team's strength of schedule.

B. Matrix Method

The formulation described in the previous subsection can be converted into a system of linear equations and solved using linear algebra techniques. If we take Equation 9 and multiply across by $(2+t_i)$ and subtract $\sum_{j\in O_i} r_j$ from both sides, we get

$$(2+t_i)r_i - \sum_{j\in O_i} r_j = 1 + \frac{w_i - l_i}{2}.$$

The solutions to this system of equations are the **Colley ratings**. The system can be written in matrix form as

$$\begin{bmatrix} 2+t_i & -n_{12} & -n_{13} & \dots & -n_{1n} \\ -n_{12} & 2+t_2 & -n_{23} & \dots & -n_{2n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ -n_{n1} & -n_{n2} & -n_{n3} & \dots & 2+t_n \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

where

$$b_i = 1 + \frac{w_i - l_i}{2}, \tag{10}$$

and n_{ij} denotes the number of times team i has played team j. We can write this problem more compactly as solving

$$\mathbf{Cr} = \mathbf{b} \tag{11}$$

where $\mathbf{r}_{n\times 1}$ is a column-vector of the Colley ratings, $\mathbf{b}_{n\times 1}$ is a column-vector consisting of Equation 10, and $\mathbf{C}_{n\times n}$ is **Colley's matrix** where

$$\mathbf{C}_{ij} = \begin{cases} 2+t_i, & \text{if } i = j. \\ -n_{ij}, & \text{if } i \neq j. \end{cases}$$

Colley's matrix, $\mathbf{C}$ is invertible, so a unique solution is guaranteed. It is apparent that $\mathbf{C}$ is real and symmetric, and Colley also proved that $\mathbf{C}$ is positive definite [9]. Because of this, $\mathbf{C}$ has a Cholesky decomposition of the form $\mathbf{C} = \mathbf{U^T U}$, where $\mathbf{U}$ is an upper triangular matrix, and Equation 11 can be solved using Cholesky factorization, which is usually more quick and stable than other solving methods. However, in practice, the system $\mathbf{Cr} = \mathbf{b}$ is typically small enough to be solved by software packages using standard numerical routines, such as Gaussian elimination or Krylov methods [11].

C. Properties

Colley's Method has a couple of notable properties based on its construction [11].

1) **Bias-free:** Unlike Massey's method, Colley's method does not use point score data. Thus, it avoids bias toward teams that try to run up the score on weaker teams for the sake of improving their rating.
2) **Conservation of average rating:** Because each team starts with an initial rating of $\frac{1}{2}$ and fluctuates around this point throughout the season based on game outcomes, the average rating for all of the teams over the span of the season is $\frac{1}{2}$. This is a result of one team's win being another team's loss.

D. Examples

In this section, we first illustrate Colley's method with the simplest possible example and then proceed with a slightly more involved scenario. Both examples were presented in Colley's original paper [9]. We use these examples to validate that our code was implemented correctly when calculating the Colley rankings for MLB teams.

Example: We have two teams: one with a 1-0 record and a ranking denoted by r_W, and one with a 0-1 record and a ranking denoted by r_L. Using Equation 9, we have

$$r_W = \frac{1+\frac{1}{2}+r_L}{2+1}$$
$$r_L = \frac{1-\frac{1}{2}+r_W}{2+1}.$$

We re-arrange this to get the following two-variable linear system:

$$3r_W - r_L = 3/2$$
$$-r_W + 3r_L = 1/2$$

Re-writing this in matrix form, we have

$$\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} r_W \\ r_L \end{bmatrix} = \begin{bmatrix} 3/2 \\ 1/2 \end{bmatrix}.$$

Multiplying both sides by the inverse of the Colley matrix, we get the following solution vector:

$$\begin{bmatrix} r_W \\ r_L \end{bmatrix} = \begin{bmatrix} 5/8 \\ 3/8 \end{bmatrix}.$$

The team that won their first game has a higher rating than the team that lost their first game, and the average rating is

$$\frac{\left(\frac{5}{8}+\frac{3}{8}\right)}{2} = \frac{1}{2},$$

as is to be expected.

Example: Now, to explore a more involved example, suppose we have the following 5 teams with the results shown in the table (x indicates the teams did not play each other):

Team	A	B	C	D	E	Record
A	x	x	W	L	L	1-2
B	x	x	L	x	W	1-1
C	L	W	x	W	L	2-2
D	W	x	L	x	x	1-1
E	W	L	W	x	x	2-1

The following Colley matrix system corresponds to this example and can be solved using our code, which gets us Colley ratings and rankings identical to those in Colley's paper:

$$\begin{bmatrix} 5 & 0 & -1 & -1 & -1 \\ 0 & 4 & -1 & 0 & -1 \\ -1 & -1 & 6 & -1 & -1 \\ -1 & 0 & -1 & 4 & 0 \\ -1 & -1 & -1 & 0 & 5 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \\ 1 \\ 1 \\ 3/2 \end{bmatrix}$$

Team	Colley Rating ($\mathbf{r}$)	Ranking	Record
A	.413	5th	1-2
B	.522	2nd	1-1
C	.500	3rd	2-2
D	.478	4th	1-1
E	.587	1st	2-1

This example explicitly shows that the Colley Method does indeed factor in strength of schedule. This is evident by the different ratings produced for Team B and Team D. Although they had identical 1-1 records, Team B was rated higher because they played a 2-2 team and a 2-1 team, while Team D played a 2-2 team and a 1-2 team. Additionally, we see the average rating is conserved at

$$\frac{(.413 + .522 + .500 + .478 + .587)}{5} = \frac{2.5}{5} = \frac{1}{2}.$$

V. Results

A. Implementation and Validation

To see which of the three methods is the best indicator of actual post-season outcomes, we calculated the winning percentage ranking, Colley ranking, and Massey ranking for each MLB team that qualified for post-season in a given year using end of regular season results and then compared those rankings with post-season results. We did this for every year from 2012-2021 (excluding 2020) using data from Baseball Reference [12].

To validate that we correctly implemented Massey's Method, we ran our code using the most current 2023 MLB data to see if our rankings matched those that Kenneth Massey publishes in real time on his website [13]. Our rankings were pretty similar and the comparison is included in Appendix B.

Colley does not have similarly published real-time rankings for the MLB on his website, so to validate that we correctly implemented Colley's Method, we ran our code using simple examples from Colley's original paper (which are included in the Colley's Method section) and reproduced the same exact results as him.

B. Comparing Methods

We did the calculations included in the next table for every year from 2012 to 2021 (excluding 2020). We include the results below for 2012 for illustration purposes and include the rest of the results in Appendix C. For each year, there are two tables containing the teams that made the post-season for that given year. The first table gives the winning percentage, Massey rating, Colley rating, and record for each team. The second table gives the rankings based on the winning percentages, Massey ratings, and Colley ratings, and we compare these to how that team actually finished in the post-season. The post-season rank is a simple ordered list based on when each team lost in the post-season. Note that the post-season rank results in sets of teams that tie because they exit the post-season during the same rounds within their respective leagues.

The measurement that we devised to determine which method most closely matched the post-season outcome is what we will call the *cumulative spread* (CS). To calculate this, we took the absolute value of the difference between the rank

Ratings and Rankings for 2012 MLB Season

Team	Win %	Massey Rating	Colley Rating	Record
WAS	0.605	0.667	0.581	98-64
CIN	0.599	0.195	0.555	97-65
NYA	0.586	1.059	0.617	95-67
OAK	0.580	0.818	0.607	94-68
SFN	0.580	0.232	0.551	94-68
ATL	0.580	0.479	0.561	94-68
TEX	0.574	0.771	0.593	93-69
BAL	0.574	0.317	0.603	93-69
DET	0.543	0.372	0.550	88-74
SLN	0.543	0.404	0.504	88-74

Team	Win % Rank	Massey Rank	Colley Rank	Post-Season
WAS	1	5	6	5
CIN	2	13	9	5
NYA	3	1	1	3
OAK	4	3	2	5
SFN	4	12	10	1
ATL	4	7	8	9
TEX	7	4	4	9
BAL	7	11	3	5
DET	9	10	11	2
SLN	9	9	14	3

that each method produced and the actual post-season rank for each team. We then added up the absolute differences for all of the teams to get the cumulative spread. For example, in the table above for 2012, we calculate the cumulative spread for winning percentage as

$$\begin{aligned} CS_{win\%} &= |1-5| + |2-5| + |3-3| + |4-5| + |4-1| \\ &\quad + |4-9| + |7-9| + |7-5| + |9-2| + |9-3| \\ &= 4+3+0+1+3+5+2+2+7+6 \\ &= 33. \end{aligned}$$

The cumulative spread results for each method for all of the years spanning 2012-2021 (excluding 2020) are listed in the table below.

Cumulative Spread (CS) Results

Year	Win % CS	Massey CS	Colley CS
2012	33	50	47
2013	19	31	16
2014	25	47	39
2015	30	57	36
2016	19	33	40
2017	17	18	18
2018	30	31	34
2019	25	31	30
2021	26	33	38
Average CS	24.9	36.8	33.1

We see that winning percentage has a significantly lower average cumulative spread than Colley's Method and Massey's Method, with Colley's Method being slightly lower than Massey's Method.

C. Discussion

The results show that the winning percentage method for ranking seems to do the best in terms of corresponding with actual post-season outcomes, and in fact, much better than Colley's Method or Massey's Method, which performed similar to each other. Since winning percentage is baseball's prevailing method of choice for ranking teams, these results are perhaps reassuring and unsurprising.

There are a few plausible explanations for these results. First, the need to factor in strength of schedule is perhaps less necessary for the MLB and professional sports leagues in general. Since every team is professionally developed, it makes sense that even the weakest team could beat the strongest team on any given day. Indeed, it happens regularly throughout the MLB season and even in the post-season. Therefore, factoring in strength of schedule to the ranking calculations, as is done in Colley's Method, may not be the best way to determine post-season outcomes. In the same way, point differential might not truly be an indicator of good and bad teams in the MLB. Even the worst professional team could have a break out day with really good hitting and/or pitching that inflates their point differential numbers. Whereas, in a level like college sports, it is more uncommon to see very bad teams run up the numbers on very good teams because of the greater disparity in skill set. Therefore, a method like Massey's method that factors in point differential may be more well-suited for leagues with a wider range of skill levels.

For the MLB, winning percentage does seem to be relatively good at matching post-season outcomes compared to the other two methods discussed. Perhaps other factors that would be more important to consider instead of strength of schedule or point differential would the the health, depth, and performance of a team's pitching staff near the end of the regular season. This is because baseball game outcomes are highly dependent on the starting pitchers. Ken Massey even suggests a method for doing this on his website [16]. Further research could be done to analyze the effectiveness of a model like this.

VI. Conclusion

Massey's Method for ranking teams hinges on incorporating the point differential between teams, and Colley's Method for ranking teams hinges on incorporating the strength of schedule among teams. They are both more sophisticated methods than just using traditional winning percentage to rank teams. However, our results showed that winning percentage was the most accurate method for matching post-season outcomes for the MLB. We hypothesize this is because neither point differential or strength of schedule are very meaningful in the MLB and most major professional leagues in general because of the spread of comparable talent across the league. All in all, the "best" ranking method for a league or division is most likely highly dependent on the underlying characteristics of that group.

Acknowledgments

Thank you to Professor Johnson (and Cookie) for a wonderful semester of learning in 18.065: Matrix Methods in Data Analysis, Signal Processing, and Machine Learning.

References

[1] Amy N. Langville and Carl D. Meyer, "Introduction to Ranking," in *Who's #1? The Science of Rating and Ranking*, Princeton, NJ, USA: Princeton University Press, 2012, ch. 1.

[2] J. Niesen, "The Computer Poll Uprising: Creators of the BCS's Most Controversial Component Look Back," Sports Illustrated. [Online]. Available: https://www.si.com/college/2018/07/11/bcs-computer-rankings-polls-formula-sagarin-billingsley

[3] "BCS Formula," Word Press. [Online]. Available: https://bcsknowhow.wordpress.com/bcs-formula/

[4] K. Massey. Statistical models applied to the rating of sports teams. *Bluefield College*, 1997.

[5] Vaziri, B., Dabadghao, S., Yih, Y., & Morin, T. L. (2018). Properties of sports ranking methods. *Journal of the Operational Research Society, 69*(5), 776-787. https://doi.org/10.1057/s41274-017-0266-8

[6] Amy N. Langville and Carl D. Meyer, "Colley's Method," in *Who's #1? The Science of Rating and Ranking*, Princeton, NJ, USA: Princeton University Press, 2012, ch. 2.

[7] Chartier, T., Kreutzer, E., Langville, A., Pedings, & K. *Bracketology: How can math help?* Department of Mathematics, College of Charleston, 2023.

[8] Massimo Franceschet and Enrico Bozzo (2018). The Massey's Method for Sport Rating: A Network Science Perspective. *CoRR*. http://arxiv.org/abs/1701.03363

[9] *Colley's Bias Free College Football Ranking Method: The Colley Matrix Explained*. Princeton University. [Online]. Available: https://www.ams.org/arc/styleguide/mit-2.pdf

[10] *The Rule of Succession*. University of Victoria. [Online]. Available: https://www.ece.uvic.ca/bctill/papers/mocap/Zabell_1989.pdf

[11] Amy N. Langville and Carl D. Meyer, "Colley's Method," in *Who's #1? The Science of Rating and Ranking*, Princeton, NJ, USA: Princeton University Press, 2012, ch. 3.

[12] *Baseball Reference*. [Online]. Available: https://www.baseball-reference.com/

[13] *Massey Ratings*. [Online]. Available: https://masseyratings.com/

[14] *NCAA Statistics*. [Online]. Available: https://stats.ncaa.org/

[15] *Massey Ratings Description*. [Online]. Available: https://masseyratings.com/theory/massey.htm#rating

[16] *Massey Making Predictions*. [Online]. Available: https://masseyratings.com/theory/predict.htm

Appendix A
Team Abbreviations

Below are the abbreviations for each team, as used in later appendices.

Team Name	Abbreviation
Arizona Diamondbacks	ARI
Atlanta Braves	ATL
Baltimore Orioles	BAL
Boston Red Sox	BOS
Chicago Cubs	CHN
Chicago White Sox	CHA
Cincinnati Reds	CIN
Cleveland Indians	CLE
Colorado Rockies	COL
Detroit Tigers	DET
Miami Marlins	MIA
Houston Astros	HOU
Kansas City Royals	KCA
Los Angeles Angels	LAN
Los Angeles Dodgers	ANA
Milwaukee Brewers	MIL
Minnesota Twins	MIN
New York Mets	NYN
New York Yankees	NYA
Oakland Athletics	OAK
Philadelphia Phillies	PHI
Pittsburgh Pirates	PIT
San Diego Padres	SDN
San Francisco Giants	SFN
Seattle Mariners	SEA
St. Louis Cardinals	SLN
Tampa Bay Rays	TBA
Texas Rangers	TEX
Toronto Blue Jays	TOR
Washington Nationals	WAS

Appendix B
Massey's Method Validation

Using data through May 11th, 2023, below is a table of our calculated Massey's rankings and the ratings that Kenneth Massey actually had on his website. As we can see, most of the rankings we produced are within a few of the rankings Massey has on his website. However, there are a few, like CHN, SLN, or MIA, that are off by a significant amount. These discrepancies are because Massey has made his method more sophisticated over the years, deviating from the simple formulation described and implemented in our paper. In particular, his website rankings now incorporate applying a Bayesian win-loss correction to the rating, which rewards teams that win consistently [15].

2023 MLB Massey Rankings Comparison

Team	Computed Rating	Computed Ranking	Website Ranking
TBA	2.827	1	1
TEX	1.941	2	5
LAN	1.320	3	4
ATL	1.277	4	2
CHN	1.190	5	19
BOS	1.006	6	6
HOU	0.854	7	9
BAL	0.741	8	3
NYA	0.501	9	8
TOR	0.500	10	7
MIL	0.360	11	10
PIT	0.345	12	12
MIN	0.245	13	11
SDN	0.207	14	13
SEA	0.172	15	18
ARI	0.117	16	14
ANA	-0.042	17	15
SLN	-0.328	18	27
WAS	-0.329	19	24
PHI	-0.600	20	16
DET	-0.683	21	20
SFN	-0.759	22	21
CIN	-0.761	23	25
COL	-0.815	24	26
NYN	-0.969	25	22
CLE	-1.053	26	23
KCA	-1.212	27	29
CHA	-1.297	28	28
MIA	-1.329	29	17
OAK	-3.426	30	30

Appendix C
Ratings and Rankings Results

Below are ratings and rankings for each MLB season from 2013 to 2021, excluding 2020 because of a different post-season format due to the COVID-19 pandemic. For each year, there are two tables containing the teams that made the post-season for that given year. The first table gives the winning percentage, Massey rating, Colley rating, and record for each team. The second table gives the rankings based on the winning percentages, Massey ratings, and Colley ratings, and we compare these to how that team actually finished in the post-season. Note that the post-season rank results in teams that tie because they exit the post-season during the same rounds within their respective leagues. These tables were used to calculate the cumulative spreads for each year.

Ratings and Rankings for 2013 MLB Season

Team	Win %	Massey Rating	Colley Rating	Record
BOS	0.599	1.268	0.605	97-65
SLN	0.599	1.032	0.588	97-65
OAK	0.593	0.843	0.578	96-66
ATL	0.593	0.580	0.569	96-66
PIT	0.580	0.328	0.573	94-68
DET	0.574	1.036	0.566	93-69
LAN	0.568	0.224	0.557	92-70
CLE	0.568	0.545	0.560	92-70
TBA	0.564	0.475	0.573	92-71
CIN	0.556	0.620	0.552	90-72

Team	Win % Rank	Massey Rank	Colley Rank	Post-Season
BOS	1	1	1	1
SLN	1	3	2	2
OAK	3	4	3	5
ATL	3	7	6	5
PIT	5	12	4	5
DET	6	2	7	3
LAN	7	13	9	3
CLE	7	8	8	9
TBA	9	9	5	5
CIN	10	5	10	9

Ratings and Rankings for 2014 MLB Season

Team	Win %	Massey Rating	Colley Rating	Record
ANA	0.605	1.023	0.611	98-64
BAL	0.593	0.763	0.602	96-66
WAS	0.593	0.650	0.570	96-66
LAN	0.580	0.385	0.550	94-68
DET	0.556	0.367	0.563	90-72
SLN	0.556	-0.059	0.540	90-72
KCA	0.549	0.226	0.557	89-73
OAK	0.543	1.094	0.555	88-74
PIT	0.543	0.137	0.529	88-74
SFN	0.543	0.112	0.516	88-74

Team	Win % Rank	Massey Rank	Colley Rank	Post-Season
ANA	1	2	1	5
BAL	2	3	2	3
WAS	2	5	3	5
LAN	4	6	7	5
DET	5	7	4	5
SLN	5	16	9	3
KCA	7	9	5	2
OAK	8	1	6	9
PIT	8	11	13	9
SFN	8	12	14	1

Ratings and Rankings for 2015 MLB Season

Team	Win %	Massey Rating	Colley Rating	Record
SLN	0.617	0.453	0.601	100-62
PIT	0.605	0.316	0.587	98-64
CHN	0.599	0.215	0.583	97-65
KCA	0.586	0.713	0.605	95-67
TOR	0.574	1.591	0.584	93-69
LAN	0.568	0.103	0.536	92-70
NYN	0.556	0.023	0.516	90-72
TEX	0.543	0.359	0.553	88-74
NYA	0.537	0.756	0.553	87-75
HOU	0.531	0.883	0.542	86-76

Team	Win % Rank	Massey Rank	Colley Rank	Post-Season
SLN	1	6	2	5
PIT	2	11	3	9
CHN	3	13	5	3
KCA	4	4	1	1
TOR	5	1	4	3
LAN	6	15	11	5
NYN	7	18	14	2
TEX	8	9	6	5
NYA	9	3	7	9
HOU	10	2	8	5

Ratings and Rankings for 2016 MLB Season

Team	Win %	Massey Rating	Colley Rating	Record
CHN	0.636	1.309	0.607	103-58
TEX	0.586	0.175	0.595	95-67
WAS	0.586	0.608	0.558	95-67
CLE	0.584	0.594	0.586	94-67
BOS	0.574	1.265	0.588	93-69
LAN	0.562	0.360	0.532	91-71
TOR	0.549	0.751	0.566	89-73
BAL	0.549	0.441	0.569	89-73
NYN	0.537	0.068	0.514	87-75
SFN	0.537	0.346	0.511	87-75

Team	Win % Rank	Massey Rank	Colley Rank	Post-Season
CHN	1	1	1	1
TEX	2	13	2	5
WAS	2	4	7	5
CLE	4	5	4	2
BOS	5	2	3	5
LAN	6	8	12	3
TOR	6	3	6	3
BAL	8	7	5	9
NYN	9	16	13	9
SFN	9	9	16	5

Ratings and Rankings for 2017 MLB Season

Team	Win %	Massey Rating	Colley Rating	Record
LAN	0.642	0.932	0.620	104-58
CLE	0.630	0.1.485	0.623	102-60
HOU	0.623	1.222	0.621	101-61
WAS	0.599	0.565	0.564	97-65
BOS	0.574	0.840	0.580	93-69
ARI	0.574	0.754	0.563	93-69
CHN	0.568	0.605	0.548	92-70
NYA	0.562	1.279	0.570	91-71
COL	0.537	0.272	0.529	87-75
MIN	0.525	0.233	0.532	85-77

Team	Win % Rank	Massey Rank	Colley Rank	Post-Season
LAN	1	4	3	2
CLE	2	1	1	5
HOU	3	3	2	1
WAS	4	8	6	5
BOS	5	5	4	5
ARI	5	6	7	5
CHN	7	7	8	3
NYA	8	2	5	3
COL	9	9	10	9
MIN	10	10	9	9

Ratings and Rankings for 2018 MLB Season

Team	Win %	Massey Rating	Colley Rating	Record
BOS	0.667	1.209	0.644	108-54
HOU	0.636	1.545	0.631	103-59
NYA	0.617	0.935	0.598	100-62
OAK	0.599	0.822	0.595	97-65
MIL	0.589	0.602	0.588	96-67
CHN	0.583	0.700	0.581	95-68
LAN	0.564	1.249	0.574	92-71
CLE	0.562	0.623	0.520	91-71
COL	0.558	0.388	0.569	91-72
ATL	0.556	0.627	0.556	90-71

Team	Win % Rank	Massey Rank	Colley Rank	Post-Season
BOS	1	3	1	1
HOU	2	1	2	3
NYA	3	4	3	5
OAK	4	5	4	9
MIL	5	9	5	3
CHN	6	6	6	9
LAN	7	2	7	2
CLE	8	8	14	5
COL	9	13	8	5
ATL	10	7	9	5

Ratings and Rankings for 2019 MLB Season

Team	Win %	Massey Rating	Colley Rating	Record
HOU	0.660	1.447	0.635	107-55
LAN	0.654	1.723	0.658	106-56
NYA	0.636	0.974	0.603	103-59
MIN	0.623	0.663	0.577	101-61
ATL	0.599	0.772	0.601	97-65
OAK	0.599	0.802	0.579	97-65
TBA	0.593	0.465	0.564	96-66
WAS	0.574	0.968	0.578	93-69
SLN	0.562	0.780	0.575	91-71
MIL	0.549	0.275	0.568	89-73

Team	Win % Rank	Massey Rank	Colley Rank	Post-Season
HOU	1	2	2	2
LAN	2	1	1	5
NYA	3	3	3	3
MIN	4	9	7	5
ATL	5	7	4	5
OAK	5	5	5	9
TBA	7	12	10	5
WAS	8	4	6	1
SLN	9	6	8	3
MIL	10	13	9	9

Ratings and Rankings for 2021 MLB Season

Team	Win %	Massey Rating	Colley Rating	Record
SFN	0.660	1.234	0.631	107-55
LAN	0.654	1.567	0.627	106-56
TBA	0.617	1.275	0.627	100-62
HOU	0.586	1.202	0.601	95-67
MIL	0.586	0.441	0.553	95-67
CHA	0.574	0.794	0.577	93-69
BOS	0.568	0.597	0.586	92-70
NYA	0.568	0.378	0.585	92-70
SLN	0.556	-0.009	0.525	90-72
ATL	0.547	0.704	0.512	88-73

Team	Win % Rank	Massey Rank	Colley Rank	Post-Season
SFN	1	3	1	5
LAN	2	1	3	3
TBA	3	2	2	5
HOU	4	4	4	2
MIL	4	9	11	5
CHA	6	6	8	5
BOS	7	8	5	3
NYA	7	11	6	9
SLN	9	13	12	9
ATL	10	7	13	1