Incorporating Pre-Game Rankings to Colley, Massey, and Elo Methods

Jiayin (Joy) Liu

Abstract

Colley, Massey, and Elo are three popular rating methods commonly used in the NFL and Major Division I NCAA sports where game results from the regular seasons can be used to seed major playoff brackets or create predictive rankings. As with most rating systems, an ample amount of game data is needed for the resulting rankings to be reflective of the competitors' strength relative to others. However, in situations such as a professional golf match play tournament, where there is a lack of available match play data, or initial weeks into regular NFL season, when it is too early to have the appropriate amount of needed data in hand, how can we produce rankings that meet the standard of accuracy? To address this problem, we propose two possible seeding methods that incorporate pre-game rankings to make up for the lack of relevant games that are adaptive to Colley, Massey, and Elo.

1. Introduction

Every spring, people watch in anticipation for one of the major championships in professional golf, the Dell Technologies Match Play. With its unique mix of formats, where 64 competitors are split into groups of 4s that engage in round-robin play, and only 16 group winners emerge and advance to the single-elimination rounds, the tournament is well deserving of its the title as the March Madness of the PGA world (Bryant, 2020). As the pairings for the final elimination rounds come around at the end of the round-robin section of the tournament, it is only natural that golf enthusiasts and analytics worldwide would make their attempts in predicting the outcome. Many rating methods are developed for this purpose of measuring the strength of competitors, whether it be for sports or other competitions, and making them comparable to produce rankings that can be predictive of outcomes. This paper focuses on three popular rating methods that can be potentially adapted to ranking the game of golf: the Colley method, based on win-loss; the Massey method, which takes account of the score differentials of the competitors; and finally, the Elo Rating System, which originated from measuring competitors' skill in the game of chess (Langville & Meyer, 2012).

Although the format of the Dell Technologies Match Play much resembles the Division I NCAA men's basketball tournament and NFL playoffs, where many weeks of regular season statistics can be employed to establish the comparative strengths of the teams within the league, previous game statistics for the PGA tour competitors are only available in the form of the Official World Golf Ranking (OWGR) and various traditional stroke play tournaments. Moreover, although some information can be extracted from the round-robin aspect of the Dell Technologies Match Play, the small groups of 4s in which the competitors are distributed can fail to establish competitors' skill differences relative to others in groups outside their own. Recent work in the field had sought to address this particular issue of lack of previous play statistics. In sports similar to golf such as tennis and squash, where not all competitors within the league had played with each other frequently enough to produce ratings that are accurate reflections of their skill differences, Elo and other matrix-based methods have been previously used to account for the insufficiency in data for these sports (Bozóki & Temesi, 2016; Clarke, 2011; Dahl, 2011).

In this paper, the aim is to propose an adaptation of the Colley, Massey, and Elo methods, in which a preexisting ranking before the game can be incorporated to better illustrate the comparative strength of the competitors in the form of pre-seeding. The goal is to determine whether this rank seeding method can address the lack of past game plays in tournaments where such information is not widely available. The method will be applied in two different sports to measure its effectiveness and where it can be best utilized. The first is seeding the OWGR to the Dell Technologies Match Play to the round-robin play results to create rankings predictive of the final single-elimination outcomes. The second includes incorporating the preseason power ranking into the weekly NFL games to create week-by-week rankings that are predictive of the outcomes in the following game week.

2. Algorithms

Colley has its basis in one of the simplest forms of rating systems—winning percentage—where for i team's rating, r_i, where $u_w$is the total number of wins and v_{tol} is the total number of game played, is equal to:

$$r_i = \frac{u_w}{v_{tol}}$$

the method moves from this formulation to a linear system. It seeks to address the biases inherent in the winning percentage such as not accounting for opponent strength and unable to give teams ratings without a single win. Furthermore, the method only accounts for the wins and losses (Langville & Meyer, 2012, 21). This aspect of the method is useful in rating golf match play competitors, where each round can be marked by 1-0 or 0-1.

Massey is also a linear system, but unlike Colley, it stems from a least-square formulation and incorporates score differentials as its primary component (Massey, 1997). With that in mind, when seeding the pre-game rankings, it can potentially better reflect the skill difference between each competitor by including the ratings that created the rankings as scores (details will be explained in section 4). For example, the OWGR is based on a point system that rewards players a certain amount of points for competing and placing, varied by the level of competition and results (Jones, Webb & Wilson, 2016). As part of incorporating the rankings, determine whether including the point differences between each player will also be an aspect discussed later in the paper (Jones, Webb & Wilson, 2016).

The Elo method, a Bradley-Terry model-based system with its current form based on a logistic function, has a player rating updating formula with new rating *R'* equal to:

$$R' = R + K(A - E)$$

where *A* is the actual score minus *E*, the expected score. *K* is the K factor that determines the maximum value in which one's rating can increase or decrease, and the higher K factor indicates a higher weight given to the pre-game performance (Langville & Meyer, 2012, 53).

The methods introduced above are the main focus of the paper and the application process. The proposed rank seeding methods are adapted to each of the three methods.

3. Methods

3.1 Dominance and Weak Dominance Graphs

To best reflect the ranking in the dataset, dominance and weak dominance graphs are the two possible methods of representation. As seen in table 1. below, the basis behind both methods is that, suppose an arbitrary ranking of four players A, B, C, and D, the dominance graph shows that A beats B, C, and D, while B beats C and D, and finally, C beats D, whereas the weak dominance graph shows that A beats B, B beats C, and C beats D (refer to figure 2.1. and 2.2. below). With these formats, we are able to generate games that are reflective of the ranking that can then be included in our dataset along with the actual games. By including either the weak dominance graph or the dominance graph in the game dataset used for calculating ratings, we can establish the strength of each competitor relative to others, especially when there is a lack of games played between the competitors in the field.

Ranking	Players
1	A
2	B
3	C
4	D

Table 1.

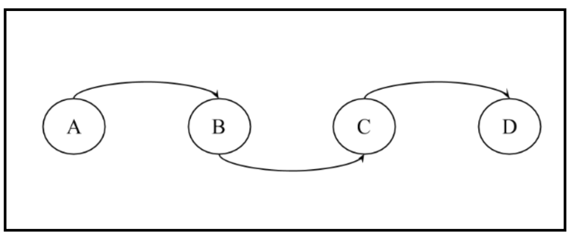


Figure 2.1. *Weak dominance graph*

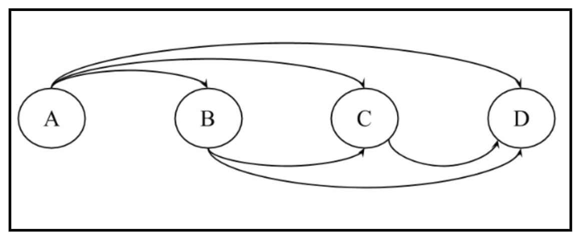


Figure 2.2. *Dominance graph*

3.2. Rank Weighting

Standard Colley and Massey weight all games equally regardless of when the games are played, and, without modification, both the pre-game ranking and actual games will be counted equally. However, several factors show that this does not accurately reflect reality. First, the competitor may improve over time. The newer games should be a better representation of a competitor's strength than past games. Second, most pre-game rankings are the product of a competitor's accumulated performance over time before the actual game. For example, the NFL's power ranking is based on weeks of preseason games, so there lies the potential debate of whether the team's performance during this period can or cannot be translated to performance during the regular season. For golf, the OWGR is based on mostly stroke play tournaments, but the Dell Technologies Match Play is in the match play format. Many may argue that one's ability to play stroke play may not be the same as one's ability to play match play. For these reasons, we will be weighing the pre-game ranking portion of the dataset used for the rating algorithms as one-tenth of the actual games (Chartier et al., 2011)

4. The Data

In determining whether the rank seeding methods either in dominance or weak dominance can be effective in producing better rating as compared to the standard methods, the paper will focus on applying the algorithms to two different games, the Dell Technologies Match Play and the regular season of the NFL. All the data will be in the Matlab game format of:

- days since 1-1-0000
- dates in YYYYMMDD
- team 1 index
- team 1homefield (1 as home, -1 as away, and 0 as neutral)
- team 1 score
- team 2 index
- team 2 homefield (1 as home, -1 as away, and 0 as neutral)
- team 2 score

4.1. Dell Technologies Match Play Ranking

Data from the year of 2015 and 2016 are used to test the rank seeding methods. The OWGR is updated weekly. For the year of 2015, the ranking used is from April 26, the same ranking used by the tournament officials for seeding and grouping. For the year of 2016, the ranking is taken from March 20. The games used for creating the predictive rankings for the single elimination section are the round-robin section of the tournament in the win-loss format (refer to table 2.1., 2.2., and 2.3. for sample of the data format used to simulate the dominance and weak dominance graph relationship).

Since the Dell Technologies Match Play uses the OWGR for seeding and grouping, the player indexes are the same as their rankings. As illustrated above in table 2.2 and 2.3., both use the win-loss format to simulate the ranks, since both Colley and Elo use only the win-loss information to produce ratings. However, Massey can take point differences into account. Since the OWGR ranks players based off of average points, which is calculated by the total points earned divided by the total number of tournaments played, an addition to the win-loss rankings, we have also created separate datasets that employ these average points for Massey (see table 3.)

Index	Player
1	Jordan Spieth
2	Jason Day
3	Rory McIlroy
4	Bubba Watson
5	Rickie Fowler
6	Adam Scott
7	Justin Rose
8	Dustin Johnson
9	Patrick Reed
10	Danny Willett

Table 2.1. *Player indexes and names used in table 2.2 and 2.3*

736408	20160320	1	0	1	2	0	0
736408	20160320	1	0	1	3	0	0
736408	20160320	1	0	1	4	0	0
736408	20160320	1	0	1	5	0	0
736408	20160320	1	0	1	6	0	0
736408	20160320	1	0	1	7	0	0
736408	20160320	1	0	1	8	0	0
736408	20160320	1	0	1	9	0	0
736408	20160320	1	0	1	10	0	0

Table 2.2. *Dominance graph generated game in the Matlab game format for the year of 2016*

736408	20160320	1	0	1	2	0	0
736408	20160320	2	0	1	3	0	0
736408	20160320	3	0	1	4	0	0
736408	20160320	4	0	1	5	0	0
736408	20160320	5	0	1	6	0	0
736408	20160320	6	0	1	7	0	0
736408	20160320	7	0	1	8	0	0
736408	20160320	8	0	1	9	0	0
736408	20160320	9	0	1	10	0	0

Table 2.3. *Weak Dominance graph generated game in the Matlab game format for the year of 2016*

736048	20150326	1	0	11.2142	2	0	9.0779
736048	20150326	1	0	11.2142	3	0	7.3498
736048	20150326	1	0	11.2142	4	0	7.1793
736048	20150326	1	0	11.2142	5	0	6.7124
736048	20150326	1	0	11.2142	6	0	6.5835
736048	20150326	1	0	11.2142	7	0	6.4455
736048	20150326	1	0	11.2142	8	0	6.1662
736048	20150326	1	0	11.2142	9	0	5.9196
736048	20150326	1	0	11.2142	10	0	5.6382

Table 3. *Dominance graph generated games with OWGR average points in the Matlab game format for the year of 2016*

4.2. NFL Week-To-Week Rankings

The data from the regular seasons from 2010 to 2019 are used for the trials of the proposed seeding methods. The rankings that are used for seeding are the corresponding years' ESPN preseason power rankings. The games are divided into 17 weeks according to the NFL calendar, and datasets are created as illustrated in figure 3. With each additional week, gameplay data from the new week is added. For example, in week 2, we have the game results from both week 1 and 2. The goal is to use the game available from these two weeks to create rankings that can be used to predict the game results in week 3. Refer to table 4.1., 4.2., and 4.3. for preseason ranking included in the week game results in the form of a dominance graph or weak dominance graph below:

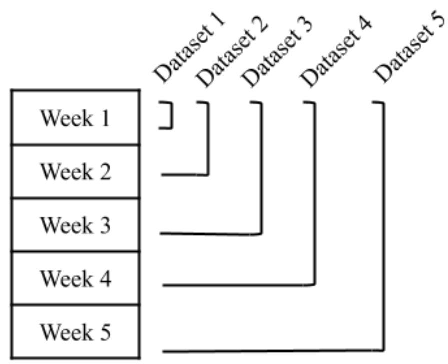


Figure 3. *A demonstration of how the games are divided*

Preseason Ranking	Index	Teams
1	26	Philadelphia
2	22	New England
3	21	Minnesota
4	18	LA Rams
5	27	Pittsburgh
6	23	New Orleans
7	15	Jacksonville
8	13	Houston
9	16	Kansas City

Table 4.1. *2018 NFL team indexes and preseason rankings used in table 4.2. and 4.3.*

365243	10000101	26	0	1	22	0	0
365243	10000101	26	0	1	21	0	0
365243	10000101	26	0	1	18	0	0
365243	10000101	26	0	1	27	0	0
365243	10000101	26	0	1	23	0	0
365243	10000101	26	0	1	15	0	0
365243	10000101	26	0	1	13	0	0
365243	10000101	26	0	1	16	0	0
365243	10000101	26	0	1	31	0	0

Table 4.2. *Dominance graph generated game in the Matlab game format*

365243	10000101	26	0	1	22	0	0
365243	10000101	22	0	1	21	0	0
365243	10000101	21	0	1	18	0	0
365243	10000101	18	0	1	27	0	0
365243	10000101	27	0	1	23	0	0
365243	10000101	23	0	1	15	0	0
365243	10000101	15	0	1	13	0	0
365243	10000101	13	0	1	16	0	0
365243	10000101	16	0	1	31	0	0

Table 4.3. *Weak Dominance graph generated game in the Matlab game format*

5. Results and Analysis

5.1. Dell Technologies Match Play

In determining whether the seeding methods affect the resulting rankings, the produced rankings are compared with the actual results. The dominance graph seeding method was used in Colley, Massey, and Elo for the year of 2015 (refer to appendix table 1A for 2015 dominance graph results). With standard Colley, Massey, and Elo, we could not produce accurate rankings since many competitors received the same ratings, leading to as many as 10 competitors tied in the same place. Moreover, the lack of connection in the game graph resulted in Massey unable to produce any rating at all. The problem lies in the lack of gameplay and the competitors not able to play with others outside of their own groups during the round-robin section of the tournament for them to establish their strengths relative to others. The 2015 results had shown that weighting the dominance graph as one-tenth of the actual gameplay versus as equal made little difference in the ranking produced, as they both produced ranking more similar to the OWGR than the actual results. The dominance graph seeding method for the year of 2016 is not shown in the appendix as it produced similar results as the year of 2015.

For weak dominance graph seeding in both Colley and Massey, the resulting rankings were somewhat predictive for the year of 2015 only (refer to appendix table 2A and 3A for 2015 and 2016 weak dominance rank results), where the rankings were able to place McIlroy, Willett, and Woodland into the top 4s regardless of how much the seeding was weighted. However, in the year of 2016 (as seen in the appendix table 3A), seedings for all three methods were unable to produce rankings remotely similar to the actual results. Though including the seeding was able to produce rankings much better than the standard methods, the small size of the tournament and the lack of gameplay ultimately resulted in inaccurate rankings.

5.2. NFL Week-To-Week Ranking

Predictability was used in determining whether the rating methods are successful in predicting games' outcomes in the following week. The ranking produced was used to compare with the actual outcome. If team A has a higher ranking than team B, and that team A had won the actual game in the following week, the methods have successfully predicted the outcome. The predictability was then calculated by the number of correct predictions divided by the total number of games.

In terms of predictability, the dominance graphs seeding method for Colley performed consistently better than the standard Colley method for the first 2 to 3 weeks for 7 out of 10 years and up to 3 to 4 weeks in some years (refer to appendix graph 4A for 2010 to 2019 predictabilities graphs for Colley). The weak dominance graph seeding methods for Colley had less of a visible difference in performance than the dominances graph method, with it performing better than the standard methods in the year of 2010 and 2019 in the week 1. In some years, the weak dominance graph seeding method underperformed in week 1 but performed better in week 2 and 3, such as the year of 2011, 2015, and 2016.

On the other hand, both the dominance graph and the weak dominance graph seeding method for Massey outperformed the standard method in the first 2 weeks for most years, since the standard method could not even produce rankings for these weeks due to the lack of gameplays and connections from teams to teams. In the year of 2016, 2018, and 2019, the dominance graph seeding methods created rankings more predictable than the standard Massey for the first 7 to 8 weeks (refer to appendix graph 5A for 2010 to 2019 predictabilities charts for Massey).

As time progressed, the dominance and weak dominance graph seeding methods for both Colley and Massey began to produce less predictive rankings than the standard methods, as evident in the decreasing performance in week 8 and after. This result may be due to the fact that as more weeks are added, the preseason rankings were no longer accurate reflections of the team's most recent performances. For Elo, the dominance graph seeding methods performed better than the standard method in the first 2 or 3 weeks for most years with the exception of 2012 and 2018 (refer to appendix graph 6A for 2010 to 2019 predictabilities chart for the Elo method). The weak dominances graphs seeding methods created rankings less predictive than the standard methods in majority of the years, with the exception of 2013 and 2017.

Interestingly, weighting the ranking for dominance and weak dominance graph seeding did not produce dramatically different results. Weighting the rank seeding as equal to the actual games produced very similar results as weighting the rank seeding one-tenth of the actual game.

7. Conclusion

With the inherent randomness as a significant component and other unaccounted factors in golf, it was somewhat expected that neither Colley, Massey, nor Elo methods with or without rank seeding could produce predictive rankings. However, the rank seeding methods were an effective way to make up for the lack of gameplay available early into the NFL season. Despite using the same seeding methods for both Dell Technologies Match Play and the NFL, the two led to very different results. Further comparing the predictability trends from year to year for the NFL, seeding used with Massey outperformed the standard methods for up to 6 or 8 weeks for the majority of the years but performed poorly in other years such as 2013 and 2014. These drastic differences highlight the significant role played by the quality of the pre-game rankings.

The rankings' abilities to reflect competitors or teams' abilities could be detrimental to the performance of the seeding methods. Though the type of games measured for the NFL power ranking was based on preseason games, which were in the same format as the regular season games, many argue that the preseason games still differ from the regular-season games, which usually have more stakes at play and different players. For golf, the OWGR itself is based on mostly stroke play tournaments as opposed to the match play format used by the Dell Technologies Match Play. As mentioned earlier in section 4.2 that one's skill in playing stroke play may not translate to skill in playing match play; this is also shown in the fact that, even within the round-robin aspect of the tournament, many of the lower-ranked players were able to beat higher-ranked players. Moreover, the top 64 players on the OWGR and participating in the Dell Technologies Match Play are all equally skilled and are all likely to win the championship regardless of their rankings. Furthermore, there are possible biases within the OWGR that favor PGA tour players more than players from other tours outside of the United States (Jones, Webb & Wilson, 2016). These varying factors ultimately led to the seeding methods' inability to provide any advantage over the standard methods.

As for the NFL, where the seeding methods were shown to produce better team rankings than the standard methods in the first several weeks, we can potentially use a mix of seeding methods and standard methods in the future. For example, removing the preseason ranking seedings around week 7 or 8 for the dominance graph seedings method for Massey, when the seeded methods are no longer performing better than the standard methods. Similar ideas such as, only using the dominance graph seeding methods for the first 2 weeks and then removing them for the remaining weeks, can also be explored for Colley.

This paper proposes potential rank seeding methods that can be of use in certain games or sports where there is a lack of past games. In incorporating the rankings in dominance and weak dominance graph to Colley, Massey, and Elo, we demonstrate that, specifically for the NFL, we can produce rankings that outperform the standard methods in week-by-week predictabilities in the initial weeks.

8. Appendix

2015 Dell Technologies Match Play – Dominance Graph Seeding Methods Results

Single-Elimination Round Finalists	WR	AR	SC	CD (0.1)	CD (1.0)	SM	MD (0.1)	MD (1.0)	SE	ED (32)
Mcllroy	1	**1**	T1	1	1	N/A	1	1	T1	1
Woodland	50	**2**	T10	49	50	N/A	53	49	T1	48
Willett	48	**3**	T1	48	47	N/A	59	48	T1	47
Furyk	5	**4**	T13	5	5	N/A	5	5	33	5
Casey	36	**T5**	T1	36	35	N/A	37	34	T9	32
Oosthuizen	29	**T5**	T1	27	27	N/A	26	28	T1	27
Senden	60	**T5**	T10	59	60	N/A	51	57	T13	58
Fleetwood	54	**T5**	T13	55	54	N/A	57	55	T13	52
Matsuyama	16	**T9**	T1	15	14	N/A	16	16	T1	13
Schwartzel	37	**T9**	T1	35	36	N/A	35	37	T1	33
Holmes	12	**T9**	T10	11	11	N/A	13	12	T13	11
Fowler	13	**T9**	T10	13	12	N/A	15	13	T9	9
Leishman	56	**T9**	T1	53	55	N/A	50	53	T9	54
Mahan	31	**T9**	T1	30	30	N/A	38	30	T1	29
Grace	38	**T9**	T13	38	38	N/A	36	38	T13	37
Westwood	26	**T9**	T1	25	25	N/A	19	22	T1	24

Table 1A. *2015 Dell Technologies Dominance Graph Seeding Methods Results, where "WR" is the OWGR of the players, "AR" is the actual results, "SC" is the standard Colley method, "CD" is the Colley method with dominance graph seeding, "SM" is the standard Massey method, "MD" is the Massey method with dominance graph seeding, "SE" is the standard Elo method, and "ED" is the Elo method dominance graph seeding. "1.0" and "0.1" are the weight given to the rank, "32" is the standard K factor used for Elo.*

2015 Dell Technologies Match Play Weak Dominance Graph Results

Actual Results		Colley (0.1)		Colley (1.0)		Massey (0.1)		Massey (1.0)		Elo (32)	
1	McIlroy	1	McIlroy	1	McIlroy	1	McIlroy	1	McIlroy	1	McIlroy
2	Woodland	2	Fowler	2	Willett	4	Fowler	2	Willett	2	Westwood
3	Willett	3	Willett	3	Schwartzel	5	Willett	3	Fowler	3	Oosthuizen
4	Furyk	4	Senden	4	Senden	6	Woodland	4	Woodland	4	Schwartzel
T5	Casey	5	Schwartzel	5	Fowler	7	Matsuyama	5	Senden	5	Woodland
T5	Oosthuizen	6	Woodland	6	Woodland	8	Senden	7	Schwartzel	6	Willett
T5	Senden	7	Westwood	7	Westwood	9	Mahan	8	Westwood	7	Matsuyama
T5	Fleetwood	8	Mahan	8	Matsuyama	11	Westwood	9	Matsuyama	8	Mahan
T9	Matsuyama	9	Matsuyama	9	Leishman	12	Leishman	10	Mahan	9	Senden
T9	Schwartzel	10	Leishman	10	Casey	14	Furyk	11	Leishman	10	Fowler
T9	Holmes	11	Oosthuizen	11	Mahan	16	Holmes	12	Casey	11	Leishman
T9	Fowler	12	Casey	12	Oosthuizen	20	Oosthuizen	16	Grace	12	Casey
T9	Leishman	16	Grace	15	Grace	22	Grace	18	Oosthuizen	14	Grace
T9	Mahan	19	Furyk	22	Furyk	25	Schwartzel	21	Furyk	15	Fleetwood
T9	Grace	25	Holmes	25	Fleetwood	28	Fleetwood	24	Holmes	16	Holmes
T9	Westwood	26	Fleetwood	26	Holmes	50	Casey	27	Fleetwood	24	Furyk

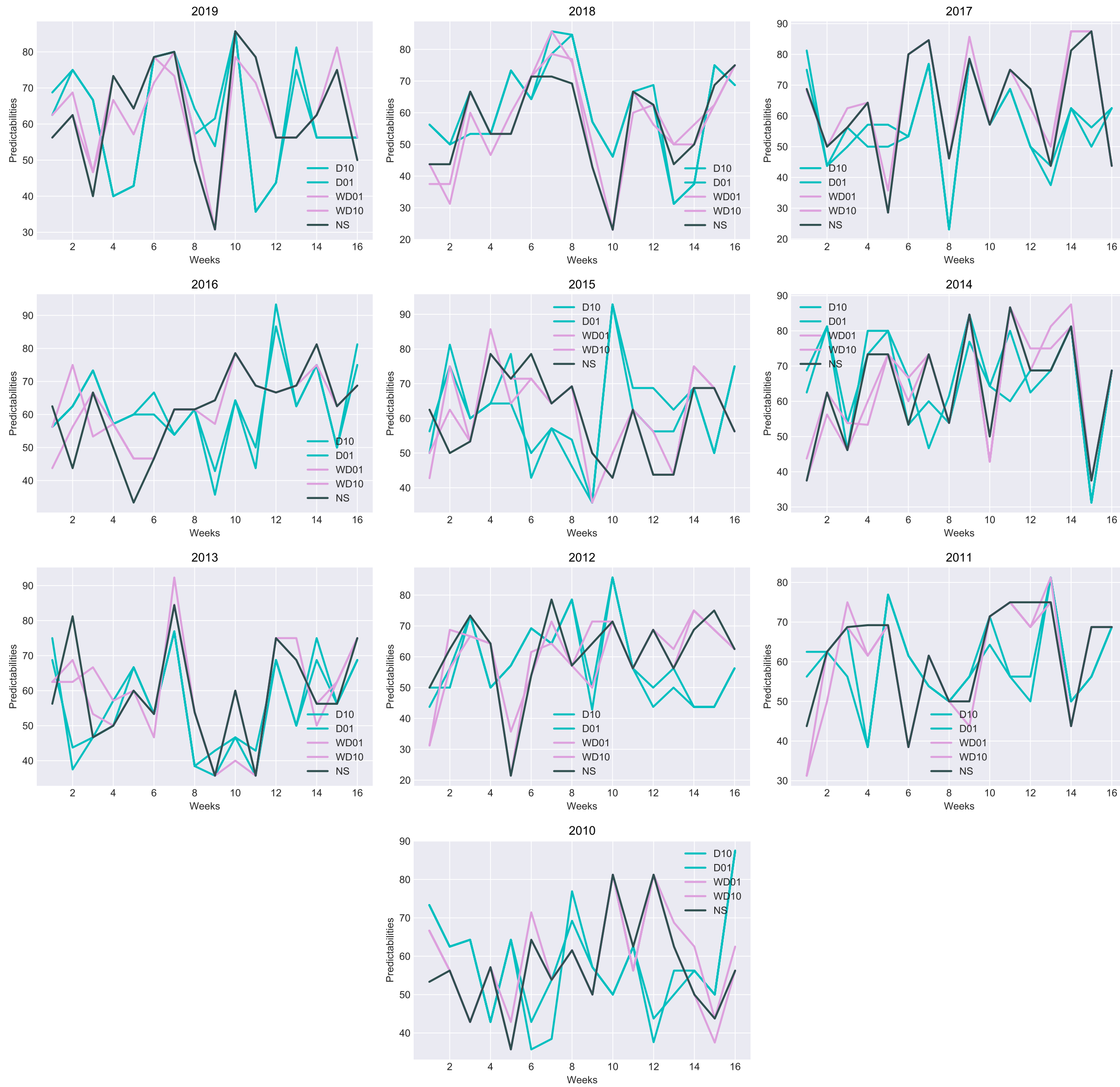
Table 2A. *"T5" or "T9" shows competitors tied for a certain place.*

2016 Dell Technologies Match Play Weak Dominance Graph Results

Actual Results		Colley (0.1)		Colley (1.0)		Massey (0.1)		Massey (1.0)	
1	Day	1	Spieth	1	Spieth	1	Spieth	1	Spieth
2	Oosthuizen	2	Oosthuizen	2	Oosthuizen	5	Reed	2	Reed
3	Cabrera-Bello	3	Reed	3	Reed	6	Koepka	3	Oosthuizen
4	McIlroy	4	Haas	4	Day	7	Haas	4	Day
T5	D Johnson	5	Z Johnson	5	Z Johnson	8	Oosthuizen	6	Snedeker
T5	Moore	6	Cabrera-Bello	6	Snedeker	9	Cabrera-Bello	8	Z Johnson
T5	Kirk	7	Day	7	Haas	14	Kirk	9	Cabrera-bello
T5	Koepka	8	Snedeker	8	Cabrera-Bello	15	Z Johnson	10	Haas
T9	Spieth	9	Kuchar	9	Kuchar	16	Snedeker	11	Koepka
T9	Reed	10	McIlroy	11	McIlroy	18	Day	13	Kuchar
T9	An	16	Koepka	16	Koepka	23	Kuchar	16	D Johnson
T9	Kizzire	17	D Johnson	17	D Johnson	24	D Johnson	19	McIlroy
T9	Haas	18	An	18	An	31	Moore	22	An
T9	Z Johnson	22	Kirk	24	Kirk	42	Kizzire	23	Kirk
T9	Snedeker	24	Moore	26	Moore	43	An	29	Moore
T9	Kuchar	42	Kizzire	44	Kizzire	61	McIlroy	44	Kizzire

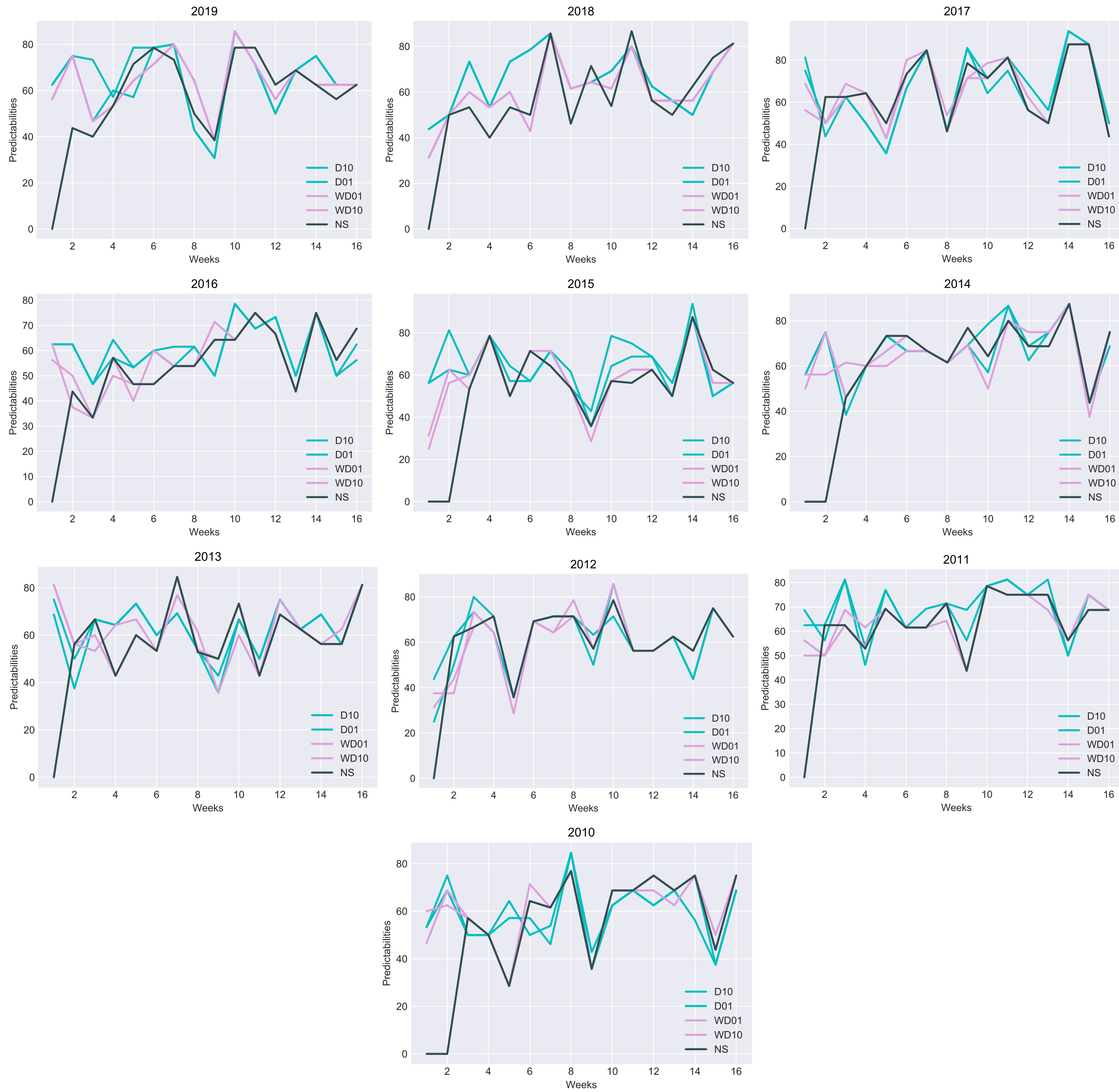
Table 3A. *Elo is not available for the year 2016, when halves in matches are introduced*

2010-2019 NFL Week-to-Week Predictabilities Graphs – Colley



Graph 4A. *"D10"is the dominance graph seeding method with weighting as 1.0, "D01" is the dominance graph seeding method with weighting as 0.1, "WD01" is the weak dominance graph seeding method with weighting as 0.1, "WD10" is the weak dominance graph seeding method with weighting as 0.1, and "NS" is the standard method.*

2010-2019 NFL Week-to-Week Predictabilities Graphs – Massey



Graph 5A. *"D10"is the dominance graph seeding method with weighting as 1.0, "D01" is the dominance graph seeding method with weighting as 0.1, "WD01" is the weak dominance graph seeding method with weighting as 0.1, "WD10" is the weak dominance graph seeding method with weighting as 0.1, and "NS" is the standard method.*

2010-2019 NFL Week-to-Week Predictabilities Graphs – Elo

Graph 6A. *"D10"is the dominance graph seeding method with weighting as 1.0, "D01" is the dominance graph seeding method with weighting as 0.1, "WD01" is the weak dominance graph seeding method with weighting as 0.1, "WD10" is the weak dominance graph seeding method with weighting as 0.1, and "NS" is the standard method.*

9. References

Bozóki, S., Csató, L., & Temesi, J. (2016). An application of incomplete pairwise comparison matrices for ranking top tennis players. *European Journal of Operational Research*, *248*(1), 211-218.

Bryant, R. (2020). *WGC - Dell Match Play | Volunteers*. Pgatourhq.Com. https://volunteers-delltechnologiesmatchplay.pgatourhq.com/tournament-info/competition-format

Chartier, T. P., Kreutzer, E., Langville, A. N., & Pedings, K. E. (2011). Sports Ranking with Nonuniform Weighting. *Journal of Quantitative Analysis in Sports*, *7*(3). https://doi.org/10.2202/1559-0410.1290

Clarke, S. (2011). Rating non-elite tennis players using team doubles competition results. *The Journal of the Operational Research Society*, *62*(7), 1385-1390.

Clarke, S. R. (1994). An adjustive rating system for tennis and squash players. *2nd Conference on Mathematics and Computers in Sport, Gold Coast, Australia*, 43-50.

Dahl, G. (n.d.). A matrix-based ranking method with application to tennis. *Linear Algebra and Its Applications*, *437*(1), 26-36.

Heavlin, W. D., & Pregibon, D. (2013). [Review of the book *Who's #1? The Science of Rating and Ranking*, by A. N. Langville & C. D. Meyer]. *The American Statistician*, *67*(2).

Jones, M. A., Webb, A., & Wilson, J. (2016). Rankings over Time. *The College Mathematics Journal*, *47*(4), 242-248.

Langville, A. N., & Meyer, C. D. (2012). *Who's #1?: The Science of Rating and Ranking*. Princeton University Press.

Massey, K. (1997). *Statistical Models Applied to the Rating of Sports Teams Honors Project: Mathematics*. https://www.masseyratings.com/theory/massey97.pdf