Dominance matrices in AFL - Do they actually work?

Name: Tengfei Wang
Student number: 22555543

Introduction

EXAMPLE

D(M)

	a	b	c
a	0	1.33	0.93
b	0.75	0	2
c	1.08	0.5	0

D(M)T

0	0.75	1.08
1.33	0	2
0.93	2	0

$$\boldsymbol{r}_1 = \boldsymbol{D}(m)\mathbf{1} - (\boldsymbol{D}(m))^T\mathbf{1}$$

$$D(m)\cdot 1 = \begin{matrix}2.26\\2.75\\1.58\end{matrix} \qquad D(m)^T\cdot 1 = \begin{matrix}1.83\\1.83\\2.93\end{matrix}$$

$$R1 = \begin{matrix}2.26-1.83 & +0.43\\2.75-1.83 & =+0.92\\1.58-2.93 & -1.35\end{matrix}$$

- **Team b** = +0.92 → Ranked highest
- **Team a** = +0.43
- **Team c** = -1.35 → Ranked lowest

- **Objective**
  - Evaluate how well dominance matrices can predict AFL rankings before the season ends.
- **What are dominance matrices and why we use it?**
  - *What*: A mathematical approach to predict team rankings based on performance metrics like score ratios or win-loss records.
  - *Why*: Not all teams have played each other in a competition. Dominance matrices provide a way to take this into account and produce a more accurate ranking.
  - In this project, ***score ratios*** was selected to predict the team ranking. It estimates team strength based on match results rather than just win-loss records.
- **Key Questions:**
  - How accurately can dominance matrices predict final AFL rankings?
  - What is the minimum number of rounds (m) needed for reliable predictions?
  - Can adjusting the weighting of ranking components improve accuracy?

Methodology

We can predict the final rankings using r_1, r_2, or higher ranking individually, or by combining two or all of them with different weightings. To determine the most effective approach, we will conduct tests to evaluate their accuracy.

1. **First-order ranking (r1):**
   - Measures direct dominance by calculating the total score ratios a team earned **minus** the score ratios conceded to others
2. **Second-order ranking (r2):**
   - Accounts for the strength of opponents, using squared dominance matrix D(m)^2.
3. **Final ranking formula:**
   - r = r1 + n1 × *r2* + *n2* × r3.
   - Tested multiple values of n1 and n2 to find the best prediction accuracy.
4. **Dominance Matrix $\boldsymbol{D(m)}$:**
   - A matrix that quantifies how much each team has outperformed others based on score ratios.
   - Higher values indicate stronger dominance over opponents.
5. $\boldsymbol{D(m)^2}$
   - Accounts for the strength of opponents by using the squared dominance matrix D(m)^2.
   - This means it not only considers who a team beats, but also how strong those beaten teams are.
   - Teams that defeat strong opponents receive higher r2 values than those beating weaker opponents.

total score ratio a team gains

total score ratio a team loses

$$\boldsymbol{r}_1 = \boldsymbol{D}(m)\mathbf{1} - (\boldsymbol{D}(m))^T\mathbf{1}$$

$$\boldsymbol{r}_2 = (\boldsymbol{D}(m))^2\mathbf{1} - ((\boldsymbol{D}(m))^2)^T\mathbf{1}.$$

$$R_3 = \left(D(m)\right)^3 1 - \left(\left(D(m)\right)^3\right)^T 1$$

$$R = R1 + R2*n1 + R3*n2$$

Data Collection & Calculation

- **Data Source:** AFL 2022 match results from FootyWire.
- **Chosen values of m:** 3, 7, 12, 17, 19, 20, 21, 22.
  - Selected based on key milestones (starting at 10% of the season, then increasing in ~20% intervals, plus key rounds).
- **Constructing the dominance matrix D(m)**
  - Used the ***score ratio*** method, where team dominance is based on points scored against each opponent.
  - A team cannot play itself, so self-play values are set to 0.

Round 16 D(16)		Geelong	Melbourne	Sydney	Collingwood	Fremantle	Brisbane	Richmond	Western Bulldog	Carlton	St Kilda	Port Adelaide	Gold Coast	Hawthorn	Adelaide	Essendon	GWS	West Coast	North Melbourne
	m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Geelong	1	0	1 4/9	77/107	1 1/7	22/23	1 1/7	1 3/86	1 13/70	0	8/9	1 35/47	0	20/23	1 42/55	1 11/12	2 18/35	1 2/7	6 59/122
Melbourne	2	9/13	0	61/73	28/41	28/47	2 11/53	1 11/27	1 26/71	0	1 38/55	1 8/9	1 13/69	1 10/81	1 29/65	1 29/70	2 14/53	2 18/19	1 47/53
Sydney	3	1 30/77	1 12/61	0	0	0	89/113	1 3/50	2 465/731	29/34	2 19/32	59/82	61/75	1 41/68	0	3 28/201	1 5/23	2 5/58	1 11/75
Collingwood	4	7/8	1 13/28	0	0	1 9/11	13/14	86/113	17/33	1 4/75	1 1/5	0	2 100/279	1 1/17	1 21/29	1 11/82	1 1/7	74/87	1 7/81
Fremantle	5	1 1/22	1 19/28	0	11/20	0	1 14/85	0	0	2 2/11	2 393/910	1 8/91	11/23	1 13/82	1 1/82	1 48/59	1 17/27	2 8/47	4 1/4
Brisbane	6	7/8	53/117	1 24/89	1 1/13	85/99	0	0	1 41/67	0	1 7/19	1 11/69	1 13/20	112/117	1 6/11	2 29/150	1 7/48	3 1/2	3 1/4
Richmond	7	86/89	27/38	50/53	1 27/86	0	0	0	1 38/61	1 387/395	28/39	1 12/65	46/47	1 23/94	82/101	1 2/3	1 36/73	4 163/505	0
Western Bulldogs	8	70/83	71/97	1 89/120	1 16/17	0	67/108	61/99	0	15/17	0	69/86	1 19/87	1 42/83	62/63	1 32/71	1 4/21	2 41/60	1 68/71
Carlton	9	0	0	1 5/29	75/79	2 113/436	0	2 68/473	1 2/15	0	26/31	1 3/91	31/46	1 1/73	1 12/17	1 13/27	1 2/5	2 10/53	1 25/32
St Kilda	10	1 1/8	55/93	32/83	5/6	1 13/16	19/26	1 11/28	0	1 5/26	0	42/43	1 26/61	1 69/73	1 7/23	72/107	1 17/60	0	2 3/50
Port Adelaide	11	47/82	9/17	1 23/59	0	91/99	69/80	65/77	1 17/69	91/94	1 1/42	0	1 2/91	7/15	23/24	1 8/25	2 26/29	3 6/11	2 1/2
Gold Coast	12	0	69/82	1 14/61	1 337/476	2 1/11	20/33	1 1/46	87/106	1 15/31	61/87	91/93	0	2 13/54	1 43/73	0	57/83	1 27/80	2 15/47
Hawthorn	13	1 3/20	81/91	68/109	17/18	82/95	1 5/112	94/117	83/125	73/74	73/142	2 1/7	54/121	0	1 16/27	3/4	25/36	0	1 10/29
Adelaide	14	55/97	65/94	0	29/50	82/83	11/17	1 19/82	1 1/62	17/29	23/30	1 1/23	73/116	27/43	0	99/103	54/113	1 31/57	1 57/58
Essendon	15	12/23	70/99	1 523/947	82/93	59/107	1 772/873	3/5	71/103	27/40	1 35/72	25/33	0	1 1/3	1 4/99	0	0	97/107	0
GWS	16	35/88	53/120	23/28	7/8	27/44	48/55	73/109	21/25	5/7	60/77	29/84	1 26/57	1 11/25	2 5/54	0	0	1 26/43	1 49/53
West Coast	17	7/9	19/56	58/121	1 13/74	47/102	2/7	1 31/470	60/161	53/116	0	11/39	80/107	0	57/88	1 10/97	43/69	0	59/74
North Melbourne	18	576/793	53/100	75/86	81/88	4/17	4/13	0	71/139	32/57	50/103	2/5	47/109	29/39	58/115	0	53/102	1 15/59	0

Note: the score ratio highlighted in red indicates that the two teams have played each other twice.

Comparison of predicted rankings vs. actual ladder

Dominance matrix approach has varying prediction accuracy across different rounds in the AFL.

1. **Low Accuracy in Early Rounds (All m Values)**
   - In Rounds 3 and 7, all methods yielded low accuracy (mostly under 22%).
   - This makes sense because early in the season, teams' strengths are uncertain, and dominance relationships are not well established.

2. **Improved Accuracy in Later Rounds (All m Values)**
   - By Round 16 and beyond, accuracy improves significantly.
   - Some models (e.g., R1 only, R1+R2*n1) reach 50%+ accuracy in later rounds (Rounds 19-21).
   - This suggests that the dominance matrix is capturing team performance trends better as more games are played.

3. **Impact of Score Margins on Prediction Accuracy**
   - Teams that win by small margins tend to have lower predicted ranking than their actual standings
   - Example: Collingwood often wins by narrow margins, making their dominance ranking appear lower than their actual ladder position

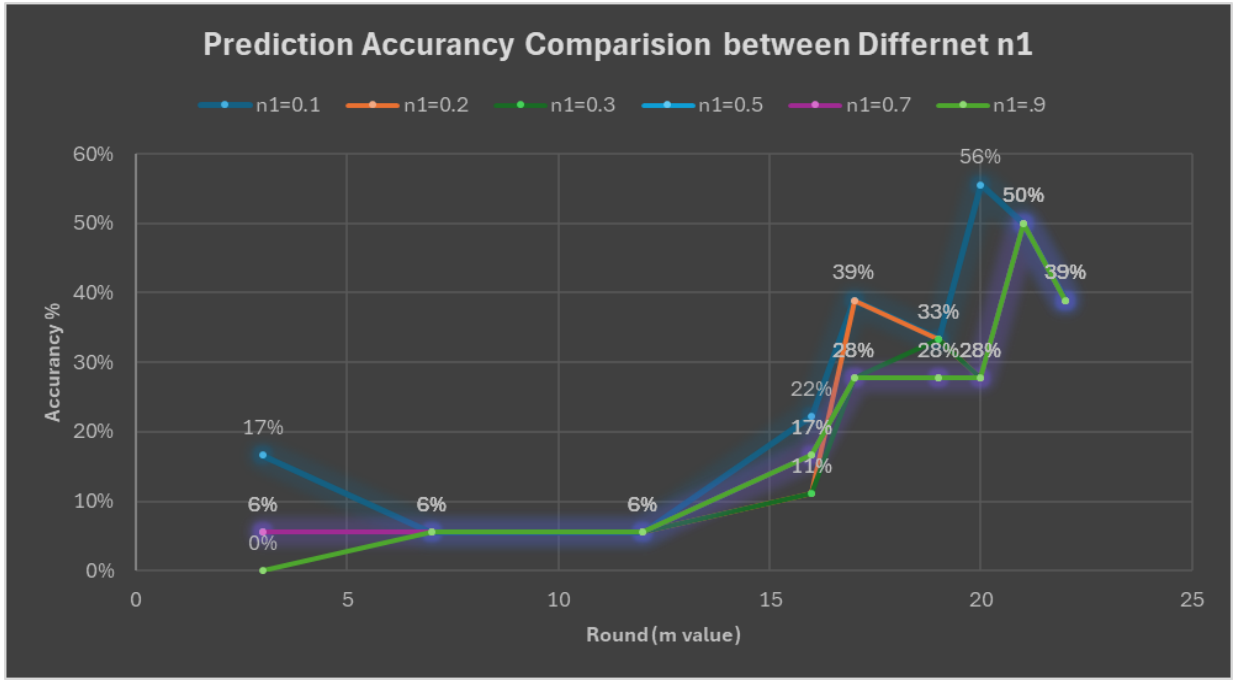
R1+R2*n1
D(3-22)

n1=0.1

Accurancy	17%	6%	6%	22%	39%	33%	56%	50%	39%	
Team	Round 3	Round 7	Round 12	Round 16	Round 17	Round 19	Round 20	Round 21	Round 22	Actual Ranking
Geelong	7	6	3	2	1	1	1	1	1	1
Melbourne	3	1	1	1	2	2	2	2	2	2
Sydney	9	5	6	5	4	4	3	3	3	3
Collingwood	4	8	11	10	11	11	10	10	10	4
Fremantle	5	3	4	4	5	5	5	5	6	5
Brisbane	1	2	2	3	3	3	4	4	4	6
Richmond	11	7	8	6	6	9	7	6	5	7
Western Bulldogs	10	10	9	8	10	7	8	8	9	8
Carlton	6	11	10	9	7	6	6	7	7	9
St Kilda	8	4	5	11	12	12	12	12	12	10
Port Adelaide	15	12	12	12	9	8	9	9	8	11
Gold Coast	13	14	7	7	8	10	11	11	11	12
Hawthorn	2	9	13	14	13	13	13	13	13	13
Adelaide	14	15	15	15	16	16	15	15	15	14
Essendon	16	16	16	16	15	14	14	14	14	15
GWS	12	13	14	13	14	15	16	16	16	16
West Coast	18	18	18	17	17	17	17	17	17	17
North Melbourne	17	17	17	18	18	18	18	18	18	18

R1+R2*n1
D(3-22)

n1=
- 0.1,
- 0.2,
- 0.3,
- 0.5,
- 0.7,
- 0.9

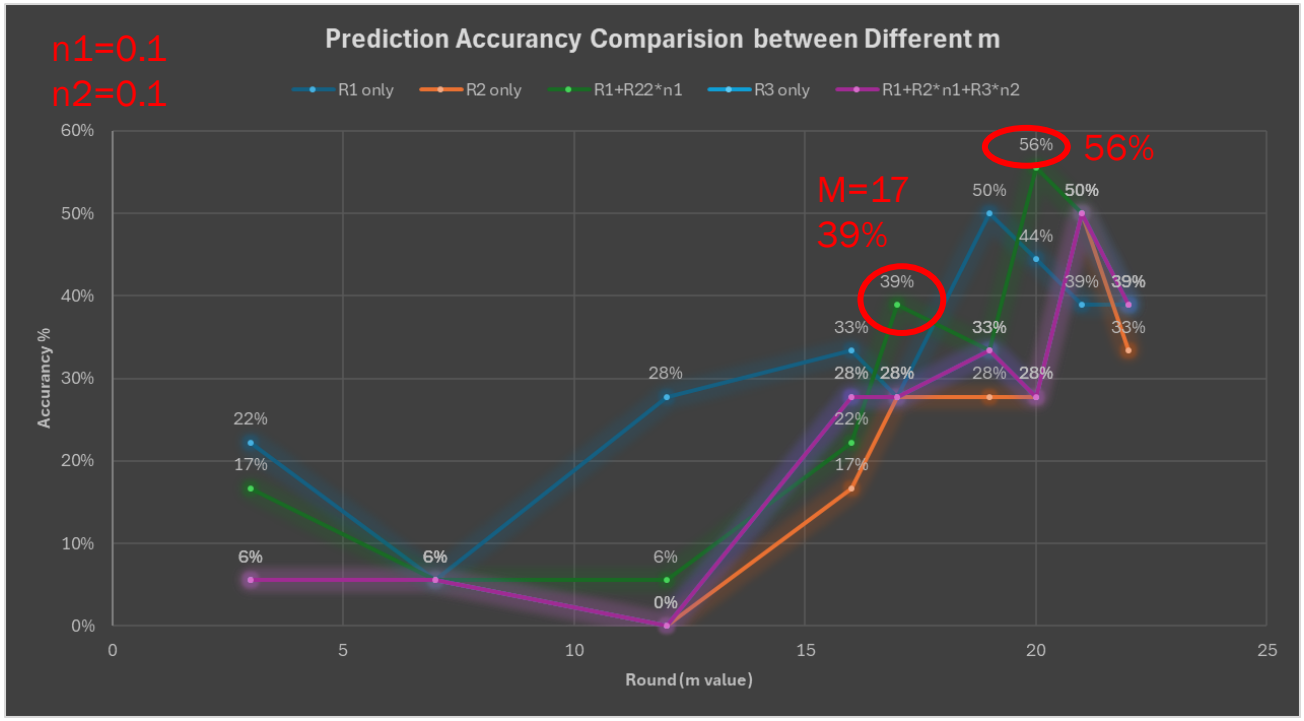


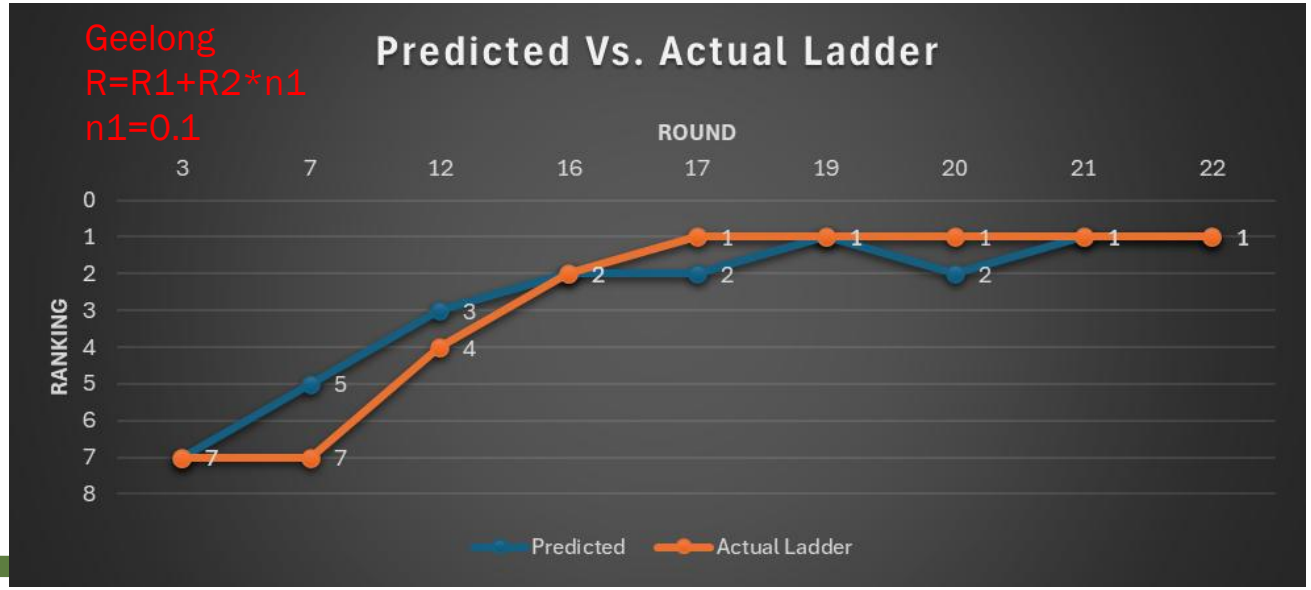
Data Analysis

Model Comparison

1. **R1 only Performs Well**
   - The R1 only model delivers strong predictions, performing similarly to R1+R2 and R1+R2+R3
2. **R2 is Less Accurate Than R1**
   - Predictions based on solely on R2 are not as reliable as those using R1
3. **Best Performing Model: R1+R2*n1**
   - This method achieved the highest accuracy (56% in Round 20),
   - This suggests that combing previous rounds with weighted adjustments improves prediction accuracy.
4. **Weighting Factor (n) Considerations**
   - Keep n small to ensure R1 has a great influence on the final prediction
5. **Limited Impact of R3**
   - R3 only and R1+R2*n1+R3*n2 show similar accuracy, indicating that R3 does not significantly enhance predictions.

Data Analysis





Reliability of the Dominance Matrix Method

1. **Can we trust this method?**
   - Yes. The dominance matrix provides reasonable predictions, especially in later rounds, but accuracy varies due to small score margins and unbalanced matchups.

2. **Prediction Accuracy**
   - Early rounds (**Rounds 3-7**) have low accuracy (**<22%**), but **50-56% accuracy** is achieved by **Round 16+**.
   - **R1+R2*n1** performs best, while **R3 adds little value**.

3. **Minimum Rounds for Reliable Predictions**
   - Predictions stabilize around **m ≈ 16-19** as team performance trends become clearer.

4. **Impact of Weighting**
   - Keeping **R1 dominant (small n1, n2)** improves accuracy. Adjusting weighting factors could further refine predictions.

5. **Future considerations:**
   - Experiment with combine both win-lose and score ratio together, with different weight adjustments to test accuracy.