MAAS' RATING SYSTEM

(last update: September 15, 2004)

Martien W. Maas

Nijmegen, The Netherlands

ABSTRACT

Maas' Rating System is based on the assumption that a win over a stronger team should be rewarded more than a win over a less stronger team. The system uses only wins as input, thus margin of victory (or home field advantage)is not taken into account. Computation of the rating scores is based on a set of equations that is solved by inverting a matrix and multiplying by a solution vector. Maas' Rating System produces Ranking Tables that highly correlate with Ranking Tables produced by other Rating Systems.

INTRODUCTION.

Already more than thirty years ago I felt a need for a Rating System that could do better than just counting the number of wins and losses. In my opinion a more sophisticated Rating System was needed, especially for unbalanced competitions, where due to draws and/or a knock-out system not all participating teams play each other and where not all participating teams even play the same number of matches. Inspired by the ELO-Rating System I developed a Rating System myself that took the strength of a team’s opponents into account and that appeared to have a number of properties similar to the ELO-Rating System.

In 1971 I published my first computer based Ranking Table in a Dutch Magazine (Goal) for European Champions Club Competitions.

Due to lack of computer facilities and problems in finding correct results I did not further maintain this Rating System.

Since the availability of powerful Personal Computers and Computer Languages and the accessibility to all information about Rating Systems via the Internet I became more and more interested again, especially after reading about College Football related Rating Systems. For this reason I like very much to share and discuss Maas’ Rating System wit other fans and ranking experts.

Hereafter I will describe the rational of Maas’ Rating System in detail and will provide a number of examples in order to show the validity of the system.

ASSUMPTIONS.

The assumptions of Maas’ Rating System are:

1. Winning is the only thing that matters. Margin of victory is not taken into account.
2. Wins against stronger opponents should be ”rewarded” more than wins against less stronger opponents (losses to stronger opponents should be ”punished” not as much as losses to less stronger opponents).

MODEL.

Maas' Rating System Model can be described as follows:

Each team's rating equals its number of wins, minus its number of losses, plus the sum of the ratings of its opponents, all divided by its number of opponents.

The above described model can be presented as formula:

$$R_i = \frac{W_i - L_i + \text{Sum}\ (R_{opp,i})}{M_i} \qquad (1)$$

Where: R_i = Rating of Team i
W_i = Wins of Team i
L_i = Losses of Team i
$R_{opp,i}$ = Rating of any Opponent of Team i
M_i = Total number of Matches of Team i

If a team plays an opponent two or more times, then I'm counting it as that many opponents.

Number of Wins is the only input parameter in the above presented formula. All other parameters can be derived from that Number of Wins and all R_i will be determined by the Rating System itself. Hence the first assumption: *"winning is the only thing"* is fulfilled.

Furthermore it can easily be derived from Formula (1) that with equal number of wins (and losses) a team's Rating depends on the strength of that team's Opponents: the higher the sum of the strength of all Opponents, the higher the team's Rating will be. And this fulfils the second assumption: "*wins against stronger opponents should be rewarded more than wins against less stronger opponents*".

COMPUTATION.

Unlike others I do not try to find a satisfactory solution by means of iterative calculations. My solution exists of the following procedure: I define a set of linear equations and determine all rating scores by inverting a matrix and multiplying by a solution vector.

For the description of the actual computation of Maas' Rating System I am pleased to refer to an example presented by Colley. The following table presents the input data set for the Colley example:

team	**A**	**B**	**C**	**D**	**E**	
A	.	.	W	L	L	**W indicates a Win**
B	.	.	L	.	W	**L indicates a Loss**
C	L	W	.	W	L	
D	W	.	L	.	.	
E	W	L	W	.	.	

For practical reasons I will refer to Rating(A) or R_a just with: **A** etc., etc
Team A played three matches (M), one against Team C, one against Team D and one against Team E.
Team A did win (W) one match and lost (L) two matches.

For Team A thus Formula (1) will result in:

A = [W – L + (C + D + E)] / M

A = [W – L + C + D + E] / 3

3 A = W - L + C + D + E

3 A – C – D – E = (W – L)

3 A – C – D – E = -1

In a similar way (as described for Team A on the previous page) equations for the other four teams can be derived which eventually will result in the following matrix equation:

3	0	-1	-1	-1		**R_a**		-1
0	2	-1	0	-1		**R_b**		0
-1	-1	4	-1	-1		**R_c**	=	0
-1	0	-1	2	0		**R_d**		0
-1	-1	-1	0	3		**R_e**		1

Since this is a so-called "singular" matrix no unique results can be obtained without any modifications.

1. A commonly used "trick" is to add 1 to the figures presented on the diagonal of the matrix, thus use 4, 3, 5, 3, 4 instead of 3, 2, 4, 2, 3 respectively. It can be done but it is a little bit subjective.

2. Another approach (in my opinion a better one) is to add a various number (0 < n <= 1) to the figures presented on the diagonal, related to the number of games per team (add 1 to diagonal value of the team that played the highest number of matches). This could be done by putting the following figures on the diagonal 3.75, 2.50, 5, 2.50 and 3.75 respectively.

3. My solution however is yet another one: I do assign a fix Rating Score to one of the (randomly chosen) teams and can thereafter solve the matrix equation in the usual way. In the "Example" on some of the next pages of this document I will demonstrate that the choice of the Team and the choice of the fixed Rating only affects the raw Ratings (which by doing so are rather meaningless statistics) but **not** the **rankings** nor the **relative differences** between the teams.

Before I will extensively present the results of an example based on a well defined set of matches, I like to stress the differences between the above presented matrix and vector when compared with the Colley Matrix Method:

- The values on the Colley Matrix diagonal are 2 higher than in my matrix (5, 4, 6, 4, and 5 respectively).
- The values in the Colley Vector consist of (W / L) where in my vector the values present (W – L).

For this small example the ranking according to both systems (Colley vs. Maas) is exactly the same while the correlation coefficient is calculated as: 0.9761.

EXAMPLE.

Assumptions and Model.

In this section I will present an example (including nine teams) in order to demonstrate the effect of a number of apaparently subjective choices of parameters. Furthermore I will describe the effects of additional input data on the rankings.

The input data set for the first Table of this example is presented in Appendix A.

Nine teams played between 19 and 38 matches.
Teams B and H did play the same number of matches against the same opponents (not against each other) with the same amount of success.
Team I played two times as much matches as Team A did, against the same opponents (not against each other) with exactly the same relative amount of success.

Table 1 in Appendix B presents the Ranking Table based on the input data set that is presented in Appendix A.
As can be expected Teams B and H have equal rating scores in Table 1. Furthermore, despite different number of matches played, Team I and Team A have the same rating scores which is in line with the described assumptions and model formula (same relative Win – Loss record with same strength of schedule).

After Ranking Table 1 was computed, the following match results were added to the original schedule as presented in Appendix A:

- Team B did win 2 matches against a higher rated team (Team C)
- Team H did win 2 matches against a lower rated team (Team G)

Table 2 in Appendix B presents the Ranking Table based on the original input data set plus four more match results.
Although Team B has the same number of matches as Team H and the same Win – Loss record, Team B has a higher rating score in Table 2 when compared to Team H. This result is in line with the described assumptions and model formula (same Win – Loss record but higher strength of schedule).

After Ranking Table 1 was computed, the following match results were added to the original schedule as presented in Appendix A:

- Team B did loss 2 matches against a higher rated team (Team C)
- Team H did loss 2 matches against a lower rated team (Team G)

Table 3 in Appendix B presents the Ranking Table based on the original input data set plus four more match results. Although Team B has the same number of matches as Team H and the same Win – Loss record, Team B has a higher rating score in Table 3 when compared to Team H. This result is in line with the described assumptions and model formula (same Win – Loss record but higher strength of schedule).

Finally I'll check the outcome of a Team's Rating Score by substitution of the team's record in Formula (1). For example: the Rating Score of Team F (Table 1 in Appendix B) appears to be: -0.116.
Team F's record is 13 Wins and 12 Losses while the mean Rating Score of its Opponents is –0.1506 (Table 1 in Appendix B).

Substitution of these values in Formula (1) reveals:

$$R_f = \frac{W_f - L_f + \mathrm{Sum}\ (R_{opp,f})}{M_f}$$

$\mathrm{Sum}\ (R_{opp,f})$ is obtained by multiplying the mean rating of Team F's opponents by the number of matches that Team F played: -0.1506 * 25 = -3.7650

$R_f =$ [(13 – 12 –3.7650) / 25] = -2.7650 / 25 = -0.1106

This result (-0.1106) is exactly the Rating Score for Team F in Table 1 of Appendix B which demonstrates that computation of the Rating Scores by means of solving Matrix Equations reveal results in line with the Model's Formula.

Parameters for computation.

From Formula (1) it can be derived that the weight of each win is (implicitly) set to 1, in other words: a team will gain **1 point** for each win (and will loss 1 point for each loss). The question can be asked why just 1 point? why not 3? or 10? or 25? etc., etc. What will be the effect on the final Rating Scores of this parameter setting?

In order to "avoid" the problem of matrix singularity two further choices were to be made:

- The first choice is related to **which of the Teams** is chosen to get a fixed Rating Score assigned (Team A? Team B? the team with the highest or lowest relative score? etc., etc.).
- The second choice is the **value of the fixed Rating Score** (0? 0.5? 1? 10? etc., etc.) that is attributed to one of the participating teams.

Since I am not a mathematician I cannot provide official mathematical "proves" for the conclusion that I will draw based on the calculations that I have performed. However, I am confident that such "prove" easily can be found by any interested mathematician.

By calculating the ratings for combinations of various values for the three parameters described above I came to the following conclusions:

- The number of points to be rewarded for a Win should be positive and larger than zero (>0). By the substitution of another value (than the default value of 1 point) for a Win, all **differences** between all Teams' Rating Scores did increase or decrease with the **same factor**, and hence the Ranking of the teams did not change.
  The conclusion that can be drawn is that the Ranking of the teams does not depend on the number of points that is gained for each Win.
- Without further comment I can report that all computations (with different teams that were assigned a fixed Rating Score) revealed similar results.
  The conclusion that can be drawn is that the Ranking of the teams does not dependent on the team that has to be chosen to be set to a fixed Rating Score.
- Computations with different Fixed Rating Scores for one of the (randomly chosen) teams showed that the **differences** between the Teams' Rating Scores did not change at all. All Rating Scores did increase or decrease with the **same constant**, and hence the Ranking of the teams did not change.

The conclusion that can be drawn is that the Ranking of the teams does not depend on the Fixed Rating Score that has to be assigned to one of the (randomly chosen) teams.

RANKING TABLE.

In the previous Sections I have described Maas' Rating System and how it works out for a well formulated example.

As mentioned before the **scale** of the original obtained raw Rating scores from Maas' Rating System is **arbitrary**: the only thing that really matters is the **Ranking Order**.

For the sake of readability I usually present Ranking Tables in which I linearly transform (rescale) the originally obtained Rating scores depending on the type of application (USA College Football, UEFA (European Football) Champions League etc.). The **presented** Rating Scores fit into well defined Ranges (e.g. from 100 to 1 or from 1 to 0) or have well defined other statistics (e.g. Mean is 0 or 0.5).

Note that this type of transformation procedures **does not change the order** in which teams originally are ranked.

DISCUSSION.

In Appendix C of this document I present a Ranking Table for **2002** College Football, Division I-A based on Maas' Rating System. Computation of correlation coefficients between Maas' Ranking and the so-called Consensus number (Massey Ratings) and other important rankings revealed values all above 0.99. The correlation coefficient with the so-called consensus was 0.99126. The so-called Ranking Violation Percentage (Massey Ratings) for Maas' Ranking Table presented in Appendix C appears to be 126 out of 772 (for the 2002 Season) which is 16.32%.

Based on the examples and the actual College Football Ranking Table that I have provided I am confident that Maas' Rating System is as effective as the "best" ones presented by Massey Ratings.

The Model of Maas’ Rating System uses a minimum of assumptions and is based on common sense: the ”reward” for a Win depends on the difference in strength between the two opponents. Maas’ Rating System adjusts for strength of schedule: a Win against a stronger team is rewarded more than a Win against a weaker team.

Like a few others I do claim that Maas’ Rating System has no subjective inputs (and hence no bias) at all. The only input contains of number of Wins for Team X versus Team Y.

Like for most other Rating Systems, the absolute value of the Rating is not important. These values are only used to determine the Ranking Order. The resulting Ranking Table for the USA College Football (Season 2002, Division I-A) shows that Maas’ Rating System produces common sense results.

COPYRIGHT STATEMENT.

DISCLAIMER.

Ranking Tables presented in this Document should not be used to predict the winners of future games.
The reader accepts complete responsibility for his/her actions based on the content of this Document.

Martien W. Maas, Nijmegen, The Netherlands, February 2003.

ACKNOWLEDGEMENTS AND REFERENCES.

I am grateful to Todd Beck, Wesley Colley, Kenneth Massey, David Wilson, John Wobus and Peter Wolfe.

I have frequently visited their websites and at some occasions did send them questions in order to get information about (College Football) Rating Systems.

[1] Todd Beck At http://tbeck.freeshell.org/
[2] Wesley Colley At www.colleyrankings.com
[3] Kenneth Massey At www.masseyratings.com/cf/compare.htm
[4] David Wilson At http://homepages.cae.wisc.edu/~dwilson/rsfc/rate/
[5] John Wobus At www.vaporia.com/sports/
[6] Peter Wolfe At www.bol.ucla.edu/~prwolfe/cfootball/

Appendix A: Data for the calculation of a sample rating list

Number of wins for all participating teams.

	A	B	C	D	E	F	G	H	I
A	.	1	2	3	2	1	0	1	0
B	2	.	0	2	2	2	2	0	4
C	3	3	.	3	3	3	3	3	6
D	0	3	0	.	1	1	4	3	0
E	1	3	0	1	.	0	3	3	2
F	0	3	0	2	1	.	4	3	0
G	1	1	1	1	1	1	.	1	2
H	2	0	0	2	2	2	2	.	4
I	0	2	4	6	4	2	0	2	.

Example: Team A wins two games against Team E.
Team E wins one game against Team A.

Note 1: Team B and Team H have exactly the same results against the same opponents.

Note 2: Team I has exactly the same win percentage against the same opponents as team A, be it that Team I played twice as much matches as Team A.

Appendix B: Results of some examples.

Table 1: Ranking Table according to Maas' Rating System.

Note: Teams B and H have same results against same opponents.
Note: Teams A and I have same results against same opponents, be it that Team I played two times as much matches as Team A.

Rank	Team Name	Matches	Wins	Losses	Percent. Wins	Average Opponent	Raw Rating	Transformed Rating
1	C	34	27	7	79.4	-0.1391	0.4491	10.00
2	A	19	10	9	52.6	-0.0526	0.0000	5.36
3	I	38	20	18	52.6	-0.0526	0.0000	5.36
4	F	25	13	12	52.0	-0.1506	-0.1106	4.21
5	B	30	14	16	46.7	-0.1081	-0.1748	3.55
6	H	30	14	16	46.7	-0.1081	-0.1748	3.55
7	E	29	13	16	44.8	-0.1000	-0.2034	3.25
8	D	32	12	20	37.5	-0.1014	-0.3514	1.72
9	G	27	9	18	33.3	-0.0880	-0.4214	1.00

Table 2: Ranking Table according to Maas' Rating System.

Changes in comparison to Table 1:
Team B did win 2 matches against Team C.
Team H did win 2 matches against Team G.

Rank	Team Name	Matches	Wins	Losses	Percent. Wins	Average Opponent	Raw Rating	Transformed Rating
1	C	36	27	9	75.0	-0.1285	0.3715	10.00
2	A	19	10	9	52.6	-0.0526	0.0000	6.00
3	I	38	20	18	52.6	-0.0526	0.0000	6.00
4	B	32	16	16	50.0	-0.0858	-0.0858	5.08
5	F	25	13	12	52.0	-0.1423	-0.1023	4.90
6	H	32	16	16	50.0	-0.1381	-0.1381	4.52
7	E	29	13	16	44.8	-0.0916	-0.1950	3.90
8	D	32	12	20	37.5	-0.0945	-0.3445	2.29
9	G	29	9	20	31.0	-0.0854	-0.4647	1.00

Table 3: Ranking Table according to Maas' Rating System.

Changes in comparison to Table 1:
Team B did loss 2 matches against Team C.
Team H did loss 2 matches against Team G.

Rank	Team Name	Matches	Wins	Losses	Percent. Wins	Average Opponent	Raw Rating	Transformed Rating
1	C	36	29	7	80.6	-0.1382	0.4729	10.00
2	A	19	10	9	52.6	-0.0526	0.0000	4.81
3	I	38	20	18	52.6	-0.0526	0.0000	4.81
4	F	25	13	12	52.0	-0.1460	-0.1060	3.65
5	B	32	14	18	43.8	-0.0607	-0.1857	2.78
6	E	29	13	16	44.8	-0.0987	-0.2022	2.60
7	H	32	14	18	43.8	-0.1116	-0.2366	2.22
8	G	29	11	18	37.9	-0.1009	-0.3423	1.06
9	D	32	12	20	37.5	-0.0977	-0.3477	1.00

Appendix C: **Table 4: Ranking Table Using Maas' Rating System.**

Rank	Team	Wins	Losses	Teams rating	Oppon. rating
1	Ohio St	14	0	1.000	0.587
2	USC	11	2	0.962	0.676
3	Georgia	13	1	0.932	0.578
4	Miami FL	12	1	0.915	0.566
5	Oklahoma	12	2	0.870	0.576
6	Michigan	10	3	0.858	0.636
7	Texas	11	2	0.851	0.565
8	Iowa	11	2	0.845	0.559
9	Notre Dame	10	3	0.816	0.594
10	Alabama	10	3	0.798	0.575
11	Washington St	10	3	0.787	0.565
12	Kansas St	11	2	0.772	0.486
13	Florida St	9	5	0.770	0.653
14	Maryland	11	3	0.768	0.532
15	Penn St	9	4	0.752	0.594
16	NC State	11	3	0.744	0.508
17	Colorado	9	5	0.738	0.620
18	Texas Tech	9	5	0.731	0.613
19	Auburn	9	4	0.728	0.570
20	Virginia Tech	10	4	0.718	0.542
21	Virginia	9	5	0.717	0.599
22	Florida	8	5	0.712	0.617
23	Pittsburgh	9	4	0.703	0.544
24	Arkansas	9	5	0.702	0.584
25	West Virginia	9	4	0.700	0.541
26	South Florida	9	2	0.696	0.433
27	Boise St	12	1	0.690	0.341
28	UCLA	8	5	0.689	0.594
29	Wisconsin	8	6	0.680	0.621
30	Boston College	9	4	0.663	0.504
31	LSU	8	5	0.659	0.564
32	Colorado St	10	4	0.657	0.480
33	TCU	10	2	0.644	0.369
34	Tennessee	8	5	0.640	0.545
35	California	7	5	0.636	0.567
36	Oregon St	8	5	0.633	0.538
37	Marshall	11	2	0.630	0.344
38	Minnesota	8	5	0.625	0.530
39	Clemson	7	6	0.625	0.593
40	Washington	7	6	0.604	0.572
41	Oklahoma St	8	5	0.601	0.506
42	Georgia Tech	7	6	0.601	0.569
43	Arizona St	8	6	0.600	0.541
44	Purdue	7	6	0.599	0.567
45	Iowa St	7	7	0.586	0.586
46	Wake Forest	7	6	0.576	0.545
47	Mississippi	7	6	0.576	0.545
48	Toledo	9	5	0.570	0.452
49	Kentucky	7	5	0.563	0.494
50	Fresno St	9	5	0.559	0.441
51	N Illinois	8	4	0.558	0.421
52	Bowling Green	9	3	0.558	0.352
53	Oregon	7	6	0.554	0.522
54	Texas A&M	6	6	0.543	0.543
55	Hawaii	10	4	0.542	0.365
56	Nebraska	7	7	0.524	0.524
57	North Texas	8	5	0.521	0.426
58	Miami OH	7	5	0.516	0.448
59	South Carolina	5	7	0.514	0.583
60	Air Force	8	5	0.508	0.413

Appendix C: **Table 4: Ranking Table Using Maas' Rating System.**

Rank	Team	Wins	Losses	Teams rating	Oppon. rating
61	Illinois	5	7	0.501	0.570
62	Cincinnati	7	7	0.490	0.490
63	Missouri	5	7	0.484	0.553
64	Louisville	7	6	0.466	0.434
65	Southern Miss	7	6	0.462	0.430
66	UCF	7	5	0.455	0.386
67	Tulane	8	5	0.450	0.355
68	New Mexico St	7	5	0.446	0.377
69	New Mexico	7	7	0.437	0.437
70	Michigan St	4	8	0.437	0.575
71	Arizona	4	8	0.413	0.550
72	North Carolina	3	9	0.412	0.619
73	San Jose St	6	7	0.397	0.429
74	Temple	4	8	0.395	0.533
75	Ball St	6	6	0.391	0.391
76	Utah	5	6	0.390	0.428
77	Syracuse	4	8	0.390	0.527
78	Stanford	2	9	0.387	0.650
79	Connecticut	6	6	0.383	0.383
80	UNLV	5	7	0.370	0.439
81	Nevada	5	7	0.362	0.431
82	Brigham Young	5	7	0.351	0.420
83	Northwestern	3	9	0.349	0.556
84	Baylor	3	9	0.339	0.546
85	Houston	5	7	0.336	0.405
86	Indiana	3	9	0.335	0.541
87	Mississippi St	3	9	0.317	0.523
88	San Diego St	4	9	0.314	0.473
89	East Carolina	4	8	0.312	0.450
90	UAB	5	7	0.312	0.380
91	W Michigan	4	8	0.303	0.440
92	Arkansas St	6	7	0.294	0.326
93	Utah St	4	7	0.293	0.406
94	Vanderbilt	2	10	0.292	0.567
95	Louisiana Tech	4	8	0.288	0.426
96	Akron	4	8	0.276	0.414
97	Duke	2	10	0.266	0.541
98	C Michigan	4	8	0.259	0.396
99	Ohio	4	8	0.258	0.396
100	Middle Tenn St	4	8	0.238	0.376
101	Rice	4	7	0.226	0.338
102	Troy St	4	8	0.224	0.361
103	LA Lafayette	3	9	0.211	0.417
104	Kent	3	9	0.202	0.408
105	Memphis	3	9	0.199	0.406
106	Kansas	2	10	0.199	0.474
107	LA Monroe	3	9	0.193	0.399
108	SMU	3	9	0.183	0.389
109	E Michigan	3	9	0.179	0.385
110	Wyoming	2	10	0.170	0.445
111	Rutgers	1	11	0.163	0.507
112	Navy	2	10	0.155	0.430
113	Idaho	2	10	0.139	0.415
114	Texas-El Paso	2	10	0.120	0.395
115	Tulsa	1	11	0.048	0.392
116	Buffalo	1	11	0.016	0.360
117	Army	1	11	0.000	0.344

